Lattice computation of the electromagnetic contributions to kaon and pion masses

S. Basak

School of Physical Sciences, NISER Bhubaneswar, Orissa 752050, India

A. Bazavov

Department of Computational Mathematics, Science and Engineering and Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA

C. Bernard[*]

Department of Physics, Washington University, St. Louis, Missouri 63130, USA

C. DeTar and L. Levkova

Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah 84112, USA

E. Freeland

Liberal Arts Department, School of the Art Institute of Chicago, Chicago, Illinois 60603, USA

Steven Gottlieb[†] and A. Torok[‡]

Department of Physics, Indiana University, Bloomington, Indiana 47405, USA

U. M. Heller

American Physical Society, One Research Road, Ridge, New York 11961, USA

J. Laiho

Department of Physics, Syracuse University, Syracuse, New York 13244, USA

J. Osborn

ALCF, Argonne National Laboratory, Argonne, Illinois 60439, USA

R. L. Sugar

Physics Department, University of California, Santa Barbara, California 93106, USA

D. Toussaint

Physics Department, University of Arizona Tucson, Arizona 85721, USA

R. S. Van de Water and R. Zhou

Theoretical Physics Department, Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA

(MILC Collaboration)

We present a lattice calculation of the electromagnetic (EM) effects on the masses of light pseudoscalar mesons. The simulations employ $2 + 1$ dynamical flavors of asqtad QCD quarks and quenched photons. Lattice spacings vary from ≈ 0.12 fm to ≈ 0.045 fm. We compute the quantity ϵ, which parametrizes the corrections to Dashen's theorem for the $K^+ - K^0$ EM mass splitting, as well as ϵ_{K^0}, which parametrizes the

[*]cb@wustl.edu
[†]sg@indiana.edu
[‡]Present address: ThermoFisher Scientific, Hillsboro, Oregon 97124, USA.

EM contribution to the mass of the K^0 itself. An extension of the nonperturbative EM renormalization scheme introduced by the BMW group is used in separating EM effects from isospin-violating quark mass effects. We correct for leading finite-volume effects in our realization of lattice electrodynamics in chiral perturbation theory, and remaining finite-volume errors are relatively small. While electroquenched effects are under control for ϵ, they are estimated only qualitatively for ϵ_{K^0} and constitute one of the largest sources of uncertainty for that quantity. We find $\epsilon = 0.78(1)_{\text{stat}}\binom{+8}{-11}_{\text{syst}}$ and $\epsilon_{K^0} = 0.035(3)_{\text{stat}}(20)_{\text{syst}}$. We then use these results on $2 + 1 + 1$ flavor pure QCD highly improved staggered quark (HISQ) ensembles and find $m_u/m_d = 0.4529(48)_{\text{stat}}\binom{+150}{-67}_{\text{syst}}$.

I. INTRODUCTION

The mass splitting between the charged and neutral kaons, $K^\pm$ and K^0, arises from two effects that give comparable contributions: the mass difference between up and down quarks, and electromagnetism. If the electromagnetic (EM) contributions can be determined and removed from the experimental meson masses, the resulting pure-QCD masses can then be used as input to a lattice QCD calculation to determine the light quark masses, and in particular the ratio m_u/m_d, a fundamental parameter of the standard model which measures the strength of strong isospin violations.

The size of the EM contributions to the $K^\pm$–K^0 mass splitting is a long-standing issue. Almost fifty years ago, Dashen [1] showed that the EM splitting of the charged and neutral kaons is equal to that of the pions in leading order (LO) of chiral $\text{SU}(3) \times \text{SU}(3)$ symmetry. In other words, at LO, $(M_{K^\pm}^2 - M_{K^0}^2)^\gamma = (M_{\pi^\pm}^2 - M_{\pi^0}^2)^\gamma$, where the superscript γ denotes the EM contribution, i.e., the difference between the quantity in the real world and in a world where all quark charges are set to zero (keeping renormalized quark masses unchanged). However, it has been known for some time that the corrections to this lowest-order result are large; see, e.g., Ref. [2] for a pedagogical review. These corrections can be estimated in a variety of continuum phenomenological models [3]. The model results differ considerably, however, and do not allow one to make controlled estimates of the systematic errors. Indeed, in lattice determinations of m_u/m_d that employ phenomenological estimates of EM contributions [4–7], the error coming from the range of EM estimates dominates all other systematic errors.

Direct lattice calculations of the EM contribution to the kaon splittings can greatly reduce the uncertainties. This approach was pioneered by Duncan, Eichten, and Thacker [8] in the quenched approximation of QCD, and has been applied in full QCD more recently by several groups [9–17]. Here we report on our lattice QCD + QED computation of $(M_{K^\pm}^2 - M_{K^0}^2)^\gamma$. We then apply our result to compute m_u/m_d in a pure QCD simulation.

There is an alternative approach to calculating EM effects on the lattice [18,19] in which one expands out QED and isospin-violating interactions to $\mathcal{O}(\alpha_{\text{EM}}, m_u - m_d)$ (where α_{EM} is the fine structure constant) and then

computes the resulting matrix elements in isospin-conserving pure QCD. We do not discuss this approach further here, but simply note that the existence of two independent methods makes possible important cross checks on the results and errors of both. See Ref. [20] for a review that covers both approaches.

In lattice simulations of QCD + QED, both the QCD and QED should in principle be *unquenched*, i.e., include all contributions from virtual sea-quark loops. However, Bijnens and Danielsson [21] have shown that QED quenching effects for mass differences such as $(M_{K^\pm}^2 - M_{K^0}^2)^\gamma$ are computable through next-to-leading order (NLO) in $\text{SU}(3) \times \text{SU}(3)$ chiral perturbation theory, with no dependence on unknown low energy constants (LECs). In other words, the sea quarks may be taken to be electrically neutral in the simulation, and the effects of their charges may be restored, correct to NLO, after the fact. We take advantage of this result here and simulate full, unquenched QCD + quenched QED (the electroquenched approximation) in order to determine the kaon EM splittings. Since the QED part of the simulation is quenched, we need only to calculate valence-quark propagators in a background consisting of pure unquenched QCD and quenched EM fields, which are free fields and therefore easily generated. For the pure QCD backgrounds, we use our large data set of ensembles generated with $2 + 1$ flavors of asqtad staggered quarks [4]. We have added a number of additional ensembles to better study finite-volume effects.

One may parametrize the kaon EM splitting by [2]

$$\epsilon \equiv \frac{(M_{K^\pm}^2 - M_{K^0}^2)^\gamma - (M_{\pi^\pm}^2 - M_{\pi^0}^2)^\gamma}{(M_{\pi^\pm}^2 - M_{\pi^0}^2)^{\text{expt}}}, \qquad (1)$$

where the experimental pion splitting is used in the denominator, rather than the EM pion splitting. The two are equal up to isospin-violating effects, which are $\mathcal{O}((m_u - m_d)^2)$, and therefore small. Determining the EM contribution to the mass of the true π^0 is costly, however, since it has quark-line disconnected EM diagrams even in the isospin limit. Instead, we drop the disconnected diagrams, which are expected to be small, and simply find the RMS average mass of $u\bar{u}$ and $d\bar{d}$ mesons. We call the pion obtained in this manner the "π^0." Both the true $(M_{\pi^0}^2)^\gamma$ and our $(M_{"\pi^0"}^2)^\gamma$ are small

because EM contributions to neutral mesons vanish in the chiral limit. For the true π^0, this is required by Dashen's arguments [1], and may be seen explicitly in chiral perturbation theory (χPT) including EM effects [22]. For the "π^0," a simple argument in partially quenched χPT, given below in Sec. III D, shows that $(M^2_{``\pi^{0}"})^\gamma$ also vanishes in the chiral limit. This means that the disconnected EM contributions that we are neglecting are themselves small. (An alternative diagrammatic proof of the small size of the disconnected terms has been given previously in Ref. [18].) Further, Zweig's rule suggests that the mass contribution from the disconnected diagrams is in fact still smaller than either $(M^2_{\pi^0})^\gamma$ or $(M^2_{``\pi^{0}"})^\gamma$ separately.

Summarizing, we use

$$\epsilon \cong \frac{(M^2_{K^\pm} - M^2_{K^0})^\gamma - (M^2_{\pi^\pm} - M^2_{``\pi^{0}"})^\gamma}{(M^2_{\pi^\pm} - M^2_{\pi^0})^{\text{expt}}} \qquad (2)$$

to compute ϵ. The systematic error coming from using the "π^0" will of course need to be estimated.

An alternative estimate of ϵ is also possible if we employ the experimental EM pion splitting in the numerator of Eq. (1) instead of our computed π splitting. This estimate is then independent of any assumptions about the disconnected diagrams in the "π^0." For a test of systematic effects in the calculation of ϵ, we can therefore look at

$$\epsilon' \equiv \frac{(M^2_{K^\pm} - M^2_{K^0})^\gamma - (M^2_{\pi^\pm} - M^2_{\pi^0})^{\text{expt}}}{(M^2_{\pi^\pm} - M^2_{\pi^0})^{\text{expt}}}, \qquad (3)$$

In Ref. [2], the contribution to the pion splitting coming from quark masses (i.e., the splitting that would be present in QCD alone) is defined to be $\epsilon_m (M^2_{\pi^\pm} - M^2_{\pi^0})^{\text{expt}}$. Then Eqs. (1) and (3) imply

$$\epsilon' = \epsilon - \epsilon_m. \qquad (4)$$

At NLO in χPT, $\epsilon_m = 0.04$ [23]. Reference [2] adds a conservative error and quotes $\epsilon_m = 0.04(2)$. In our calculation, $\epsilon - \epsilon'$ appears to be positive. However, because our systematic errors in both ϵ and ϵ' are significantly larger than 0.04, we are only able to use the difference $\epsilon - (\epsilon' + 0.04)$ as one estimate of those errors, and have nothing to report about ϵ_m itself.

We also calculate the EM contribution to the squared mass of the neutral kaon, $(M^2_{K^0})^\gamma$. It is convenient to express this quantity in terms of the experimental pion splitting, just as we have done for the kaon splitting. We follow Ref. [2] and define the dimensionless quantity ϵ_{K^0} by

$$\epsilon_{K^0} \equiv \frac{(M^2_{K^0})^\gamma}{(M^2_{\pi^+} - M^2_{\pi^0})^{\text{expt}}}. \qquad (5)$$

The following is an outline of the remainder of the paper: Section II gives the details of the $2+1$ flavor asqtad staggered QCD ensembles [4] on which we compute

(quenched) EM effects. In addition, we describe the pure QCD $2+1+1$ highly improved staggered quark (HISQ) ensembles [24] on which we calculate m_u/m_d, with input on EM effects from the asqtad simulations. We discuss infinite volume χPT in QCD + QED in Sec. III. Modifications for partial quenching [21] and staggered discretization errors [25] are detailed, and the staggered result for the meson masses at NLO is presented and explained. In Sec. IV, we describe how we define QED in finite volume (FV). Finite-volume effects are then calculated at one loop in staggered χPT in Sec. V. We show that the resulting formulas give an excellent description of our lattice data over a wide range of volumes. We can therefore correct for FV effects, with a small residual systematic error. Section VI C then presents a variety of chiral fits to the FV-corrected lattice data, and Sec. VII describes our results and systematic errors for the EM contributions to the kaon massess, and the parameters ϵ and ϵ_{K^0}. Finally, in Sec. VIII, we use our EM results are to adjust the experimental kaon masses to their values in a pure-QCD world, which are then taken as input to the calculation of m_u/m_d following Ref. [26].

Our final results are

$$\epsilon = 0.78(1)_{\text{stat}} \binom{+8}{-11}_{\text{syst}},$$

$$\epsilon_{K^0} = 0.035(3)_{\text{stat}}(20)_{\text{syst}},$$

$$m_u/m_d = 0.4529(48)_{\text{stat}} \binom{+150}{-67}_{\text{syst}}.$$

Preliminary versions of this work have appeared in Refs. [12–15,27].

We note that m_u/m_d may be computed on the lattice in other ways that do not depend on knowing the EM contributions to the kaon masses. In particular, Ref. [28] uses a dispersive treatment of the experimental input from the decay $\rho \to 3\pi$ instead of kaon splittings to obtain the ratio m_u/m_d from their lattice determination of m_s/m_l, where $m_l \equiv (m_u + m_d)/2$. Since the $\rho \to 3\pi$ decay violates isospin but is known to be fairly independent of EM corrections, it gives a handle on m_u/m_d that does not require EM input, at least to some level of accuracy.

II. LATTICE DETAILS

We calculate meson masses on the $(2+1)$-flavor MILC asqtad ensembles, with quenched photon fields, and with lattice spacings ranging from ≈ 0.12 fm to ≈ 0.045 fm. Table I shows the ensembles employed. On all ensembles, we generate propagators for valence quarks that have charges 0, $\pm 1/3e$, or $\pm 2/3e$, where $e \approx 0.303$ is the physical electron charge, and we compute the masses of mesons made from various combinations of these quarks. On many ensembles we also have mesons made from quarks with charges greater than physical: $\pm e$ and $\pm 4/3e$. On some ensembles, we even have quarks with charges

TABLE I.   Parameters of the $(2 + 1)$-flavor asqtad ensembles used in this study. The quark masses m'_l and m'_s are the light and strange dynamical masses used in the runs. The number of configurations listed as '132 + 52' for the $a \approx 0.12$ fm, $48^3 \times 64$ ensemble gives values for two independent streams, the first in single precision, and the second in double. We treat them as separate data, and do not average the results. The r_1/a values are mass-independent, in that they are extrapolated to physical quark masses, rather than the sea mass of the simulations. The errors listed for r_1/a are the sum in quadrature of the statistical errors and the extrapolation errors. We use the $a \approx 0.12$, $m'_l = 0.01$, $m'_s = 0.05$ result for r_1/a for those $a \approx 0.12$ ensembles where no r_1/a value has been directly computed.

$\approx a$ [fm]	Volume	β	m'_l/m'_s	# Configs.	L (fm)	$m_\pi L$	r_1/a
0.12	$12^3 \times 64$	6.76	0.01/0.05	1000	1.4	2.7	$\cdots$
	$16^3 \times 64$	6.76	0.01/0.05	1303	1.8	3.6	$\cdots$
	$20^3 \times 64$	6.76	0.01/0.05	2254	2.3	4.5	2.739(12)
	$28^3 \times 64$	6.76	0.01/0.05	274	3.2	6.3	$\cdots$
	$40^3 \times 64$	6.76	0.01/0.05	115	4.6	9.0	$\cdots$
	$48^3 \times 64$	6.76	0.01/0.05	$132 + 52$	5.5	10.8	$\cdots$
	$20^3 \times 64$	6.76	0.007/0.05	1261	2.3	3.8	2.739(13)
	$24^3 \times 64$	6.76	0.005/0.05	2099	2.7	3.8	2.739(13)
0.09	$28^3 \times 96$	7.09	0.0062/0.031	1930	2.3	4.1	3.789(6)
	$40^3 \times 96$	7.08	0.0031/0.031	1015	3.3	4.2	3.755(6)
0.06	$48^3 \times 144$	7.47	0.0036/0.018	670	2.8	4.5	5.353(12)
	$56^3 \times 144$	7.465	0.0025/0.018	798	3.3	4.4	5.330(12)
	$64^3 \times 144$	7.46	0.0018/0.018	826	3.8	4.3	5.307(12)
0.045	$64^3 \times 192$	7.81	0.0028/0.014	801	2.8	4.6	7.208(25)

$\pm 2e$, although charges that high are not included in the analysis at this time.

Quenched photon fields are generated in momentum space in the finite-volume Coulomb gauge QED_{TL} defined in detail in Sec. IV. The momentum-space distribution is Gaussian, and is generated and Fourier transformed to position space by a serial program. The spectrum program reads the photon fields from disk and, for each desired charge, converts the field to a $U(1)$ phase factor with that charge. The $SU(3)$ links are multiplied by the $U(1)$ links, and then the same gauge smearing that we use for $SU(3)$ alone is applied. This amounts to an a^2-improved action, but without any tadpole improvement of $U(1)$.

A. New ensembles

To study finite-volume errors, which were found to be quite important in our prior work, we have generated a number of new ensembles that are not detailed in Ref. [4]. Our prior finite-volume work used two volumes corresponding to spatial size $L = 20$ and 28. We have added $L = 12$, 16, 40, and 48 in order to have data on both larger and smaller volumes. For $L = 12$, we have generated the ensemble using the R algorithm [29–31] in a single stream of 5200 time units of evolution. Each trajectory consists of 150 steps with a step size of 0.00667. The first 200 time units are dropped and every 5th time unit is then archived for analysis, yielding 1000 configurations in the ensemble. For $L = 16$, we have four separate streams. Three of them use the RHMC [32–36] algorithm with a 3G1F Omelyan integrator [37,38]. The step size is 0.05, and there are twenty steps per trajectory. Each of these streams has 334

or 335 configurations separated by 6 time units. A fourth stream employs the R algorithm with the same parameters as for $L = 12$ and has 300 configurations separated by 5 time units. For $L = 40$, we use the RHMC algorithm with 40 steps of size 0.025 and analyze 115 configurations separated by 6 time units. All of the above ensembles are generated by single-precision code, except that accumulations are done in double precision. For $L = 48$, we use two streams, one in single precision (as above), and one in double precision. In each case, archived configurations are separated by 6 time units. From the single-precision ensemble, 132 configurations are used for the spectrum analysis, whereas 52 are analyzed from the double-precision ensemble. These have not been combined in the finite-volume study. Table II summarizes information about the new ensembles.

TABLE II.   Characteristics of the new ensembles generated to study finite volume effects. Each ensemble has a volume of $L^3 \times 64$ with the value of L in the first column. The second column indicates the algorithm used to generate the ensemble. The third and fourth columns contain the molecular dynamics step size and the number of steps in each trajectory, respectively. The fifth column indicates how many trajectories separate archived lattices on which the spectrum analysis is done. The last column contains additional comments.

L	Algorithm	δt	Steps	Trajectories	Comment
12	R	0.00667	150	5	
16	R	0.00667	150	5	
16	RHMC	0.05	20	6	3 streams
40	RHMC	0.025	40	6	
48	RHMC	0.025	40	6	Single precision
48	RHMC	0.025	40	6	Double precision

TABLE III.   Details of the charges and valence quark masses used for the meson spectrum. The last column indicates how many charge and mass combinations were used to construct mesons.

$\approx a$ [fm]	Volume	m_l'/m_s'	Charges	am_v	Channels
0.12	$12^3 \times 64$	0.01/0.05	$\pm 2/3, \pm 1/3, 0$	0.005, 0.007, 0.01, 0.02, 0.03, 0.04, 0.05	700
	$16^3 \times 64$	0.01/0.05	$\pm 2/3, \pm 1/3, 0$	0.005, 0.007, 0.01, 0.02, 0.03, 0.04, 0.05	700
	$20^3 \times 64$	0.01/0.05	$\pm 2, \pm 4/3, \pm 1, \pm 2/3, \pm 1/3, 0$	0.005, 0.007, 0.01, 0.02, 0.03, 0.04, 0.05	532
	$28^3 \times 64$	0.01/0.05	$\pm 2, \pm 4/3, \pm 1, \pm 2/3, \pm 1/3, 0$	0.005, 0.007, 0.01, 0.02, 0.03, 0.04, 0.05	532
	$40^3 \times 64$	0.01/0.05	$\pm 2/3, \pm 1/3, 0$	0.005, 0.007, 0.01, 0.02, 0.03, 0.04, 0.05	700
	$48^3 \times 64$	0.01/0.05	$\pm 2/3, \pm 1/3, 0$	0.005, 0.007, 0.01, 0.02, 0.03, 0.04, 0.05	700
	$20^3 \times 64$	0.007/0.05	$\pm 4/3, \pm 1, \pm 2/3, \pm 1/3, 0$	0.005, 0.007, 0.01, 0.02, 0.03, 0.04, 0.05	364
	$24^3 \times 64$	0.005/0.05	$\pm 4/3, \pm 1, \pm 2/3, \pm 1/3, 0$	0.005, 0.007, 0.01, 0.02, 0.03, 0.04, 0.05	364
0.09	$28^3 \times 96$	0.0062/0.031	$\pm 4/3, \pm 1, \pm 2/3, \pm 1/3, 0$	0.0031, 0.0062, 0.0093, 0.0124, 0.0155, 0.0186, 0.031	364
	$40^3 \times 96$	0.0031/0.031	$\pm 2, \pm 4/3, \pm 1, \pm 2/3, \pm 1/3, 0$	0.0031, 0.0062, 0.0093, 0.0124, 0.0155, 0.0186, 0.031	532
0.06	$48^3 \times 144$	0.0036/0.018	$\pm 2, \pm 4/3, \pm 1, \pm 2/3, \pm 1/3, 0$	0.0036, 0.0054, 0.0072, 0.009, 0.0108, 0.0126, 0.018	532
	$56^3 \times 144$	0.0025/0.018	$\pm 2/3, \pm 1/3, 0$	0.0018, 0.0025, 0.0036, 0.0044, 0.0054, 0.0072, 0.0108, 0.0144, 0.018	1125
	$64^3 \times 144$	0.0018/0.018	$\pm 2/3, \pm 1/3, 0$	0.0018, 0.0025, 0.0036, 0.0044, 0.0054, 0.0072, 0.0108, 0.0144, 0.018	1125
0.045	$64^3 \times 192$	0.0028/0.014	$\pm 2/3, \pm 1/3, 0$	0.0014, 0.0021, 0.0028, 0.0035, 0.0042, 0.0056, 0.0084, 0.0112, 0.014	1125

B. Spectrum calculations

In order to calculate the meson spectrum, we read an archived dynamical $SU(3)$ gauge configuration and a quenched $U(1)$ gauge configuration and proceed to cast quark propagators from a corner wall source. We use a variety of valence quark charges and masses. A multishift solver is employed so that for each desired charge all desired masses are found with one iterative process.

The calculation of the meson spectrum has been primarily done on GPU based computers at the Texas Advanced Computing Center, National Center for Supercomputing Applications, and Indiana University using the QUDA approach pioneered at Boston University [39], but enhanced to support staggered quarks [40–43].

In Table III, we summarize the quark charges, masses, and number of channels we study on each ensemble. Figure 1 shows the Goldstone pion propagators as a function of Euclidean time for the $a \approx 0.045$ fm ensemble, which is our finest lattice spacing. We show four charge combinations for our lightest valence quark mass on that ensemble. Using the notation further detailed in Sec. III B, the quark charges are q_x and q_y in units of the fundamental charge e, and the meson charge q_{xy} is $q_x - q_y$ since the meson is made from an x-quark and y-antiquark. The combinations (q_x, q_y) we plot are $(0,0)$, $(2/3, 2/3)$, $(1/3, -2/3)$, and $(2/3, -2/3)$, with total charges $q_{xy} = 0, 0, 1$ and $4/3$, respectively. We see a nice linear decrease of the propagators in this semilogarithmic plot over a large range of t, before the periodic boundary conditions result in curvature at large t. In Fig. 2,
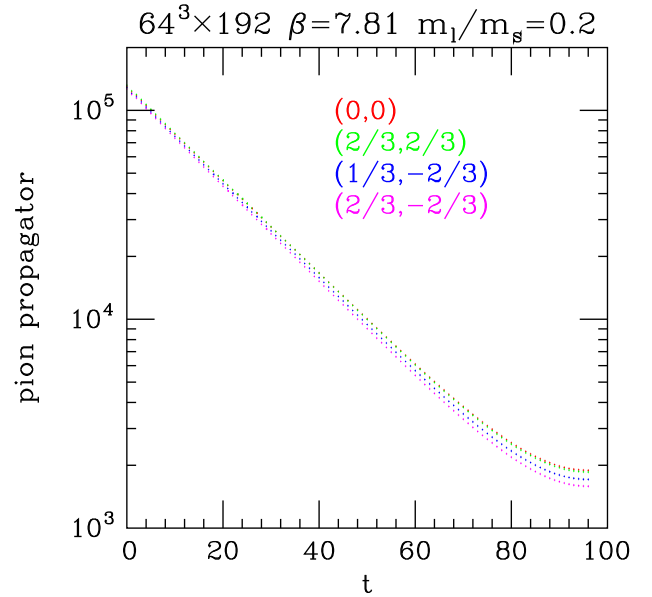

FIG. 1.   The Goldstone pion propagator as a function of Euclidean time on the $a \approx 0.045$ fm ensemble with $am_l' = 0.0028$ and $am_s' = 0.014$. The grid size is $64^3 \times 192$. The meson propagators are periodic in time and have been folded over, so the maximum time is 96. Four charge combinations are plotted: $(0,0)$, $(2/3, 2/3)$, $(1/3, -2/3)$, and $(2/3, -2/3)$, with total charges $q_{xy} = 0, 0, 1$ and $4/3$, respectively. These charges are all in units of e. The valence quark and antiquark masses are both 0.0014 in lattice units.
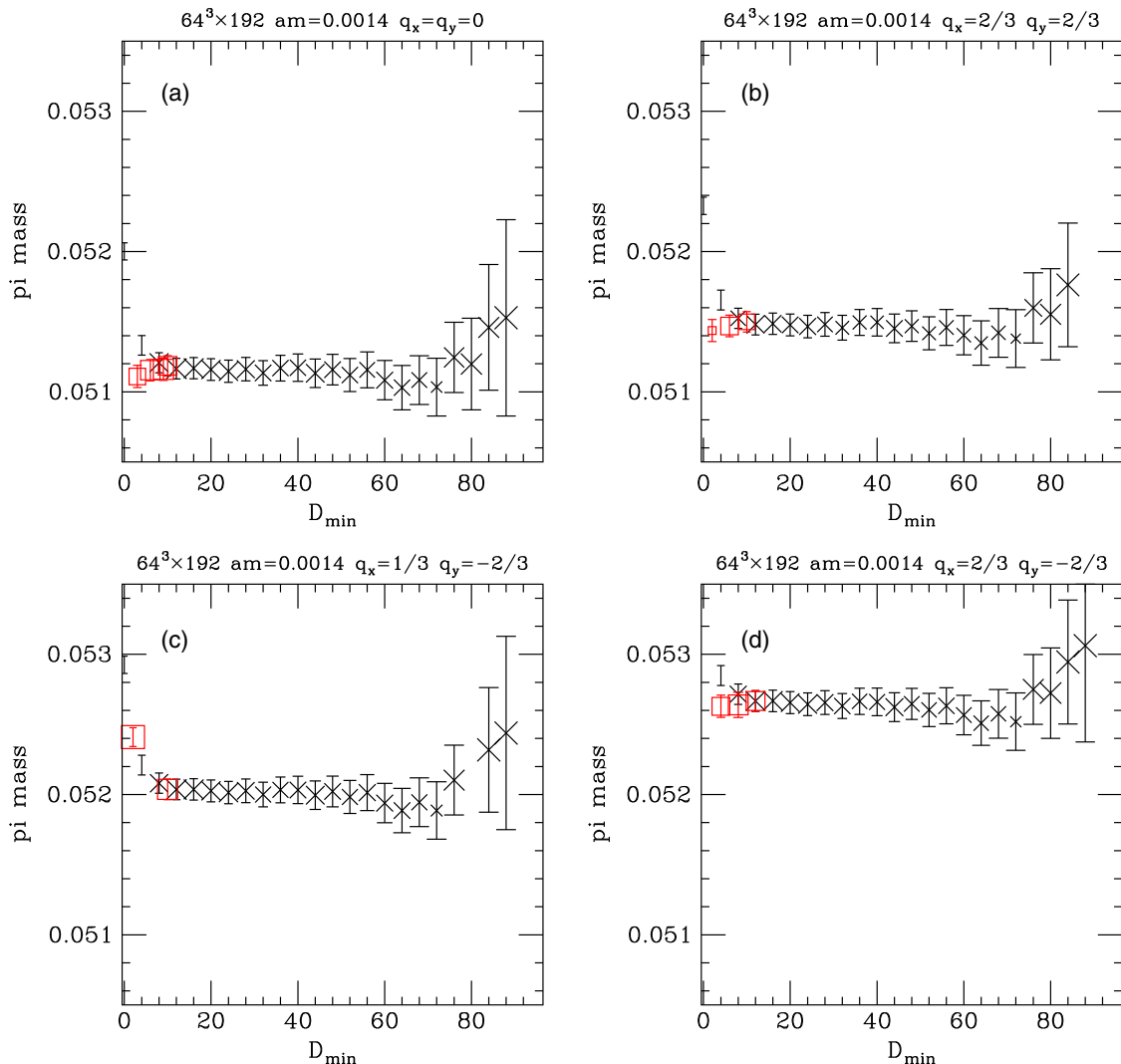
FIG. 2. Fits of the four pion propagators shown in Fig. 1. Fits are from D_{min} to the center of the lattice. The symbol sizes are proportional to the p value of the fit. Black crosses denote single-particle fits and red squares denote two-particle fits.

we show the results of fitting the propagators in Fig. 1. Each plot shows a series of fits starting from D_{min} and extending to the center of the lattice. The symbol size is proportional to the p value of the fit. Crosses are fits with a single particle (two free parameters), and squares correspond to two particles (four parameters). We see that there are many fits with good p values, and that the meson masses depend significantly on the total charge. We can even see a difference between the two cases of a neutral meson, one with uncharged quarks and the other made from a quark and an antiquark whose charges cancel each other. Much of this difference is unphysical, coming from the effect of EM quark-mass renormalization at fixed bare mass—see Sec. III C. Note that the quality of the plateaus for mesons with charged quarks is virtually identical to that for the meson with uncharged quarks; we return to this point in Sec. V. The masses corresponding to fits with $D_{min} = 50$, which is the value chosen for this ensemble in our final

analysis, are detailed in Table IV. Figure 3 plots these masses vs the square of the meson charge.

In order to construct the correlations among the masses of all the channels on a specific ensemble, we use a single

TABLE IV. The masses of the four mesons plotted in Fig. 1. On this ensemble, we fit the propagator from $t = 50$ to the center of the lattice assuming a single particle. The first two columns are the charges of the two quarks in units of e. Since the meson is made from a quark and an antiquark, the meson charge q_{xy} in the third column is $q_x - q_y$. The mass and its error are in the fourth column. Each fit has 45 degrees of freedom (d.o.f.). χ^2 and the p value of the fit are in columns five and six, respectively.

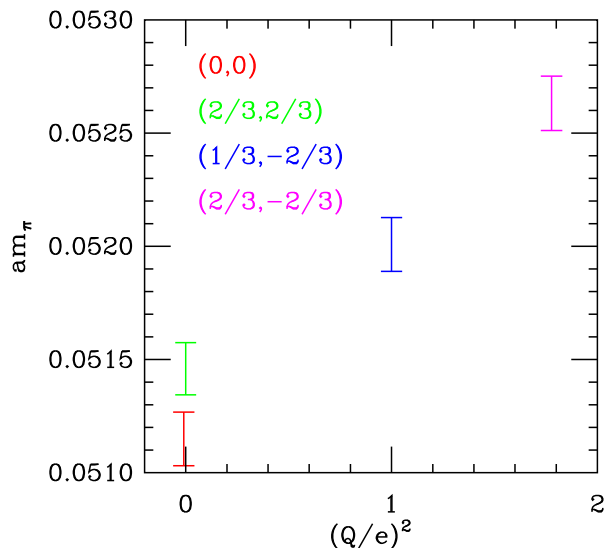| q_x	q_y	q_{xy}	am	χ^2	p
0	0	0	0.05115(12)	39.26	0.713
2/3	2/3	0	0.05146(12)	42.32	0.586
1/3	−2/3	1	0.05201(12)	39.43	0.706
2/3	−2/3	4/3	0.05263(12)	40.27	0.672

FIG. 3. The Goldstone pion mass as a function of the square of the meson charge on the $a \approx 0.045$ fm ensemble with $am'_l = 0.0028$ and $am'_s = 0.014$. Four charge combinations are plotted: $(0,0)$, $(2/3, 2/3)$, $(1/3, -2/3)$, and $(2/3, -2/3)$, with total charges $Q = 0$, 0, 1 and 4/3, respectively. These charges are all in units of e. The valence quark and antiquark masses are both 0.0014, in lattice units.

elimination jackknife fitting procedure. On each ensemble, a single value of $D_{\min}$ is used for all channels. In the subsequent analysis, we subtract the squared meson mass for $q = 0$ quarks from the corresponding squared meson mass with nonzero quark charges, properly taking in account the correlations. These correlations are expected to be very large, especially for mesons with the same valence quark masses but different valence quark charges, because the QCD contributions are identical in the two cases, and only the small QED effects are different. Because of this high degree of correlation, the errors of the subtracted quantities are much smaller than one would find by the naive propagation of errors from the masses themselves. For example, the correlation between the $q_x = 0$, $q_y = 0$ and the $q_x = 1/3$, $q_y = -2/3$ masses in Table IV is 0.998, and the error in the mass difference is 0.85%. If the error in the difference were propagated naively, omitting the correlation, the error in the mass difference would be about 20%.

The small errors in the subtracted masses is illustrated by an alternative analysis shown in Fig. 4 for the same data as in Figs. 1 and 2. Here, rather than fitting individual propagators, we fit the ratio of each propagator for a meson made of charged quarks with the corresponding propagator for the meson made from neutral quarks. Because of the effect of the periodic boundary conditions in time, the ratio depends not only on the meson mass difference, but also on the meson masses themselves. The latter dependence is mild, but still non-negligible, and makes fits with three unconstrained parameters (the mass

difference, the mass of $q_x = q_y = 0$ meson, and the overall amplitude) somewhat unstable. Instead, we have constrained, with Bayesian priors, the mass of the $q_x = q_y = 0$ meson to 0.05115(12), as given in Table IV. The plots show the resulting mass differences as a function of the minimum distances in the fits, $D_{\min}$. The horizontal solid and dotted lines (in red online) show the mass differences and errors computed from the individual masses using $D_{\min} = 50$.

C. Scale setting

We use the intermediate quantity r_1 [44,45] to set the relative scale of our ensembles, and take $r_1 = 0.3117(22)$ fm [46] as the absolute scale. From the smoothing fit to r_1/a values described in Ref. [4], we extrapolate the r_1/a values at the simulated quark masses to the physical quark masses (given below in Table V), holding β fixed. This defines a mass-independent scale-setting scheme, which is needed in order to apply chiral perturbation theory. The scheme is mass independent because it gives an r_1/a value that depends only on β and not on the simulated quark masses m'_l and m'_s. Mass-independent values of r_1/a for our ensembles are listed in Table I. The errors shown are a sum (in quadrature) of statistical errors and errors of the extrapolation to the physical quark masses.

III. CHIRAL PERTURBATION THEORY WITH ELECTROMAGNETISM

A. Continuum chiral theory for QCD + QED

In the continuum, the chiral effective theory for QCD + QED was worked out by Urech [22]. Along with all hadrons heavier than the pseudoscalar mesons, high-momentum photons are integrated out of the chiral theory, resulting in a single effective meson-interaction term at LO. Photons and mesons with low momentum (less than the chiral cutoff Λ_χ), are treated explicitly in this chiral perturbation theory (χPT).

The partially quenched version of the chiral theory is relevant here. In partial quenching, the valence and sea quarks are treated as distinct; when EM is included, this means that valence and sea quarks may have different electric charges and/or masses.[1] Bijnens and Danielsson [21] have calculated the meson masses and decay constants at NLO (one loop) in partially quenched χPT in QCD + QED with three flavors of sea quarks (u, d, s). A key insight of Ref. [21] is that sea-quark charges affect meson masses in particularly simple ways at NLO. In analytic terms involving the sea-quark charges, only the sums of the squared sea-quark charges appear—there are no cross terms between sea-quark and valence-quark charges. (It is necessary to assume here that the sum of the three sea-quark

---

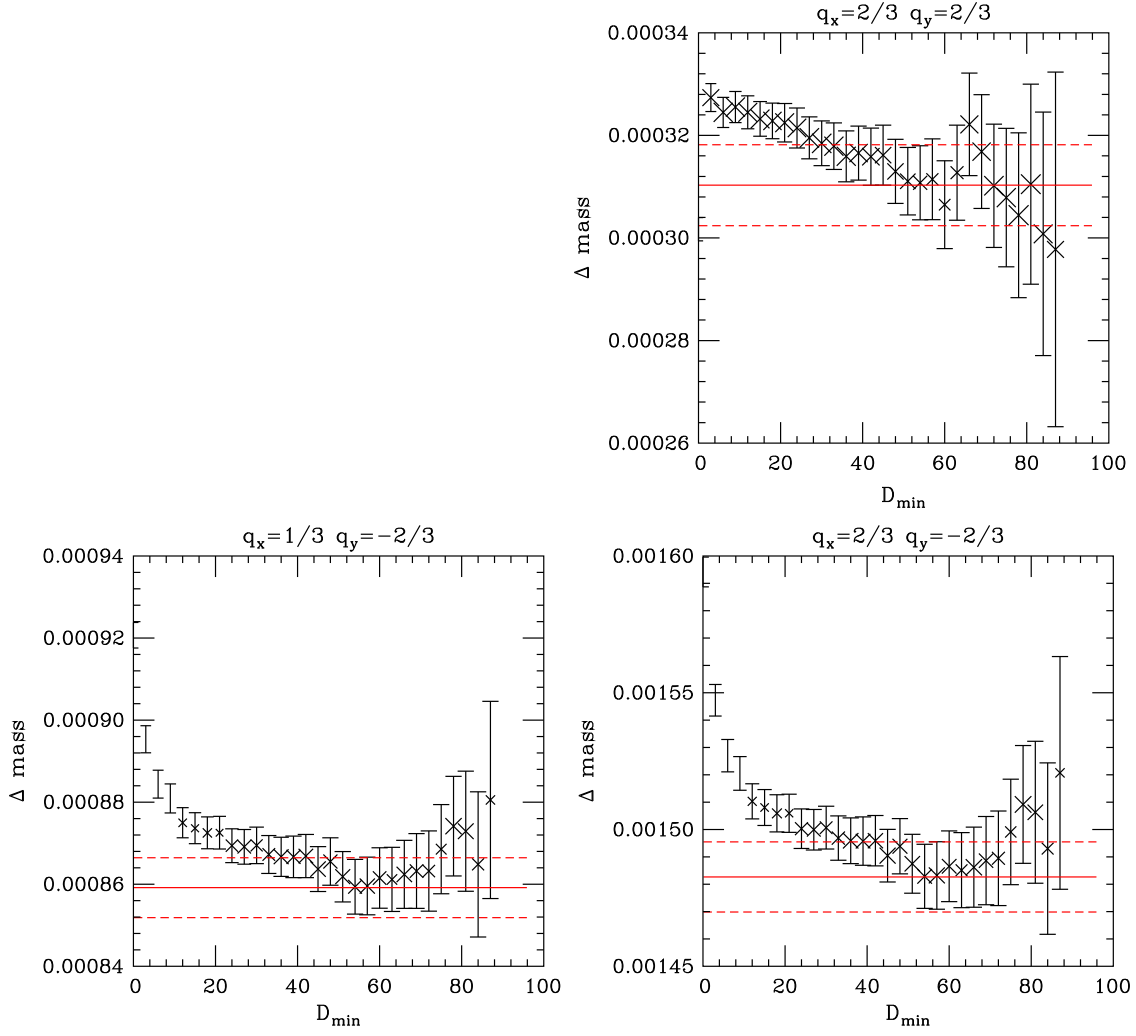[1]We will call the limit where valence- and sea-quark masses and charges are equal "full" QCD + QED.

FIG. 4. Fits of the ratio of the propagators shown in Fig. 1 for mesons with charged quarks, divided by the propagator for the meson with neutral quarks. The vertical axis gives the mass difference between the two propagators in the ratio. Fits are from D_{min} to the center of the lattice. The symbol sizes are proportional to the p value of the fit. The horizontal, red (color online), solid and dashed lines show the central value and error for the mass difference given by subtracting the masses in Table IV and propagating the errors using the covariance matrix determined by jackknife, as described in the text.

charges vanishes, as it does in the real world.) Sea-quark charges may also appear in the one-loop chiral logarithms, but these are completely determined in terms of the LO LECs. This implies that in the difference of squared mass of

two mesons with the same valence quark masses but different valence quark charges, the analytic terms depending on sea-quark charges cancel. Thus the difference may be reliably computed on the lattice with a simulation in

TABLE V. Quantities used as inputs in the chiral-discretization fits and/or their extrapolation. The first three columns identify the ensemble, and then we list, in r_1 units, the physical values of the light quark and strange quark mass, the slope B_0 [Eq. (13)], and the taste splittings for axial, tensor, vector, and singlet tastes, respectively. The last row is labeled "Cont." for "continuum"; see text for how this is defined. The errors for the quark masses are from the chiral extrapolation and the absolute scale, respectively; statistical errors are negligible. For the other quantities the errors given are statistical only.

$\approx a$ [fm]	β	m'_l/m'_s	$r_1 m_l$	$r_1 m_s$	$r_1 B_0$	$r_1^2 a^2 \Delta_A$	$r_1^2 a^2 \Delta_T$	$r_1^2 a^2 \Delta_V$	$r_1^2 a^2 \Delta_I$
0.12	6.76	0.005/0.05	0.00333(6)(5)	0.0919(16)(13)	6.832(4)	0.230(2)	0.371(5)	0.487(6)	0.609(17)
0.09	7.08	0.0031/0.031	0.00338(6)(5)	0.0927(16)(13)	6.639(6)	0.075(5)	0.124(6)	0.160(10)	0.222(18)
0.06	7.46	0.0018/0.018	0.00343(7)(5)	0.0937(16)(13)	6.487(6)	0.027(1)	0.044(2)	0.058(2)	0.071(3)
0.045	7.81	0.0028/0.014	0.00342(6)(5)	0.0936(16)(13)	6.417(6)	0.010(2)	0.017(3)	0.023(3)	0.028(3)
Cont.	7.08	$\cdots$	0.00361(7)(5)	0.0990(17)(14)	6.015(6)	0	0	0	0

which the sea quarks are uncharged (the electroquenched approximation). The sea-quark charge dependence, which comes only from one-loop chiral logarithms, may be put in after the fact. All dependence on unknown NLO LECs cancels in the difference.

Note that the quantities we need to calculate to determine ϵ in Eq. (2), namely $(M^2_{K^\pm} - M^2_{K^0})^\gamma$ and $(M^2_{\pi^\pm} - M^2_{``\pi^{0}"})^\gamma$, are squared-mass differences of the type required to make them reliably calculable with electroquenched simulations, in the sense described in the previous paragraph. We emphasize, however, that the calculability depends on using SU(3) (3-flavor) χPT at NLO, which will have non-negligible systematic corrections that need to be estimated. The alternative, treating only the u, d quarks as light (2-flavor SU(2) χPT), is not necessarily an improvement, despite the fact that in SU(2) χPT the errors are generically much smaller at a given order than in SU(3). The reason is that the calculability of squared-mass differences in the electroquenched approximation depends on the tracelessness of the quark charge matrix, which holds in SU(3), but not in SU(2). Thus, if SU(2) is used, the chiral errors are likely to be smaller, but one must include a separate quenching error that needs to be estimated in some independent fashion. That is the approach taken in Ref. [17].

In this paper, we compute the EM effect on the neutral kaon mass, $(M^2_{K^0})^\gamma = \epsilon_{K^0}(M^2_{\pi^+} - M^2_{\pi^0})^{\text{expt}}$, in addition to ϵ. In this case, χPT does not allow us to control the electroquenching error, because that error it is not computable at lowest nontrivial χPT order. The quantity $(M^2_{K^0})^\gamma$ is the difference between the squared mass of a neutral kaon made out of charged valence quarks, with a charged sea, and the squared mass of a kaon made out of neutral valence quarks, with a neutral sea. Even effects that depend on the sea-quark charges alone do not cancel here. Our estimate of the electroquenching error in $(M^2_{K^0})^\gamma$ is therefore based on large-N_c power counting only ($N_c = 3$ is the number of QCD colors), and must be considered a rough guide only.

B. Staggered chiral perturbation theory with EM

With the staggered lattice action, each quark flavor appears as four species, known as "tastes." This is a remnant of the 16-fold doubling of species of naive lattice fermions. To obtain standard QCD in the continuum limit, it is necessary to eliminate the unwanted taste d.o.f. in the sea. Our simulations accomplish this by taking the fourth-root of the fermion determinant for each quark flavor [47]. Numerical and theoretical arguments for the validity of this procedure in the continuum limit can be found in Refs. [48–59]. The appropriate chiral theory for staggered quarks with the rooting procedure is called "rooted staggered" χPT (rSχPT) [60,61]. Starting with the staggered chiral Lagrangian of Ref. [61], it is straightforward [25] to include EM effects following Ref. [22].

At leading order, the Euclidean, staggered QCD + QED chiral Lagrangian is[2]

$$\mathcal{L}^{(\text{LO})} = \frac{1}{4} F_{\mu\nu} F_{\mu\nu} + \frac{\lambda}{2} (\partial_\mu A_\mu)^2 + \frac{f^2}{8} \text{Tr}(d_\mu \Sigma^\dagger d_\mu \Sigma)$$
$$- \frac{B_0 f^2}{4} \text{Tr}(\mathcal{M}\Sigma + \mathcal{M}\Sigma^\dagger)$$
$$+ \frac{m_0^2}{24} (\text{Tr}(\Phi))^2 + a^2 \mathcal{V} - e^2 C \, \text{Tr}(Q\Sigma Q\Sigma^\dagger), \quad (6)$$

where Tr denotes a trace over flavor and staggered taste indices. The quantities A_μ, $F_{\mu\nu}$, and λ are the photon gauge potential, the EM field strength, and the gauge-fixing parameter, respectively. The meson fields are contained in

$$\Sigma = \exp(i\Phi/f), \quad \Phi = \begin{pmatrix} U & \pi^+ & K^+ \\ \pi^- & D & K^0 \\ K^- & \bar{K}^0 & S \end{pmatrix}, \quad (7)$$

where diagonal entries U D, and S, are the quark-antiquark pairs $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ respectively. Each of the meson fields $U, \Pi^+, K^+, \ldots$ in Eq. (7) are composed of 16 tastes, as in $\pi^+ \equiv \sum_{b=1}^{16} \pi_b^+ T_b$, where the T_b are the Hermitian taste generators

$$T_b = \{\xi_5, i\xi_{\mu 5}, i\xi_{\mu\nu}(\mu < \nu), \xi_\mu, \xi_I\}. \quad (8)$$

Here ξ_μ are a set of 4 Euclidean gamma matrices, $\xi_{\mu\nu} \equiv \xi_\mu \xi_\nu$, $\xi_{\mu 5} \equiv \xi_\mu \xi_5$, and $\xi_I \equiv I$ is the 4×4 identity matrix. The term $a^2 \mathcal{V}$ in Eq. (6) is the taste-violating potential [61], with a the lattice spacing. The anomaly term $\frac{1}{24} m_0^2 \langle \Phi \rangle^2$ gives mass to the η', and causes mixing of the flavor-neutral fields U, D, S through "hairpin" (quark-line disconnected) diagrams [62,63]. As usual in partially quenched and/or staggered calculations, it is convenient to keep this term and use the simple U, D, S basis along the diagonal of Φ. At the end of the calculation, we can take $m_0 \to \infty$ [64] and decouple the η'.

In Eq. (6), $\mathcal{M}$ is the quark mass matrix,

$$\mathcal{M} = \text{diag}(m_u, m_d, m_s), \quad (9)$$

and Q is the quark (electric) charge matrix

$$Q = \text{diag}(q_u, q_d, q_s) = \text{diag}(2/3, -1/3, -1/3), \quad (10)$$

with the property $\text{Tr}(Q) = 0$. The covariant derivative d_μ is given by

---

[2]Reference [25] used this Lagrangian but, because of space limitations, did not explicitly display it.

$$d_\mu \Sigma = \partial_\mu \Sigma - ieQA_\mu \Sigma + i\Sigma eQA_\mu, \qquad (11)$$

where we have set vector and axial source terms to zero since they are not needed for present purposes. Electromagnetic effects on the meson masses come both directly, from the low-energy photon field A_μ, and indirectly, through the term $e^2 C \text{Tr}(Q\Sigma Q\Sigma^\dagger)$ (with e the fundamental electric charge and C an LEC), which represents the effects of high-energy photons that have been integrated out.

With p a typical meson 4-momentum, and M and m generic meson and quark masses, respectively, the standard power-counting scheme of rSχPT is $p^2 \sim M^2 \sim m \sim a^2$, where factors of the chiral scale Λ_χ (to make the dimension of each quantity the same) are implicit. Including EM, χPT becomes a joint expansion in p^2 and e^2. The Lagrangian of Eq. (6) is LO in the sense that it includes the leading terms both in p^2 and in e^2. Even though EM corrections are in general smaller or much smaller than typical SU(3) chiral corrections,[3] we are interested here in EM quantities, which start at $\mathcal{O}(e^2)$, so e^2 terms are rightly included in the LO Lagrangian. One-loop diagrams from Eq. (6) then produce $\mathcal{O}(e^2 p^2)$ corrections, which we consider NLO. Higher nonanalytic (chiral log) corrections have not been computed in rSχPT, but it will be necessary to add higher-order analytic terms ($\mathcal{O}(e^2 p^4)$ and sometimes $\mathcal{O}(e^4)$ and $\mathcal{O}(e^4 p^2)$) in order to get acceptable chiral fits. We will refer to $\mathcal{O}(e^2 p^4)$ and $\mathcal{O}(e^4)$ terms as next-to-next-to leading order (NNLO), and those of $\mathcal{O}(e^4 p^2)$ or $\mathcal{O}(e^2 p^6)$ as N^{3}LO; this counting treats $e^2 \sim p^4$. Terms that go like e^4 ultimately have negligible impact on our results for ϵ, but can be necessary to describe small, but statistically significant, effects in our lattice data, especially when we include data for quarks with larger-than-physical charges.

We consider a generic pseudoscalar meson composed of two different valence quarks x and y with masses m_x and m_y. In units of e, the quark (*not* antiquark) charges are q_x and q_y, so that the meson charge is $q_{xy} = q_x - q_y$. At LO, the squared mass of such a meson with taste b is

$$M^2_{xy,b} = \chi_{xy,b} + q^2_{xy} e^2 \Delta_{\text{EM}}, \qquad (12)$$

$$\chi_{xy,b} = B_0(m_x + m_y) + a^2 \Delta_b, \qquad (13)$$

where $\chi_{xy,b}$ is the LO squared mass without EM effects, Δ_b is the taste splitting coming from the staggered potential $\mathcal{V}$, and

$$\Delta_{\text{EM}} \equiv \frac{4C}{f^2}. \qquad (14)$$

Dashen's theorem is immediately evident from Eq. (12) since the LO EM contribution proportional to Δ_{EM} is independent of quark masses.

We remark that Eqs. (12) and (13) are in general complete LO masses (with and without EM) only when the meson is flavor charged ($x \neq y$). For $x = y$, there are additional contributions in the taste-singlet case (coming from the anomaly, m_0, term) and the taste-vector or axial vector cases (coming from taste-violating hairpins in $\mathcal{V}$). As is standard in partially quenched or staggered χPT calculations, such terms are treated as separate two-meson vertices, giving rise to disconnected contributions to flavor-neutral propagators.

Beyond LO, the fourth-root procedure needs to be implemented. This can be done systematically at the level of the chiral theory by using a replica trick [65] for the sea quarks: replicating them n_r times and setting $n_r = 1/4$ at the end of the calculation [54,55,66]. (Additional, un-replicated valence quarks, here called x and y, must also be introduced.) We do not show the replications explicitly in Eq. (6); in practice it is actually more convenient at the one-loop level to use quark-flow techniques [67] to keep track of diagrams with sea-quark loops, and multiply them by hand by a factor of 1/4. Since both the replica and the quark-flow approaches distinguish sea and valence quarks, it is straightforward to take into account, in the chiral calculations, the fact that our simulations are partially quenched.

From Eq. (6), it is straightforward to compute the squared mass of a pseudoscalar meson to order NLO ($\mathcal{O}(p^4)$, one-loop). We focus on the taste-ξ_5 (pseudoscalar taste) meson because it is the valence meson that we have simulated. The taste-ξ_5 meson is a true Goldstone boson in the massless limit and in the absence of EM (for electrically charged mesons). From now on, we always mean the taste-ξ_5 meson if we do not otherwise specify the meson's taste.

We are interested in the EM contribution to the squared mass,

$$(M^2_{xy})^\gamma \equiv M^2_{xy} - M^2_{xy}\big|_{q_x=q_y=q_u=q_d=q_s=0} \quad \text{[fixed renorm. mass]},$$
$$(15)$$

where the second term on the right-hand side is the squared mass in a world without EM, where all quark charges, both valence (q_x, q_y) and sea (q_u, q_d, q_s) vanish. The difference should be taken at fixed renormalized quark masses, so that only physically meaningful EM effects contribute to $(M^2_{xy})^\gamma$. This is a nontrivial requirement because the masses of quarks with different charges, e.g., u and d, have different EM renormalization. It is much more convenient to work with an intermediate quantity

$$\Delta M^2_{xy} \equiv M^2_{xy} - M^2_{xy}\big|_{q_x=q_y=q_u=q_d=q_s=0} \quad \text{[fixed bare mass]},$$
$$(16)$$

---

[3]For example, a typical $\mathcal{O}(p^2)$ chiral correction is $(f_K - f_\pi)_{\text{NLO}}/f_\pi \sim 20\%$, while the $\mathcal{O}(e^2)$ (and higher) correction to the charged pion mass is $(M^2_{\pi^\pm} - M^2_{\pi^0})^\gamma / M^2_{\pi^0} \sim 5\%$, and is much less than that, on a percentage basis, for the kaon.

where the two terms on the right-hand side are computed at the same values of the bare quark masses.

On the lattice, we have computed M_{xy} for various choices of valence quark charges, including vanishing charges, for each valence bare quark mass studied. This means that it is straightforward to construct the quantity ΔM_{xy}^2, as well as its correlated errors with other choices of quark charges and valence masses. On the other hand, the construction of $(M_{xy}^2)^\gamma$ would require theoretical assumptions about the EM mass renormalization, coupled with interpolation or extrapolation of the data to adjust the bare masses in the subtraction in Eq. (15). It is much easier to postpone the renormalization step until after the chiral fit, when we will have the ability to make these adjustments easily. Fortunately, the functional form of the chiral fit that is appropriate to the physical quantity $(M_{xy}^2)^\gamma$ may also be applied to a fit of the unphysical intermediate quantity ΔM_{xy}^2. As we will see, the only consequence of fitting ΔM_{xy}^2 instead of $(M_{xy}^2)^\gamma$ is that the former will have unphysical contributions to two LECs that are affected by EM renormalization. We therefore postpone detailed discussion of renormalization until Sec. III C. Except for some comments about the affected LECs, we ignore the difference between ΔM_{xy}^2 and $(M_{xy}^2)^\gamma$ in the current section.

Separating orders in the chiral expansion, we write the difference in Eq. (16) as

$$\Delta M_{xy}^2 = \Delta_{\text{LO}} M_{xy}^2 + \Delta_{\text{NLO}} M_{xy}^2 + \Delta_{\text{NNLO}} M_{xy}^2 + \cdots, \quad (17)$$

$$\Delta_{\text{LO}} M_{xy}^2 = q_{xy}^2 e^2 \Delta_{\text{EM}}, \quad (18)$$

$$\Delta_{\text{NLO}} M_{xy}^2 = \Delta_{\text{NLO}}^{\log} M_{xy}^2 + \Delta_{\text{NLO}}^{\text{analytic}} M_{xy}^2, \quad (19)$$

where $\Delta_{\text{LO}} M_{xy}^2$ is independent of taste. Equation (18) follows from Eq. (12), and Eq. (19) divides the NLO contribution into logarithmic (nonanalytic) and analytic contributions. For NNLO and higher orders, the chiral logarithms are not known; when such orders are needed in the chiral fits, we therefore include the analytic contributions only.

The mass of the Goldstone meson has been computed to NLO [one loop, $\mathcal{O}(p^4, e^2 p^2)$] in rS$\chi$PT with EM in Ref. [25]. Figure 5 shows the NLO contributions to the meson mass. The photon tadpole diagram does not contribute here since it vanishes in dimensional regularization; in FV, however, the momentum integral becomes a sum, and the photon tadpole is nonzero, as discussed in Sec. IV. The photon sunset diagram is essentially the same as in the continuum, since the meson-photon vertex is taste-conserving, and the external pseudoscalar-taste meson is also the meson in the loop. The calculation of the contribution from the meson tadpole, Fig. 5(c), is very similar to that in Ref. [61], with the addition of a new 4-meson vertex from the C term in Eq. (6).
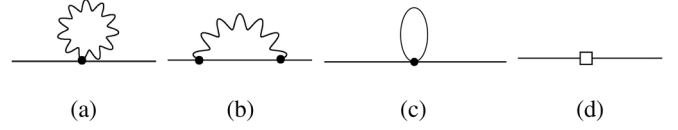

FIG. 5. Feynman diagrams that contribute to the meson-mass at $O(p^4, e^2 p^2)$. Straight lines are the pseudoscalar meson propagator and wiggly lines are the photon. A filled dot represents a vertex from the $O(p^2, e^2)$ Lagrangian, $\mathcal{L}^{(\text{LO})}$, while an open square represents an insertion of the $O(p^4, e^2 p^2)$ Lagrangian, $\mathcal{L}^{(\text{NLO})}$. (a) photon tadpole; (b) photon sunset; (c) meson tadpole; (d) $O(p^4, e^2 p^2)$ tree-level insertion.

The result of the calculation is that the NLO contribution to the squared mass splits into an EM contribution proportional to e^2 and a non-EM contribution, which is identical to that in Ref. [61], and which cancels in the difference $\Delta M_{xy,5}^2$, where we include the subscript 5 to emphasize here that we are talking about the meson with taste ξ_5. The one-loop diagrams Fig. 5(a)–5(c) give

$$\Delta_{\text{NLO}}^{\log} M_{xy,5}^2 = -\frac{1}{16\pi^2} e^2 q_{xy}^2 \chi_{xy,5} [3\ln(\chi_{xy,5}/\Lambda_\chi^2) - 4]$$
$$- \frac{2e^2 \Delta_{\text{EM}}}{16\pi^2 f^2} \left(\frac{1}{16}\right)$$
$$\times \sum_{\sigma,b} [q_{x\sigma} q_{xy} \ell(\chi_{x\sigma,b}) - q_{y\sigma} q_{xy} \ell(\chi_{y\sigma,b})], \quad (20)$$

where sea-quark flavors and the 16 meson tastes are labeled by σ and b, respectively, Λ_χ is the chiral scale, and $\ell(\chi)$ is the renormalized loop integral

$$\int \frac{d^4 k}{\pi^2} \frac{1}{k^2 + \chi} \to \ell(\chi) \equiv \chi \ln(\chi/\Lambda_\chi^2). \quad (21)$$

The result in the first line in Eq. (20) is from the photon sunset diagram, Fig. 5(b), and that in the second and third lines is from the meson tadpole, Fig. 5(c). We have put the squared masses on the right-hand side to their values in the absence of EM ($\chi_{xy,b}$), rather than the full LO masses, Eq. (12). This change makes only a higher order, $\mathcal{O}(e^4)$, difference. In chiral fits, we have also tried replacing $\chi_{xy,b}$ by the full LO masses in the one-loop terms; the small difference does not change either the quality of fits or the physical results significantly.

The contributions from Fig. 5(d) lead to analytic contributions with unknown LECs. It is useful to write these contributions in terms of natural dimensionless variables of rSχPT [7]

$$\mu_i = \frac{2B_0 m_i}{8\pi^2 f_\pi^2},$$
$$\mu_{a^2} = \frac{a^2 \bar{\Delta}}{8\pi^2 f_\pi^2}, \quad (22)$$

where i labels quark flavors, $i \in \{x, y, u, d, s\}$, and $\bar{\Delta}$ is the mean of the taste splittings Δ_b (weighted by multiplicities). It is straightforward to find the possible analytic terms using the standard spurion approach, but it is much quicker simply to write down all polynomials of a given order using the rules that follow from the symmetries:

(1) Charge conjugation symmetry implies that a valence $x\bar{y}$ meson has the same mass as its antiparticle, the $y\bar{x}$ meson, so terms must be symmetric under the interchange $q_x, \mu_x \leftrightarrow q_y, \mu_y$.

(2) In the absence of EM, the partially conserved staggered axial symmetry that rotates x into y quarks guarantees that $M^2_{xy,5}$ is proportional to $m_x + m_y$ (times possible additional mass factors). When EM is turned on, the symmetry is explicitly broken, but only for charged mesons ($q_x \neq q_y$). Thus, when $q_x = q_y$, all terms must either vanish or be proportional to $m_x + m_y$.

(3) The fact that the sea quarks couple equally to valence quarks implies that terms must be symmetric under sea-quark interchange: $q_u, \mu_u \leftrightarrow q_d, \mu_d \leftrightarrow q_s, \mu_s \leftrightarrow q_u, \mu_u$.

(4) The sum of sea-quark charges vanishes in the two cases of interest here, the physical case and the electroquenched case. Therefore, terms proportional to the sum $q_u + q_d + q_s$ may be dropped.

Given these rules, there are six independent analytic contributions possible at $\mathcal{O}(e^2 p^2)$ (NLO):

$$e^2 q^2_{xy}\mu_{a^2}, \qquad e^2 q^2_{xy}(\mu_u + \mu_d + \mu_s), \qquad e^2(q^2_x + q^2_y)(\mu_x + \mu_y),$$

$$e^2 q^2_{xy}(\mu_x + \mu_y), \qquad e^2(q^2_x\mu_x + q^2_y\mu_y), \qquad e^2(q^2_u + q^2_d + q^2_s)(\mu_x + \mu_y). \tag{23}$$

Of these, the last contribution will cancel for ΔM^2_{xy} since it is independent of the valence charges. The remaining contributions are independent of the sea-quark charges. That means that sea-quark-charge dependence only enters at NLO in the chiral logarithms, Eq. (20), and hence is computable, as discovered by Bijnens and Danielsson [21].[4]

The result of an $\mathcal{O}(1)$ shift in the scale of the chiral logarithms suggests that an appropriate scale for the analytic contributions in Eq. (23) is f^2_π [from the first line in Eq. (20)] or Δ_{EM} [from the second line in Eq. (20)]. In fact, these two quantities are the same order of magnitude, as can be seen by estimating Δ_{EM} by assuming that the experimental π^+–π^0 mass splitting comes entirely from the leading-order contribution $e^2\Delta_{EM}$. We therefore choose the scale f^2_π for all the NLO analytic contributions. In addition, we find it helpful to include mean taste splittings in analytic terms that absorb the chiral-scale dependence coming from the meson tadpole, which has an average over tastes in the third line in Eq. (20). As in Ref. [7], this definition of the LECs at nonzero lattice spacing simplifies the chiral-scale dependence of the LECs, and also tends to capture much of the lattice-spacing dependence of the lattice data, reducing the size of the pure discretization term (proportional to μ_{a^2}) in the fit. The NLO analytic contribution to ΔM^2_{xy} is then

$$\Delta^{analytic}_{NLO} M^2_{xy} = e^2 f^2_\pi [\kappa_1 q^2_{xy}\mu_{a^2} + \kappa_2 q^2_{xy}(\mu_u + \mu_d + \mu_s + 3\mu_{a^2})$$
$$+ \kappa_3(q^2_x + q^2_y)(\mu_x + \mu_y)$$
$$+ \kappa_4 q^2_{xy}(\mu_x + \mu_y) + \kappa_5(q^2_x\mu_x + q^2_y\mu_y + q^2_{xy}\mu_{a^2})]. \tag{24}$$

The usual expectation would be that the dimensionless LECs κ_i are $\mathcal{O}(1)$. However, several features of the current problem indicate that the expectation may be violated. First of all, previous work, both in the continuum [3] and on the lattice [10–14,16–18], suggests that ϵ is large ($\mathcal{O}(1)$, rather than $\ll 1$), which would imply that the NLO terms produce $\mathcal{O}(1)$ corrections to the LO result, and hence that at least some of the NLO LECs may be expected to be significantly larger than 1. A second issue arises from the nature of our data set. Because the ensembles we study here all have a strange quark mass tuned to near the physical value ($m'_s \approx m_s$), and a light quark mass significantly lighter than that ($m'_l \lesssim 0.2m'_s$), the κ_2 term in Eq. (24) is approximately a constant up to discretization errors, and may therefore compete in the fit with the LO term $q^2_{xy}e^2\Delta_{EM}$. In most fits, in fact, κ_2 has a tendency to get large and Δ_{EM} to get small—even negative in some cases. This is a typical problem that occurs with SU(3) fits to data sets in which m'_s does not take a significant range of values less than m_s. Fortunately, the final results for physical quantities depend only mildly on the relative sizes of the LO and κ_2 terms. In most of our fits, including the central fit, we simply set $\kappa_2 = 0$, but leave Δ_{EM} unconstrained. However we also consider fits where both Δ_{EM} and κ_2 are unconstrained, as well as ones in which κ_2 is constrained by a prior that enforces $\kappa_2 \lesssim 1$. Differences between results of these fits and the central one are included in an estimate of the systematic error of the chiral extrapolation.

---

[4]The fact that $q_u + q_d + q_s = 0$ is crucial to this conclusion. If the sum of the sea-quark charges were not zero, calculability of sea-quark charge dependence of ΔM^2_{xy} at NLO would be spoiled, e.g., by a term in M^2_{xy} proportional to $(q_x+q_y)(q_u+q_d+q_s) \times (\mu_x+\mu_y)$. Such a term would be generated, in the notation of Ref. [21], by a contribution to the Lagrangian of the form $\mathrm{Tr}(\mathcal{Q}_L)\mathrm{Tr}(\mathcal{Q}_R u^\nu u_\nu) + (L \leftrightarrow R)$.

A final complication is the fact that $\Delta M_{xy,5}^2$, the quantity we are fitting, includes unphysical contributions because it has not been adjusted for the effects of EM quark-mass renormalization. In particular, the term multiplied by κ_5 in Eq. (24) is precisely of the form that would be induced by the $\mathcal{O}(\alpha_{EM})$ EM renormalization of the quark masses m_x and m_y, so κ_5 will have an unphysical renormalization contribution. Indeed, all fits that do not include an additional correction for renormalization give $\kappa_5 \approx 12$, with $\kappa_5 = 12.2(2)$ in the central fit. After renormalization is taken into account in some way, this effective value of κ_5 is significantly reduced. On our central fit, the preferred nonperturbative scheme described in Sec. III C is nearly equivalent to simply setting $\kappa_5 = 0$ after the fit. With an $\overline{MS}$ scheme and a perturbative determination of the renormalization constant at one loop, κ_5 is reduced, effectively, by a factor of 2 but remains clearly nonzero.

Beyond NLO, the SχPT logarithms have not been calculated, so we are unable to continue the chiral expansion in a systematic fashion. However, for acceptable chiral fits to the lattice data, we must include some or all of the NNLO analytic terms, and at least one N³LO term. Following the symmetry rules above, the independent NNLO terms (for vanishing sea-quark charges) are

$$
\begin{aligned}
\Delta_{\text{NNLO}}^{\text{analytic}} M_{xy}^2 = {}& e^2 f_\pi^2 [\rho_1 q_{xy}^2 \mu_{a^2}^2 + \rho_2 q_{xy}^2 \mu_{a^2}(\mu_u + \mu_d + \mu_s + 3\mu_{a^2}) + \rho_3(q_x^2 + q_y^2)\mu_{a^2}(\mu_x + \mu_y) \\
& + \rho_4 q_{xy}^2 \mu_{a^2}(\mu_x + \mu_y) + \rho_5 q_{xy}\mu_{a^2}[q_x(\mu_x + \mu_{a^2}) - q_y(\mu_y + \mu_{a^2})] \\
& + \rho_6 q_{xy}^2(\mu_u + \mu_d + \mu_s + 3\mu_{a^2})^2 + \rho_7 q_{xy}^2(\mu_u^2 + \mu_d^2 + \mu_s^2) \\
& + \rho_8 q_{xy}^2(\mu_x + \mu_y)(\mu_u + \mu_d + \mu_s + 3\mu_{a^2}) + \rho_9 q_{xy}(\mu_u + \mu_d + \mu_s + 3\mu_{a^2})[q_x(\mu_x + \mu_{a^2}) - q_y(\mu_y + \mu_{a^2})] \\
& + \rho_{10}(q_x^2 + q_y^2)(\mu_x + \mu_y)(\mu_u + \mu_d + \mu_s + 3\mu_{a^2}) + \rho_{11}(q_x^2 + q_y^2)(\mu_x + \mu_y)^2 \\
& + \rho_{12}(q_x^2 - q_y^2)(\mu_x^2 - \mu_y^2) + \rho_{13} q_{xy}^2(\mu_x + \mu_y)^2 + \rho_{14} q_{xy}^2(\mu_x^2 + \mu_y^2)] \\
& + e^4 f_\pi^2 [\rho_1' q_{xy}^2(q_x^2 + q_y^2) + \rho_2'(q_x^2 - q_y^2)^2],
\end{aligned}
\tag{25}
$$

where the terms with ρ_i coefficients are $\mathcal{O}(e^2 p^4)$, and those with ρ_i' coefficients are $\mathcal{O}(e^4)$. Taste-splitting terms (μ_{a^2}) have been added to mass terms (μ_j) in plausible ways based on the example of the NLO chiral logarithms, but of course these choices are merely guesses of how best to absorb discretization errors into the mass terms.

Equation (25) includes taste-violating analytic terms, such as the term multiplied by ρ_1, that arise naturally in rSχPT. However, lattice-spacing dependence can also arise simply from "generic" discretization effects that break no continuum symmetries and therefore produce no new LECs. Rather, they induce a-dependence in the LECs that are already present. While the leading taste violations in QCD with asqtad quarks are $\mathcal{O}(\alpha_S^2 a^2)$, the leading generic errors are $\mathcal{O}(\alpha_S a^2)$. The quark couplings to EM do not change the leading generic errors because the combination of paths in the asqtad action removes $\mathcal{O}(a^2)$ terms as always. However, the EM gauge action we use is unimproved and therefore induces $\mathcal{O}(a^2)$ generic errors.[5]

Generic discretization errors of the NLO analytic parameters $\kappa_1, ..., \kappa_5$ in Eq. (24) may produce effects of a size comparable to that from the NNLO parameters, so should be included. Even more important, a generic error on the LO parameter Δ_{EM} may induce effects comparable to NLO and is therefore required in our fits. We thus include six generic variation parameters $\psi_0, ..., \psi_5$ that give a-dependence to the LO and NLO LECs:

$$
\Delta_{EM}(a) = \Delta_{EM}\left(1 + \psi_0 \frac{a^2}{r_1^2}\right),
\tag{26}
$$

$$
\kappa_i(a) = \kappa_i\left(1 + \psi_i \frac{a^2}{r_1^2}\right), \qquad (i = 1, ..., 5).
\tag{27}
$$

The parameters Δ_{EM} and κ_i on the right-hand side here are the continuum ($a = 0$) values. In Eqs. (26) and (27), we have assumed $\mathcal{O}(a^2)$ generic errors. However, we also make fits assuming $\mathcal{O}(\alpha_S a^2)$ errors, and include the results of those fits in our systematic error estimates. In practice, it makes little difference whether we assume $\mathcal{O}(a^2)$ or $\mathcal{O}(\alpha_S a^2)$ generic errors. Fits with the former actually tend to have slightly lower p values and slightly larger statistical errors. Nevertheless, they are preferred because the leading errors are $\mathcal{O}(a^2)$.

At N³LO, possible terms are $\mathcal{O}(e^4 p^2)$ or $\mathcal{O}(e^2 p^6)$. The latter are not necessary for good fits on any subsets of our data that we have considered, and we do not discuss them further here. The former are necessary, especially when we include data with quark charges larger than their physical values. The independent N³LO $\mathcal{O}(e^4 p^2)$ terms are

---

[5]The fact that the a^2 errors occur in the EM sector, and therefore automatically come with a factor of α_{EM} in quark quantities, does not help here because we are focusing on EM quantities, which have that same overall α_{EM} factor.

$$\Delta_{\mathrm{N^3LO}}^{\mathrm{analytic}} M_{xy}^2 = e^4 f_\pi^2 [\lambda_1 q_{xy}^2 (q_x^2 + q_y^2) \mu_{a^2} + \lambda_2 (q_x^2 - q_y^2)^2 \mu_{a^2} + \lambda_3 q_{xy}^2 (q_x^2 + q_y^2)(\mu_u + \mu_d + \mu_s + 3\mu_{a^2})$$
$$+ \lambda_4 (q_x^2 - q_y^2)^2 (\mu_u + \mu_d + \mu_s + 3\mu_{a^2}) + \lambda_5 (q_x^4 + q_y^4)(\mu_x + \mu_y) + \lambda_6 (q_x^4 \mu_x + q_y^4 \mu_y) + \lambda_7 q_x q_y (q_x^2 + q_y^2)(\mu_x + \mu_y)$$
$$+ \lambda_8 q_x q_y (q_x^2 \mu_x + q_y^2 \mu_y) + \lambda_9 q_{xy}^4 (\mu_x + \mu_y)]. \tag{28}$$

When the charges of the quarks in the mesons are limited to physical values or smaller ($\pm 2e/3$, $\pm e/3$, or 0), only the λ_6 term is necessary for acceptable fits, and its value is ≈ 4. (The central fit gives $\lambda_6 = 4.1(1)$.) Note that this term has the form of an $\mathcal{O}(e^4)$ quark mass renormalization. This implies that λ_6, like the NLO LEC κ_5 [Eq. (24)], has an unphysical renormalization contribution. We note that, even though fits with λ_6 set to zero have very low p values, $<10^{-10}$, the term has little effect on the physical quantities studied here. In particular, if we simply set $\lambda_6 = 0$ after the fit, these quantities change by amounts less than or equal to their statistical errors, and much less than their total (systematic plus statistical) errors.

C. Electromagnetic quark-mass renormalization

In this section, we discuss the renormalization of quark masses due to EM effects, i.e., $\mathcal{O}(e^2)$ or higher. This is important because the multiplicative renormalization factor Z_m is different for quarks with different EM charges, and thus affects how we separate "true" EM effects from quark mass effects such as isospin violations. Because we are not interested here in determining absolute, physical quark masses (e.g., $\overline{\mathrm{MS}}$ quark masses in MeV, say), renormalization due to the strong interactions alone can be ignored since the corresponding Z_m is the same for all quark flavors. Therefore, when we refer in this paper to "renormalized" or "bare" quark masses, we mean renormalized or bare with respect to EM. All quark masses discussed are bare as far as the strong interactions are concerned.

It is instructive first to estimate the size of the EM renormalization effect on the determination of ϵ. At fixed lattice spacing a, let δu and δd be the fractional shift in the u and d bare masses such that their renormalized EM masses are both equal to m_l. At $\mathcal{O}(e^2)$, we have

$$m_l(\delta u - \delta d) = C \frac{(q_u e)^2 - (q_d e)^2}{4\pi} m_l = C \frac{\alpha_{\mathrm{EM}}}{3} m_l. \tag{29}$$

Assuming that the size of any logarithms in $a\mu$ remains modest (μ is the scale of the renormalized masses), the constant C is expected to be of order 1. With $\alpha_{\mathrm{EM}} \sim 0.01$, this gives $\delta u - \delta d \sim 0.003$. Compared to the experimental pion splitting, the induced mass-squared splitting between a K^+ and a K^0 is then approximately

$$\frac{B_0 m_l(\delta u - \delta d)}{m_{\pi^+}^2 - m_{\pi^0}^2} \sim \frac{0.003 m_\pi^2 / 2}{m_{\pi^+}^2 - m_{\pi^0}^2} \sim 0.02. \tag{30}$$

Our estimate of the EM renormalization effect on ϵ is thus quite small, 0.02. The reason the effect is small is that the residual chiral symmetry of staggered quarks guarantees that the renormalization is multiplicative, so that the shifts in the u- and d-quark masses are small. The shift in the s-quark mass is much larger; however, its effect cancels in ϵ between $M_{K^+}^2$ and $M_{K^0}^2$. On the other hand, for quantities such as $(M_{K^0}^2)^\gamma$, the EM effect on the squared K^0 mass itself, the fractional systematic error from not including renormalization effects is at least an order of magnitude larger than for ϵ. One must also keep in mind that the estimate in Eq. (30) is qualitative, and could easily be off by a factor of 3 or more if C is larger or smaller than naively expected.

We now proceed to more detailed discussion of perturbative renormalization, which converts bare quark masses to $\overline{\mathrm{MS}}$ renormalized masses at some convenient scale, here taken to be $\mu = 2$ GeV. Only a one-loop determination is available in the literature. For staggered quarks in QCD, the renormalized $\overline{\mathrm{MS}}$ mass is given at this order in terms of the bare mass $m(a)$ at lattice spacing a by [68]

$$m^{\overline{\mathrm{MS}}}(\mu) = Z_m m(a) = (1 + \alpha_V(q^*) Z_m^{(2)}(a\mu)) m(a), \tag{31}$$

$$Z_m^{(2)}(a\mu) = b - \frac{4}{3\pi} - \frac{2}{\pi} \ln(a\mu), \tag{32}$$

where $\alpha_V(q^*)$ is the strong coupling in the V scheme [69] evaluated at scale q^*, and b is a constant depending on the details of the staggered action. We have neglected discretization corrections of $\mathcal{O}((am)^2)$.

In order to find the corresponding EM renormalization for staggered quarks, we merely have to remove the overall SU(3) Casimir factor of 4/3 from $Z_m^{(2)}$ and to replace $\alpha_V(q^*)$ with $\alpha_{\mathrm{EM}} = e^2/(4\pi)$. Issues such as the proper scheme and scale q^* for α_{EM} are irrelevant since α_{EM} is so small compared to α_S, and hence runs very slowly. Because we do not include EM corrections to the QCD tadpole factors in the asqtad action, we take $b = 2.27$ [68], which corresponds to the case of asqtadlike smearing without tadpole improvement. The one-loop EM renormalization is then

$$\delta m \equiv (m^{\overline{\mathrm{MS}}}(\mu) - m(a))_{\mathrm{EM}}$$

$$= q^2 e^2 m(a) \left(c - \frac{3}{8\pi^2} \ln(a\mu) \right), \qquad c = 0.110, \quad (33)$$

where q is the charge of the quark in units of e.

The EM renormalization first affects ΔM_{xy}^2 at NLO in χPT. To include one-loop renormalization in the chiral fit at this order, we simply add

$$\Delta_{\mathrm{renorm}} M_{xy}^2 = B_0(\delta m_x + \delta m_y) \qquad (34)$$

to Eq. (17). Note that changes in μ can then be absorbed in the chiral fits by changes in the NLO LECs: κ_5 and (if discretization effects are important) κ_1, Eq. (24). After the fit, the effect of Eq. (34) is removed from the result. This procedure is equivalent to readjusting the bare quark masses so that the renormalized masses have the desired value, so that, in particular, $m_u^{\overline{\mathrm{MS}}}(\mu) = m_d^{\overline{\mathrm{MS}}}(\mu)$. As discussed below in Sec. VII, the net result is that including the one-loop EM renormalization would shift ϵ by 0.03, with small variations depending on the details of the fit. This is consistent with (but somewhat larger than) the order-of-magnitude estimate of the effect made above. Based on this small shift, which is significantly less than the other systematic errors in our result, our approach in preliminary calculations [12–14,27] was to omit renormalization in the central value, and simply include an estimate of the effect in the systematic errors. However, Eq. (33) will get strong corrections starting at two loops, i.e., $\mathcal{O}(\alpha_S e^2)$, and experience from pure-QCD quark mass renormalization suggests that we would need the corrections through $\mathcal{O}(\alpha_S^2 e^2)$ to be able to be confident of the coefficient of e^2 at the few percent level.[6] We are thus only able to take Eq. (33) as a qualitative estimate of the EM renormalization effect in the $\overline{\mathrm{MS}}$ scheme.

In the absence of high-order perturbative calculations, a nonperturbative determination of the EM renormalization is necessary to get reliable results. As we will see below, such a nonperturbative approach yields an estimate for the effect of EM renormalization on ϵ of approximately 0.07, a bit more than twice as large as the one-loop perturbative estimate.

The nonperturbative method we use has been proposed by the BMW Collaboration [71]. The idea is to compare the masses of neutral π^0-like mesons constructed from $u\bar{u}$ quarks and $d\bar{d}$ quarks with quark-line connected propagators only (no intermediate states with only gluons and/or photons are allowed).

We first introduce the needed connected correlators for arbitrary valence quarks x and y. The connected $x\bar{x}$ and $y\bar{y}$ correlators are explicitly constructed in PQQCD by

---

[6]Compare, e.g., the one-loop result for the strange quark mass in Ref. [68] with the two-loop result of Ref. [70].

introducing additional valence flavors x' and y' with $q_{x'} = q_x$, $m_{x'} = m_x$ and $q_{y'} = q_y$, $m_{y'} = m_y$. The connected correlators are then

$$G_{xx'}(t) = \frac{1}{2} \sum_{\vec{z}} \langle \bar{x}(t, \vec{z}) \gamma_5 x'(t, \vec{z}) \bar{x}'(0) \gamma_5 x(0) \rangle, \quad (35)$$

$$G_{yy'}(t) = \frac{1}{2} \sum_{\vec{z}} \langle \bar{y}(t, \vec{z}) \gamma_5 y'(t, \vec{z}) \bar{y}'(0) \gamma_5 y(0) \rangle, \quad (36)$$

where disconnected contributions are absent since x and x' are different quarks, so x cannot contract with $\bar{x}'$ (and similarly for y and $\bar{y}'$). We let $M_{xx'}$ and $M_{yy'}$ be their masses. These mesons are each of the form discussed in rule 2 above Eq. (23): neutral mesons composed of two different, but equally charged, quarks. The EM contributions to $M_{xx'}$ and $M_{yy'}$ must therefore be proportional to $B_0 q_x^2 e^2 (m_x + m_{x'})$ and $B_0 q_y^2 e^2 (m_y + m_{y'})$, respectively, where we have inserted the factor of B_0 to put these contributions in units of squared meson mass. For $q_x = 2/3$, $m_x \sim m_u$, and $q_y = -1/3$, $m_y \sim m_d$, the contributions are of order $\alpha_{\mathrm{EM}} M_\pi^2$. This is much smaller than the effect of isospin violation on the squared mass difference $M_{xx'}^2 - M_{yy'}^2$, which is $B_0(m_x - m_y) \sim M_\pi^2$ for approximately physical mass of the quarks, since the quark mass difference is of the same order as the masses themselves.

To lowest nontrivial order in α_{EM}, we may therefore define an isospin limit by adjusting the bare masses m_u and m_d such that $M_{uu'}^2 = M_{dd'}^2$ [71]. This is not by itself a sufficient renormalization condition, however, since it does not fix the overall scale of the light quark masses. We can do that by demanding that the renormalized mass of the u and d quarks is the same as their mass in pure isospin-symmetric QCD, the theory onto which we are matching our QCD + (quenched)QED theory. Since chiral symmetry requires that the EM effects on the mass of the physical π^0 are also of order $\alpha_{\mathrm{EM}} M_\pi^2$, the pion mass in pure QCD may be taken to have the experimental mass of the π^0, $M_{\pi^0,\mathrm{expt}}$. This leads to a nonperturbative EM renormalization condition. In the QCD+(quenched)QED theory, we adjust the bare masses m_u and m_d to enforce

$$M_{uu'}^2 = M_{dd'}^2 = (M_\pi^2)^{\mathrm{QCD}} \equiv M_{\pi^0,\mathrm{expt}}^2. \quad (37)$$

We call the renormalization scheme defined by this condition the "BMW scheme." A related "Dashen scheme" has been introduced by the QCDSF Collaboration [16]. In their scheme, the masses of connected $u\bar{u}$, $d\bar{d}$ and $s\bar{s}$ mesons are all set equal at a symmetric point.

We define the mass m_l as the common u, d mass such that the charged pion in our pure QCD simulations has mass $(M_\pi)^{\mathrm{QCD}}$. Therefore, Eq. (37) may be enforced by setting

$$m_u = m_l(1 - \delta_u), \qquad m_d = m_l(1 - \delta_d) \qquad (38)$$

and choosing δ_u and δ_d so that the EM contributions to $M_{uu'}^2$ and $M_{dd'}^2$ vanish:

$$(M_{uu'}^2)^\gamma = 0 = (M_{dd'}^2)^\gamma. \qquad (39)$$

Recall that $(M^2)^\gamma$ is defined as the difference between the squared mass of the meson composed of charged quarks with that composed of uncharged quarks, but with the same renormalized masses. In Eq. (39), the EM renormalized mass is m_l, so that the neutral-quark (pure QCD) subtraction terms in the definition of $(M_{uu'}^2)^\gamma$ and $(M_{dd'}^2)^\gamma$ [see Eq. (15)] are equal to $(M_\pi^2)^{\text{QCD}}$. Thus Eq. (38) should be interpreted as defining the bare masses m_u and m_d such that the EM renormalized mass of each quark is m_l.

The condition Eq. (39) must then be rewritten in terms of $\Delta M_{uu'}^2$ and $\Delta M_{dd'}^2$, the EM effects at fixed bare mass [see Eq. (16)], which are the quantities we directly compute and fit in our simulations. With the bare mass fixed at m_u in $\Delta M_{uu'}^2$, and at m_d in $\Delta M_{dd'}^2$, the charged-quark terms in $\Delta M_{uu'}^2$ and $\Delta M_{dd'}^2$ are the same as in $(M_{uu'}^2)^\gamma$ and $(M_{dd'}^2)^\gamma$, respectively, but the neutral quark subtraction terms are different. Within the approximation that $M_{xy}^2 = B(m_x + m_y)$ in pure QCD, we may easily correct for the changed subtraction terms and rewrite Eq. (39) as

$$\Delta M_{uu'}^2(m_u) - 2Bm_l\delta_u = 0 = \Delta M_{dd'}^2(m_d) - 2Bm_l\delta_d. \qquad (40)$$

After a chiral fit to the data for $\Delta M_{xy}^2(m_x, m_y)$, we solve these conditions iteratively for δ_u and δ_d at each lattice spacing, or in the fit extrapolated to the continuum. Iteration is in principle necessary because m_u and m_d depend nonlinearly on $\Delta M_{uu'}^2$ and $\Delta M_{dd'}^2$, respectively, at fixed m_l. However, since δ_u and δ_d are $\mathcal{O}(\alpha_{\text{EM}})$, one could simply evaluate $\Delta M_{uu'}^2$ and $\Delta M_{dd'}^2$ in Eq. (40) at m_l with negligible changes to our final results. For B, we use the derivative with respect to $2m_l$ of the NLO SU(2) χPT result for M_π^2 in QCD:

$$B = \frac{(M_\pi^2)}{2m_l}\left(1 - \bar{\ell}_3 \frac{M_\pi^2}{16\pi^2 f_\pi^2}\right) \qquad (41)$$

with $\bar{\ell}_3 = 2.81(64)$ [2]. Systematic errors associated with the value of B are included in our error analysis in Sec. VII.

The residual chiral symmetry of staggered quarks implies that quark mass normalization is multiplicative. That means that once we know δ_d, we can use it to renormalize any charge-1/3 quark. In particular, in this scheme the bare strange quark mass m_S whose EM renormalized mass is m_s, the known physical strange mass in pure QCD, is

$$m_S = m_s(1 - \delta_d). \qquad (42)$$

Once the strange quark mass has been renormalized, we may compute $(M_{K^0}^2)^\gamma$, the EM effect on the neutral kaon, from

$$(M_{K^0}^2)^\gamma = \Delta M_{K^0}^2 - B_s(m_s - m_S) - B_l(m_l - m_d), \quad (43)$$

where B_s and B_l are the derivatives of $(M_K^2)^{\text{QCD}}$ with respect to m_s and m_l, respectively. Unfortunately, because a large fraction of $\Delta M_{K^0}^2$ is unphysical, and removed when constructing $(M_{K^0}^2)^\gamma$ in the renormalization step, the resulting systematic error in $(M_{K^0}^2)^\gamma$ [or equivalently ϵ_{K^0}, Eq. (5)] is relatively large ($\sim 35\%$). The result is particularly sensitive to the uncertainty in the derivative B_s.

We emphasize here two contrasting points about our renormalization scheme. On the one hand, if we keep δ_u and δ_d in Eq. (38) fixed, we can replace m_l, the average physical u, d mass, with any mass m_l', and thereby find bare masses m_u and m_d that both have renormalized masses equal to m_l'. On the other hand, for masses $m_l' > m_l$ it is *not* true that the resulting EM contributions to $(M_{uu'}^2)^\gamma$ and $(M_{dd'}^2)^\gamma$ vanish or even remain equal to each other. The condition in Eq. (40) may only be enforced at one value of m_l', and it is only when we enforce it at or near $m_l' = m_l$, as we do, that the terms we set to zero are necessarily small, of second order in a joint expansion in α_{EM} and isospin violations. As a numerical test of the latter point, we computed $(M_{uu'}^2)^\gamma$ and $(M_{dd'}^2)^\gamma$ for $m_l' = m_s/2$, i.e., for a heavy pion with mass approximately equal to the mass of the kaon. We obtain $(M_{uu'}^2)^\gamma \approx 82\,(\text{MeV})^2$ and $(M_{dd'}^2)^\gamma \approx 21\,(\text{MeV})^2$, which are non-negligible and of the same order of magnitude as our result for $(M_{K^0}^2)^\gamma$.

A final renormalization scheme that we have tried consists of simply setting to zero after the chiral fit the two LECs, κ_5 and λ_6, that are dominated by unphysical renormalization effects at $\mathcal{O}(\alpha_{\text{EM}})$ and $\mathcal{O}(\alpha_{\text{EM}}^2)$, respectively. Interestingly, this "LEC scheme" gives results for the central fit that are extremely close to those obtained from the BMW scheme: ϵ differs only by 0.03%; ϵ_{K^0}, by 0.2%. However, the results from different chiral fits vary much more with the LEC scheme than with the BMW one; this is especially true of ϵ_{K^0}, which can differ by more than 100% as we change the details of the fit, or the ranges of valence masses and charges included. For this reason we do not consider the LEC scheme further here.

D. The neutral pion

The mass of the (partially quenched) π^0 comes from the correlator

$$G_{\pi^0}(t) = \frac{1}{2}\sum_{\vec{z}}\langle[\bar{x}(t, \vec{z})\gamma_5 x(t, \vec{z}) - \bar{y}(t, \vec{z})\gamma_5 y(t, \vec{z})]$$
$$\times [\bar{x}(0)\gamma_5 x(0) - \bar{y}(0)\gamma_5 y(0)]\rangle, \qquad (44)$$

where x is an up-type valence quark with $q_x = 2/3$, y is a down-type valence quark with $q_y = -1/3$, and we work in the isospin limit $m_x = m_y$. (For simplicity, all quark masses in this subsection should be interpreted as renormalized masses.) This true π^0 has quark-line disconnected EM contributions because $q_x \neq q_y$. As mentioned in the introduction, such disconnected contributions would be costly to compute numerically, so we drop them. We define the squared mass "$M^2_{\pi^0}$" as a simple average of the squared masses coming from the two connected correlators, one for x and one for y, obtained from Eqs. (35) and (36), respectively. We can now define

$$M^2_{``\pi^0"} = \frac{1}{2}(M^2_{xx'} + M^2_{yy'}). \tag{45}$$

It is then easy to see that chiral symmetry implies that $M^2_{``\pi^0"}$ vanishes in the (two-flavor) chiral limit. That is because both $M^2_{xx'}$ and $M^2_{yy'}$ are of the form discussed in rule 2 above Eq. (23): neutral mesons composed of two different, but equally charged, quarks. The EM contributions to their masses must therefore be proportional to $e^2(m_x + m_{x'}) = e^2(m_y + m_{y'}) \propto e^2 M^2_\pi$. Chiral symmetry also implies that the EM contributions to the true $M^2_{\pi^0}$ must be proportional to $e^2 M^2_\pi$, but the reasoning is slightly different because $M^2_{\pi^0}$ is not of the form $M^2_{xx'}$ with x and x' different flavors. The spontaneously broken chiral symmetry associated with the π^0 is diagonal and is not broken explicitly by the also-diagonal quark-charge matrix Q. Hence the EM contribution to its mass must vanish as usual in the two-flavor chiral limit. We may make a rough estimate of the size of $(M^2_{\pi^0})^\gamma$ by using the chiral logarithm contribution calculated in [22], $e^2 \Delta_{\text{EM}} M^2_\pi (\ln(M^2_\pi/\Lambda^2_\chi) + 1)/ (8\pi^2 f^2)$, and taking $\Delta_{\text{EM}} = 4C/f^2 \approx 4C/f^2_\pi \cong 9900 \, (\text{MeV})^2$ from [21] and $\Lambda_\chi = m_\rho = 0.77$ GeV. This gives a magnitude of about 30 MeV^2.

The π^+ has totally different behavior from either the "π^0" or the π^0. Since its chiral symmetry is broken explicitly by the quark charges, $\Delta M^2_{\pi^+}$ is nonvanishing in the two-flavor chiral limit at leading order, and equal to $e^2 \Delta_{\text{EM}}$. At NLO, Eqs. (20) and (24) show that there are both a chiral log and an analytic contribution (from the κ_2 term) proportional to $e^2 M^2_K$. We may estimate the size of $\Delta M^2_{\pi^+}$ from the LO term, Eq. (18), and the NLO chiral logarithm contribution proportional to $e^2 M^2_K$ in the continuum limit. This gives $\Delta M^2_{\pi^+} \approx 1050$ MeV^2. Alternatively, since $\Delta M^2_{\pi^+}$ is so much larger than $\Delta M^2_{\pi^0}$, and since the u–d quark mass difference contributes so little to the π^+–π^0 splitting, we may simply use the experimental splitting $M^2_{\pi^+} - M^2_{\pi^0} = 1261$ MeV^2 as an estimate of $\Delta M^2_{\pi^+}$. Either way, it is clear that $\Delta M^2_{\pi^+} \gg \Delta M^2_{\pi^0}$.

Since both $\Delta M^2_{\pi^0}$ and $\Delta M^2_{``\pi^0"}$ are $\mathcal{O}(\alpha_{\text{EM}} M^2_\pi)$, the error due to the simulation of the "π^0" rather than the π^0 is also $\mathcal{O}(\alpha_{\text{EM}} M^2_\pi)$. We estimate the size of this systematic error in Sec. VII D.

IV. QED IN FINITE VOLUME

With the noncompact realization of QED on the lattice, which we use, it is necessary to drop some zero-modes in a finite volume in order to have a convergent path integral. In particular, the action in Coulomb gauge for the zero component of the vector potential, A_0, is $\frac{1}{2}\int(\partial_i A_0)^2$. Since the A_0 mode with spatial momentum $\vec{k} = 0$ has vanishing action, it must be dropped. Similarly, the action for the spatial components A_i is $\frac{1}{2}\int[(\partial_0 A_i)^2 + (\partial_j A_i)^2]$. Here only the mode with 4-momentum $k_\mu = 0$ must be dropped, and that is what we do. This version of QED in FV was first introduced by Duncan, Eichten and Thacker [8]; following the nomenclature in Borsanyi et al. [72], we call the resulting theory QED$_{TL}$. Summarizing, QED$_{TL}$ is defined in Coulomb gauge by

$$A_0(k_0, \vec{k} = 0) = 0, \quad \forall k_0,$$
$$\vec{k} \cdot \vec{A}(k_0, \vec{k}) = 0, \quad \forall k_0, \vec{k}, \quad [\text{QED}_{TL}]$$
$$\vec{A}(k_0 = 0, \vec{k} = 0) = 0. \tag{46}$$

Hayakawa and Uno, in their calculation of EM FV effects in χPT [73], introduce a different FV action, called QED$_L$, in which they drop all modes with $\vec{k} = 0$, both for A_0 and for A_i. Again in Coulomb gauge, QED$_L$ is defined by

$$A_0(k_0, \vec{k} = 0) = 0, \quad \forall k_0,$$
$$\vec{k} \cdot \vec{A}(k_0, \vec{k}) = 0, \quad \forall k_0, \vec{k}, \quad [\text{QED}_L]$$
$$\vec{A}(k_0, \vec{k} = 0) = 0, \quad \forall k_0. \tag{47}$$

The difference between Eqs. (46) and (47) is solely in the last line of each, in the treatment of $\vec{A}$ when $\vec{k} = 0$. This difference implies that the FV effects in the MILC calculations are different from those computed in Ref. [73].

To make explicit the difference between our set-up (QED$_{TL}$) and that of Ref. [73] (QED$_L$), we give the Coulomb gauge photon propagator in each case:

$$\mathcal{D}_{ij}(k) \equiv \langle A_i(k) A_j(-k) \rangle$$
$$= \begin{cases} \frac{1}{k^2}\left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2}\right), & \vec{k} \neq 0; \\ 0, & \vec{k} = 0. \end{cases} \quad [\text{QED}_L] \tag{48}$$

$$\mathcal{D}_{ij}(k) \equiv \langle A_i(k) A_j(-k) \rangle$$

$$= \begin{cases} \dfrac{1}{k^2}\left(\delta_{ij} - \dfrac{k_i k_j}{\vec{k}^2}\right), & \vec{k} \neq 0; \\[2ex] \dfrac{1}{k^2}\delta_{ij}, & \vec{k}=0, k_0 \neq 0; \quad [\text{QED}_{TL}] \\[2ex] 0, & \vec{k}=0, k_0=0. \end{cases}$$

$$(49)$$

$$\mathcal{D}_{00}(k) \equiv \langle A_0(k) A_0(-k) \rangle$$

$$= \begin{cases} \dfrac{1}{\vec{k}^2}, & \vec{k} \neq 0; \\[2ex] 0, & \vec{k}=0. \end{cases} \quad [\text{QED}_L \quad \text{and} \quad \text{QED}_{TL}] \quad (50)$$

The violation of Gauss's law induced by the absence of the $\vec{k}=0$ A_0 mode makes it possible to have net charges on a FV torus with periodic boundary conditions [73]. But Gauss's law has no implications for the spatial modes A_i, so does not distinguish between Eqs. (48) and (49).

Borsanyi *et al.* [72] have independently studied QED in FV, using both the QED_L and QED_{TL} versions. They define QED_L by

$$\sum_{\vec{x}} A_{\mu,x_0,\vec{x}} = 0, \quad \forall x_0, \mu. \quad (51)$$

This is in fact a partial gauge specification, because spatially-independent, but time-dependent, gauge transformations would violate the $\mu=0$ condition $A_0(k_0,\vec{k}=0)=0$ (written here in momentum space). One can bring any EM gauge field that satisfies Eq. (51) into Coulomb gauge, as was assumed in writing Eq. (48). The necessary gauge transformation is, in momentum space:

$$A_\mu(k_0,\vec{k}) \to A_\mu(k_0,\vec{k}) - ik_\mu \Lambda(k_0,\vec{k}) \quad (52)$$

$$\Lambda(k_0,\vec{k}) = \begin{cases} -\dfrac{i\vec{k}\cdot\vec{A}(k_0,\vec{k})}{\vec{k}^2}, & \vec{k} \neq 0; \\[2ex] 0, & \vec{k}=0. \end{cases} \quad (53)$$

Borsanyi *et al.* define QED_{TL} by

$$\sum_x A_{\mu,x} = 0, \quad \forall \mu. \quad (54)$$

Unlike Eq. (51), this definition is gauge invariant, as can be immediately seen from Eq. (52). Equation (54) can be put into a special Coulomb gauge that satisfies Eq. (46) by the transformation:

$$\Lambda(k_0,\vec{k}) = \begin{cases} -\dfrac{i\vec{k}\cdot\vec{A}(k_0,\vec{k})}{\vec{k}^2}, & \vec{k} \neq 0; \\[2ex] -\dfrac{iA_0(k_0,\vec{k})}{k_0}, & \vec{k}=0, k_0 \neq 0; \\[2ex] 0, & k_0=0, \vec{k}=0. \end{cases} \quad (55)$$

Thus the two definitions of QED_{TL}, Eqs. (46) and (54), are equivalent.

V. FINITE VOLUME EFFECTS IN CHIRAL PERTURBATION THEORY

Before discussing the FV calculations, it is important to make some remarks on the literature. The first calculation for the FV EM effects on pseudoscalar meson masses that we are aware of was performed by Hayakawa and Uno [73]. They worked in QED_L exclusively and used χPT at one loop. Again for QED_L, Davoudi and Savage [74] showed, using nonrelativistic effective field theory, that the leading $1/L$ and $1/L^2$ terms found in Ref. [73] are in fact universal, independent of the internal structure of the particle of interest. They related higher-order terms directly to the structure, parametrized in terms of EM multipole moments and polarizabilities, and extended the calculations to include spin-1/2, as well as spin-0, particles. Shortly after Ref. [74] appeared, Borsanyi *et al.* [72], and our own work [14] independently completed the FV calculations for QED_{TL}. Where they overlap, the results of Ref. [72] and Ref. [14] agree. However we have focused only on pseudoscalar mesons and have not worked out the analytic form of the asymptotic expansions in powers of L and T, which Ref. [72] does very nicely for both QED_{TL} and QED_L. Further, Borsanyi *et al.* found a discrepancy with the results of Ref. [74] for the first nonuniversal $(1/L^3)$ terms for spin-1/2. The issue involved is in fact quite subtle, but it seems to have been resolved [75,76] in favor of the result in [72].

In FV, defined here by spatial extent L and temporal extent T, the momentum components take on discrete values

$$k_i = \frac{2\pi n_i}{L}, \qquad k_0 = \frac{2\pi n_0}{T}, \quad (56)$$

with n_i ($i=1,2,3$) and n_0 integers. Through NLO in χPT, the meson mass squared in FV may then be calculated simply by replacing the momentum integrals in the diagrams of Fig. 5 by sums:

$$\int \frac{d^4 k}{(2\pi)^4} \to \frac{1}{L^3 T} \sum_{n_\mu}. \quad (57)$$

Because the Feynman diagrams are divergent, it is as usual convenient to perform the renormalization in infinite volume, and, in FV, calculate only the difference between the momentum sums and the integrals. This difference, if treated carefully, is finite and does not require renormalization.

We thus stipulate that the EM effect ΔM_{xy}^2 defined in Eq. (16) is the appropriately renormalized infinite-volume result, and write

$$(\Delta M_{xy}^2)_{\text{FV}} = \Delta M_{xy}^2 + \delta_{\text{FV}}, \qquad (58)$$

$$\delta_{\text{FV}} = \delta_{\text{FV}}^{\text{meson}} + e^2 q_{xy}^2 m^2 \delta_{\text{FV}}^{\gamma}(mL, mT) \qquad (59)$$

where δ_{FV} is the complete NLO FV correction, $\delta_{\text{FV}}^{\text{meson}}$ is the contribution from the meson tadpole, Fig. 5(c), and $\delta_{\text{FV}}^{\gamma}$ is the contribution from photon loops, Figs. 5(a) and 5(b). The factors $e^2 q_{xy}^2 m^2$ have been taken out of $\delta_{\text{FV}}^{\gamma}$ for convenience. For notational simplicity in this section, m will denote the tree-level mass of the meson of interest in the absence of EM; ultimately we put $m^2 = \chi_{xy,5}$ in the results. With the factor of m^2 removed, $\delta_{\text{FV}}^{\gamma}$ is dimensionless, and hence is a function of mL and mT (or T/L) only, rather than m, L, T separately.

The FV effects from the meson tadpole come from pions that loop around the volume, and hence the effect is suppressed by a factor of $\exp(-mL)$. Because of this suppression, $\delta_{\text{FV}}^{\text{meson}}$ is of negligible size on our ensembles, $\lesssim 0.2\%$. However, since the calculation of the effect is completely standard, it is straightforward to include it. In the notation of Ref. [77], we just have to make the substitution $\ln(m^2/\Lambda^2) \to \delta_1(mL)$, where δ_1 is a sum over Bessel functions, to obtain the FV correction. From Eq. (20), this gives

$$\delta_{\text{FV}}^{\text{meson}} = \frac{-2e^2 \Delta_{\text{EM}}}{16\pi^2 f^2}\left(\frac{1}{16}\right)\sum_{\sigma,b}[q_{x\sigma}q_{xy}\chi_{x\sigma,b}\delta_1(\sqrt{\chi_{x\sigma,b}}L)$$
$$- q_{y\sigma}q_{xy}\chi_{y\sigma,b}\delta_1(\sqrt{\chi_{y\sigma,b}}L)]. \qquad (60)$$

In contrast to the meson tadpole effects, the FV effects from photon diagrams, parametrized by $\delta_{\text{FV}}^{\gamma}$, are large: $\sim 5\%$–20%, depending on the ensemble and valence masses. Since the results are nontrivial, we describe the calculation in some detail, starting with the sunset diagram, Fig. 5(b). We work in Coulomb gauge, choose the external meson to be at rest $[p = (p_0, 0, 0, 0)]$, and route the loop momentum k along the interior meson line, with momentum $p - k$ on the photon line.[7] Because spatial $\vec{p} = 0$ and $k_i \mathcal{D}_{ij} = 0$ (for both QED$_L$ and QED$_{TL}$), only the 00 component of the photon propagator contributes to the sunset diagram. This diagram's contribution to the self-energy then has integrand (summand)

$$\mathcal{I}_{\text{s}} = -\frac{k_0^2 + p_0^2}{\vec{k}^2(k^2 + m^2)}, \qquad [\vec{k} \neq 0], \qquad (61)$$

where we have omitted an overall factor of $e^2 q_{xy}^2$. A linear term in k_0 in the numerator has been dropped because k_0 and $-k_0$ contributions cancel for both the infinite-volume integral and the FV sum.

Since $\mathcal{I}_{\text{s}}$ goes to a constant as $k_0 \to \infty$, the difference between the sum and integral over k_0 (not to mention the integral itself) is divergent, so the FV effect from this diagram alone (in Coulomb gauge) is not well defined. However, once this diagram is combined with the photon tadpole, the problem goes away. What is needed is in fact only the $\mathcal{D}_{00}$ contribution to the tadpole, which has the integrand $1/\vec{k}^2$. Adding this to Eq. (61), gives

$$\mathcal{I}_{\hat{\text{s}}} = \frac{\vec{k}^2 + m^2 - p_0^2}{\vec{k}^2(k^2 + m^2)}, \qquad [\vec{k} \neq 0], \qquad (62)$$

where the "hat" on the subscript s indicates that the sunset diagram has been modified by a piece of the photon tadpole. It is useful to keep the rest of the tadpole separate, because it gives different contributions in the QED$_L$ and QED$_{TL}$ cases, unlike $\mathcal{I}_{\hat{\text{s}}}$.

The FV effect on the self energy coming from Eq. (62) is

$$m^2 \delta_{\text{FV}}^{\hat{\text{s}}}(p_0/m, mL, mT) = \frac{1}{L^3 T}{\sum_{k_0,\vec{k}}}' \mathcal{I}_{\hat{\text{s}}} - \int \frac{d^4 k}{(2\pi)^4}\mathcal{I}_{\hat{\text{s}}}, \qquad (63)$$

where the prime on the summation symbol means that the $\vec{k} = 0$ term is dropped, but there is no restriction on k_0. As in Eq. (59), we take out a factor of m^2 to make $\delta_{\text{FV}}^{\hat{\text{s}}}$ dimensionless.

From Eq. (48), the remaining (spatial) components of the photon tadpole in QED$_L$ give the integrand and corresponding FV effect

$$\mathcal{I}_{\text{t,QED}_{\text{L}}} = \frac{2}{k^2}, \qquad [\vec{k} \neq 0], \qquad (64)$$

$$m^2 \delta_{\text{FV}}^{\text{t,QED}_L}(mL, mT) = \frac{1}{L^3 T}{\sum_{k_0,\vec{k}}}' \mathcal{I}_{\text{t,QED}_L} - \int \frac{d^4 k}{(2\pi)^4}\mathcal{I}_{\text{t,QED}_L}. \qquad (65)$$

In QED$_{TL}$, there is an extra contribution coming from the nonzero value of $\mathcal{D}_{ij}$ when $\vec{k} = 0$ but $k_0 \neq 0$, see Eq. (49).

$$m^2 \delta_{\text{FV}}^{\text{t,QED}_{TL}}(mL, mT)$$
$$= m^2 \delta_{\text{FV}}^{\text{t,QED}_L}(mL, mT) + m^2 \delta_{\text{FV}}^{\text{t,+}}(mL, mT). \qquad (66)$$

$$\delta_{\text{FV}}^{\text{t,+}}(mL, mT) = \frac{1}{m^2 L^3 T}\sum_{k_0 \neq 0}\frac{3}{k_0^2}$$
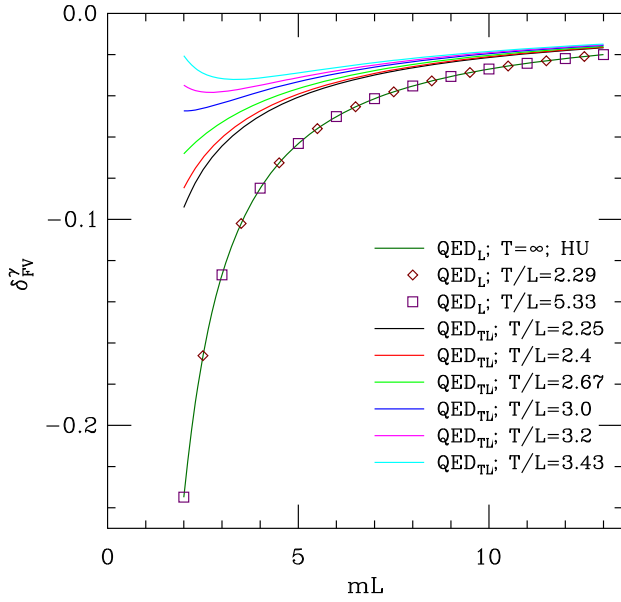$$= \frac{mT}{(mL)^3}\frac{3}{4\pi^2}2\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{mT}{4(mL)^3}, \qquad (67)$$

---

[7]The final result is of course independent of the momentum routing. However, when T is not infinite, there are interesting subtleties, which can lead to apparent routing-dependence if treated incorrectly. See the Appendix for a discussion.

FIG. 6. The FV effect from photon diagrams, δ_{FV}^{γ} for QED_L and QED_{TL}, as a function of mL. In the QED_L case, the dark green line shows the result when $T = \infty$ from Hayakawa and Uno [73], while the dark red diamonds and purple squares show our evaluation at $T/L = 2.29$ and $T/L = 5.33$, respectively. For QED_{TL}, the lines give our results for six values of T/L ranging between 2.25 and 3.43, which are the values relevant to the bulk of our data. The numerical errors in the points and lines are too small to be seen on this scale.

where we have used the well-known result $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$ [78].

When T is infinite, we can obtain the correction to the meson mass-squared by evaluating the self energy at $p_0 = im$. For finite T, however, this prescription is not obviously correct, and indeed is wrong in some cases. Here, we will simply assume that we may use the prescription, and leave it to the Appendix to explain the point in detail and show that plugging $p_0 = im$ into the integrand in Eq. (62) gives the desired answer. The complete contributions from the photon diagrams to the FV effect on the meson mass-squared are then

$$\delta_{FV}^{\gamma,QED_L}(mL, mT) = \delta_{FV}^{\hat{s}}(i, mL, mT) + \delta_{FV}^{t,QED_L}(mL, mT) \tag{68}$$

$$\delta_{FV}^{\gamma,QED_{TL}}(mL, mT) = \delta_{FV}^{\gamma,QED_L}(mL, mT) + \frac{mT}{4(mL)^3}. \tag{69}$$

It now is necessary only to evaluate the difference of sums and integrals given in Eqs. (63) and (65). This can be done straightforwardly using an importance-sampling integration program such as VEGAS [79]. The sum may be treated as an integral by defining the "finite-volume integrand" at the arbitrary point k as the average of the infinite-volume integrand at the 16 corners of the FV hypercube containing

k, weighted appropriately by the distances in each direction to the corners. For example, if $\tilde{k}$ is the closest point in the sum "below" k ($\tilde{k}_\mu < k_\mu$) then the weight of the integrand at $\tilde{k}$ is $[1 - (k_0 - \tilde{k}_0)T/(2\pi)] \prod_{i=1}^{3}[1 - (k_i - \tilde{k}_i)L/(2\pi)]$. When a corner is a special point (e.g., $\vec{k} = 0$, k_0 arbitrary) that should be dropped from the sum, we simply put in 0 for the integrand there. One could also use the value at the closest corner of the FV hypercube rather than the weighted average, but the resulting integrand has discontinuities on the midplanes of the hypercube, and the numerical integration therefore has larger errors.

We have checked that our result for δ_{FV}^{γ}, the sum of the sunset and the photon tadpole diagrams, agrees with that of Ref. [73] in the QED_L case. In Fig. 6, we plot in dark green the result calculated from the results of Hayakawa and Uno [73], and superimpose points calculated by us at representative values of mL. Hayakawa and Uno work at infinite T, whereas our points have been computed at $T/L = 2.29$ and $T/L = 5.33$. It is clear that for such values of T/L the finite-T effects are negligible in QED_L. (See the Appendix for further discussion.)

The difference in the QED_{TL} case is the extra term $\delta_{FV}^{t,+}$ in Eq. (67). Figure 6 also shows our QED_{TL} results for ranges in values of mL and T/L that cover all of our data used in the final analysis; data with more extreme T/L values (≥ 4 and ≤ 1.6) are used later in this section in testing the applicability of our formulas.

Unlike Ref. [73] and the present calculation, Davoudi and Savage [74] and Borsanyi *et al.* [72] do not compute the FV effects in the context of χPT, but instead work first with the universal terms that describe a pointlike particle, and then consider corrections coming from the particle structure. Aside from the contribution from the meson tadpole, Eq. (60), which is suppressed by $\exp(-mL)$, the one-loop χPT calculation is in fact identical to the pointlike approximation of Refs. [72,74] because there are no corrections to the photon-meson vertices or internal meson lines in Figs. 5(a) and 5(b). In QED_{TL}, for pointlike mesons, Ref. [72] finds

$$\delta_{FV}^{\gamma} \underset{T,L\to\infty}{\sim} -\frac{\kappa}{4\pi mL} - \frac{\kappa}{2\pi(mL)^2} + \frac{mT}{4(mL)^3}, \tag{70}$$

where the last term is what we call $\delta_{FV}^{t,+}$, Eq. (67), and the other terms come from the asymptotic expansion of δ_{FV}^{γ} in QED_L. The constant κ is defined in Ref. [72] by

$$\kappa \equiv \int_0^{\infty} \frac{d\lambda}{\lambda^{3/2}} \{\lambda^{3/2} + 1 - [\theta_3(0, e^{-\frac{\pi}{\lambda}})]^3\}, \tag{71}$$

where $\theta_3(u, q) = \sum_{n=-\infty}^{+\infty} q^{n^2} e^{i2nu}$ is a Jacobi theta function. By numerical integration, one finds $\kappa \approx 2.8373$. An equivalent definition of κ in Ref. [73] is
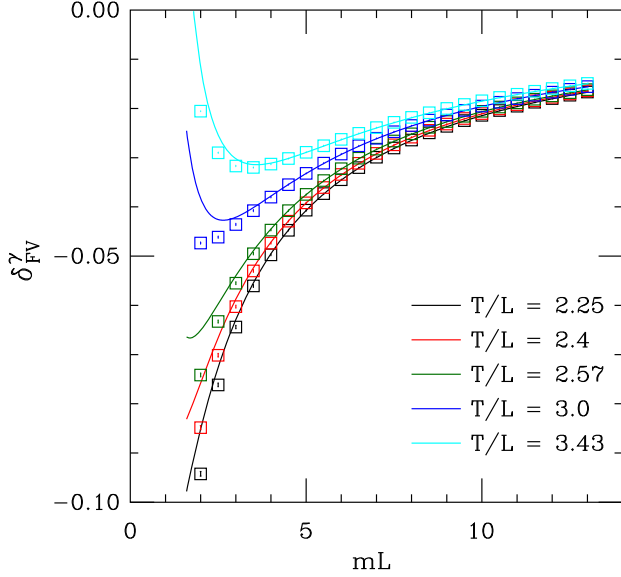
FIG. 7. A comparison of the full FV effect for QED_{TL}, coming from one-loop photon diagrams with a pointlike meson, and the corresponding asymptotic forms determined in Ref. [72], for various values of T/L. The squares show our calculations of the full effect, while the lines are the asymptotic forms. The numerical errors in the points are small and are just barely visible in some of the points at the left.

$$\kappa = \int_0^\infty \frac{d\lambda}{\lambda^2} \{\lambda^{3/2} + 1 - [\theta_3(0, e^{-\frac{\pi}{\lambda}})]^3\}. \qquad (72)$$

The equivalence of Eqs. (71) and (72) follows from the identity [80]

$$\theta_3(0, e^{-\pi x}) = \frac{1}{\sqrt{x}} \theta_3(0, e^{-\pi/x}), \qquad (73)$$

which can easily be proved using the Poisson summation formula.
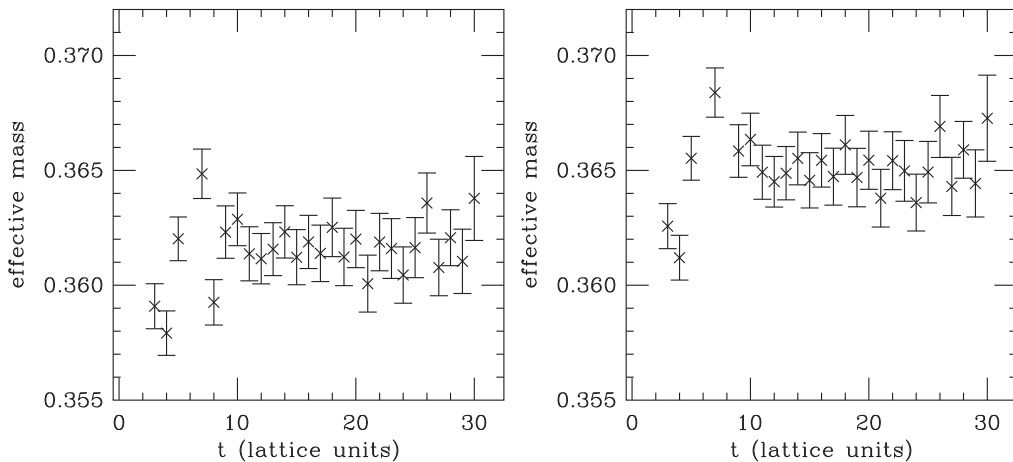
In Fig. 7, we compare our results for QED_{TL} with the asymptotic form Eq. (70). For $mL \geq 3.8$, which describes the unitary points in our data used in the final analysis, the differences with the asymptotic form are negligible. However, a few valence points in that analysis have $mL \gtrsim 2.9$, for which the differences ($\lesssim 6\%$) are important to include. In our test of FV effects described later in the section, we have points as low as $mL = 2.7$ and aspect ratio of $T/L = 5.33$ for which the differences are a bit bigger, $\approx 7\%$. For convenience, we use our full results everywhere in the analysis, even where the differences with the asymptotic form are negligible.

We emphasize here that the term $\delta_{FV}^{t,+} = mT/(4(mL)^3)$ in Eq. (67) indicates that the large-volume limit is rather subtle in QED_{TL}. The result is acceptable if the limit $L \to \infty$ is taken before $T \to \infty$, or if the limits are taken together at fixed aspect ratio T/L, but not if the limit $T \to \infty$ is taken first. In other words, the QED_{TL} set-up is not well defined in finite spatial volume at zero temperature. This fact has also been pointed out by Borsanyi et al. [72]. They make the further point that QED_{TL} violates reflection positivity because the constraint required to set the single $k_\mu = 0$ mode of A_i to zero involves the square of the integral over all space-time of A_i. Although many actions used in lattice QCD violate reflection positivity, one might worry that in this case the violation leads to problems with defining or isolating the lowest states in correlation functions. Reference [72] did have problems from close excited states in extracting masses in pure quenched QED_{TL}. In our QCD plus quenched QED_{TL} simulations, however, this does not seem to be a problem. As illustrated in Fig. 8, we find no significant differences between the qualities of plateaus in correlation functions in $QCD + QED_{TL}$ vs those for QCD alone. The example shown is for a putative "worst case" in our data because the aspect ratio $T/L = 5.33$ has the largest value, and L is the smallest. See also the plots for our ensemble with $a \approx 0.045$ fm and $T/L = 3$, shown in Fig. 2.

FIG. 8. Effective mass plots for a "K^+ meson" in pure QCD (left) and in $QCD +$ quenchedQED_{TL} (right). The data is from the ensemble listed first in Table I, with $a \approx 0.12$ fm, $L/a = 12$ and $T/a = 64$. The valence masses are 0.01 and 0.04.
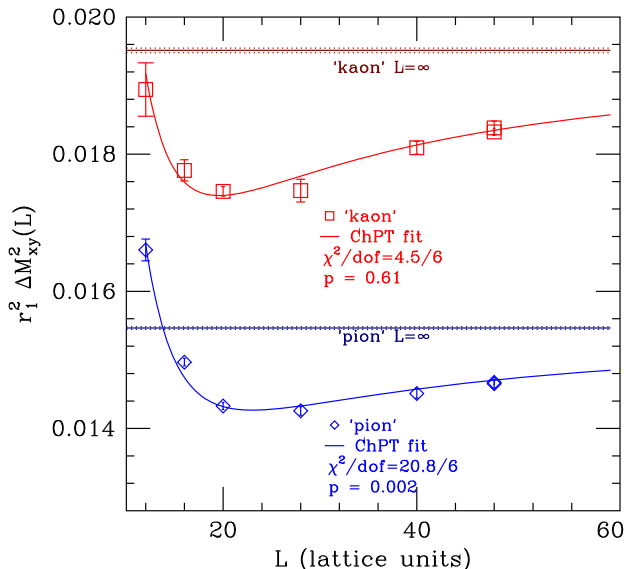
FIG. 9.   Finite volume effects at $a \approx 0.12$ fm and $am'_l = 0.01$, $am'_s = 0.05$ as a function of spatial lattice length L for two different meson masses: a unitary 'pion' (blue) with degenerate valence masses $m_x = m_y = m'_l$, and a 'kaon' (red) with valence masses $m_x = m'_l$ and $am_y = 0.04$, close to the physical strange quark mass. The fit lines are to our FV form for QED_{TL} (omitting the negligible meson tadpole term), and have one free parameter each, the infinite volume value (shown by horizontal solid lines with dotted lines for errors).

Again, no significant differences in plateau quality between $QCD + QED_{TL}$ and pure QCD are visible.

Despite the fact that we have not found any evidence of problems due to the lack of reflection positivity in QED_{TL}, the reader may wonder why we did not just use QED_L or a massive-photon infrared regulator [81], both of which are reflection-positive. The reason is straightforward: When this project was begun [27], and by the time most of the numerical computations were completed [12], the issues with QED_{TL} were not known. We simply followed the QED_{TL} approach of the original paper on the subject, Ref. [8]. The fact that QED_{TL} has smaller FV corrections than QED_L in the relevant range of parameters (as seen in Fig. 6) is a nice accidental benefit of our choice of QED_{TL}, but it was also not known when this project started and therefore had no influence on the choice.

To test our understanding of the FV effects, we have generated ensembles with a wide range of spatial sizes at $\beta = 6.76$ ($a \approx 0.12$ fm) with sea-quark masses $m'_l = 0.01$ and $m'_s = 0.05$ (see Table I). In Fig. 9, we show fits, for two different meson masses on these ensembles, to our calculated FV correction, given by Eqs. (58) and (59), with $\delta^\gamma_{FV} = \delta^{\gamma,QED_{TL}}_{FV}$, Eq. (69). We neglect the meson tadpole term in Eq. (58) for convenience, since its effect is not visible on this scale. This means that the FV correction used here is the same as in the pointlike approximation for the mesons. The shape of the fit curves are completely

determined by the FV calculation; the only free parameter in each fit is the overall height of the curve given by the value in infinite volume. The theory gives a good description of the data, and we use it to correct the data for FV effects. We estimate the remaining systematic error associated with FV effects in Sec. VII B.

One can now understand why it was difficult to observe FV effects directly in the data set available in Ref. [12]. At that time, we had only the $L = 20$ and $L = 28$ ensembles to compare. From Fig. 9, one sees that the minima of the curves are in this region of L or close to it, and therefore the difference expected between these volumes is small compared to the statistical errors in the data.

VI. CHIRAL-DISCRETIZATION FITS AND CHIRAL-CONTINUUM EXTRAPOLATIONS

In this section, we first discuss the quantities that have been determined from pure QCD computations, and are used here as inputs to the chiral-discretization fits. We then show (a small subset of) the data we fit, both before and after FV corrections. Finally, we describe the fits themselves.

A. Inputs

In addition to the lattice values of r_1/a that set the relative scales, we need other lattice-dependent quantities as inputs to the QCD + QED calculations. Table V lists the values of these quantities for one ensemble at each of our (approximate) lattice spacings. The first three columns serve to identify the ensembles. Columns four and five give the light and strange physical quark masses in r_1 units, which are determined from chiral fits to pure QCD lattice data [4]. These masses are "physical" in the sense that they have been determined by demanding that the π and K mesons take their (isospin-averaged) experimental values in absence of EM.[8] They are, however, bare masses, in that no renormalization (perturbative or otherwise) has been applied.

Two errors are shown for the masses. The first is the systematic error coming from the chiral extrapolation. It is determined by comparing the results of fits that include chiral logarithms through NNLO (plus higher-order analytic terms) and those that include the chiral logarithms only through NLO. Other changes in the fits give similar estimates of the errors. The second error in the masses comes from the uncertainty in the absolute scale, i.e., the error in the physical value of r_1.

---

[8]There is an apparent circularity here, in that we are computing in this paper the EM effects on the π and K masses. In practice, we have used earlier, phenomenological estimates of EM effects (see Ref. [82]) to remove them at this stage. We can iterate to make the calculation self-consistent, but it is unnecessary, because the EM effects make only a small change in the estimates of the strange and isospin-averaged light masses.
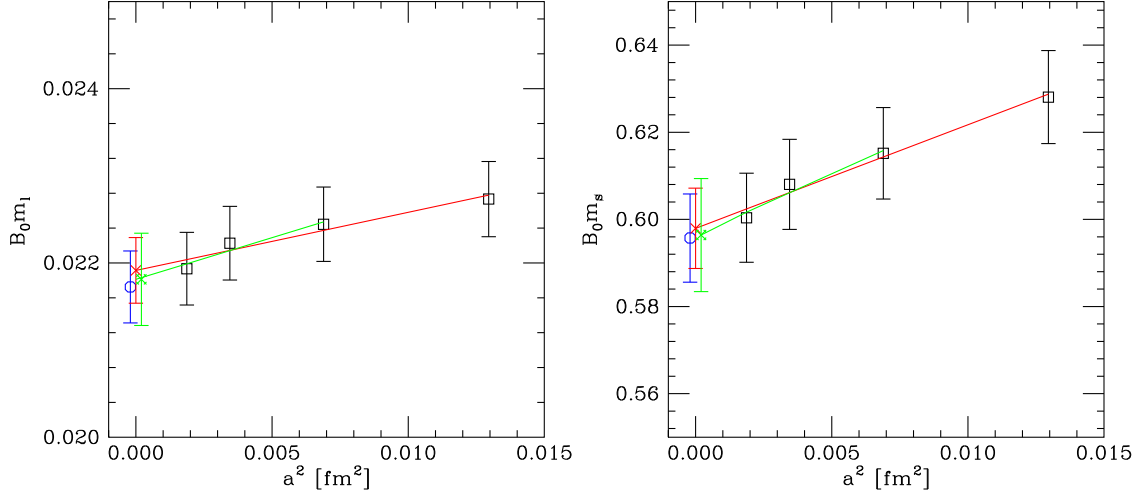
FIG. 10.   Products $B_0 m_l$ (left) and $B_0 m_s$ (right) vs a^2. The black squares show our data at nonzero lattice spacing, while the blue octagons show our continuum values. Two simple linear extrapolations are shown for comparison. The red line in each plot is a fit to all four lattice spacings, and the red cross is its extrapolation. Similarly, the green line and fancy cross in each plot comes from a fit that omits the coarsest lattice spacing.

In the final row of the table, "Cont." stands for "continuum." It is convenient for us to view the continuum not as the $\beta \to \infty$, $a \to 0$ limit, but as another ensemble with fixed β and a, in which all discretization effects have been extrapolated away. In other words, we view the continuum as a lattice with a perfect action. This allows us to continue to employ bare lattice masses to describe the physical point, just as we do at nonzero lattice spacing. Here we have chosen the continuum to have $\beta = 7.08$, the same as the $a \approx 0.09$ ensemble with simulation masses $m'_l/m'_s = 0.0031/0.031$. The scale of the two is however slightly different, since extrapolating away the discretization effects changes the estimates of the physical quark masses. This in turn affects the r_1/a value, which is adjusted to be at physical masses. The 0.0031/0.031 ensemble has $r1/a = 3.755$, while $r1/a = 3.744$ for the "continuum," a 0.3% difference. This difference shows that the discretization effects in our mass-independent scale-setting scheme are small.

The LEC B_0 is given in column six of Table V. It is obtained from a fit of the squared masses of the Goldstone (taste ξ_5) mesons to Eq. (13). The fit is performed for each lattice spacing over the full range of meson masses that enter this analysis. This LO result is used for the meson masses in the NLO (and higher-order) expressions in ΔM_{xy}^2, Eq. (17); $\Delta_{\mathrm{LO}} M_{xy}^2$ is of course mass independent. Like the quark masses, the B_0 values shown here are bare (unrenormalized).

Both B_0 and the quark masses need to be renormalized before we can properly compare values at different lattice spacings and extrapolate to the continuum. We use the one-loop renormalization from [68], Eq. (31), to do the extrapolation. As discussed in the context of EM mass renormalization in Sec. III C, this means that there are substantial errors from renormalization affecting the continuum values of B_0, m_l, and m_s in Table V. This is true even

though we take out the renormalization factors, defined for the continuum to be the same as those of the $\beta = 7.08$, $m'_l/m'_s = 0.0031/0.031$ ensemble. Such errors would be important if we wanted to extract quark masses or B_0 in a continuum scheme such as $\overline{\mathrm{MS}}$. However the renormalization errors are irrelevant here and not included in Table V because only the renormalization-invariant products $B_0 m_l$ and $B_0 m_s$ enter into the results from our χPT fits. This is illustrated in Fig. 10, which shows these products computed from the values in Table V and compares the continuum values (blue octagons) to the values that would have been obtained by linear extrapolation in a^2 from all four of our lattice spacings (red lines and crosses) or the three spacings with $a \lesssim 0.09$ fm (green lines and fancy crosses). Although the continuum values of m_l, m_s, or B_0 were not obtained from such extrapolations,[9] the figure shows that the products have small discretization errors and smooth behavior with a^2. Renormalization factors, along with their large one-loop errors, cancel out.

The values of quark masses and B_0 shown in the table may be used for any ensemble in the same group of approximate lattice spacings as the ones listed. For example, for the $a \approx 0.06$ fm, $\beta = 7.47$, $m'_l/m'_s = 0.0036/0.018$ ensemble, which is not listed in Table V, one should just use the values listed for the $a \approx 0.06$ fm, $\beta = 7.46$, $m'_l/m'_s = 0.0018/0.018$ ensemble. The small changes in β, and hence in lattice scale, result in even smaller changes in discretization effects and renormalization constants. Thus even though the quantities shown are unrenormalized, their differences among a group of ensembles with the same

---

[9]The physical quark masses come from two-loop chiral fits described in [4], while B_0 comes from linear extrapolation of the values in Table V after renormalization.

group of approximate lattice spacings are negligible. Note from the table that even when the lattice scale changes from ≈ 0.12 fm to ≈ 0.045 fm, the changes in the masses in r_1 units are less than 3% and those in B_0 are less than 7%.

The final needed input for our fits are the values of the taste-splittings $a^2 \Delta_b$ in Eq. (13). Table V gives these splittings in r_1 units. For unlisted ensembles, the explicit factor of a^2 in the splittings results in changes of a few percent from the listed ones. We include these changes in our fitting routines, even though they are smaller than the current statistical errors on the splittings. One can make the adjustment simply by multiplying the listed value by the ratio $(r_1/a)^2_{\text{listed}}/(r_1/a)^2_{\text{unlisted}}$, with the r_1/a values taken from Table I. For the two $a \approx 0.06$ fm ensembles mentioned in the previous paragraph, the adjustment in splittings is about 2%.

B. FV corrections to our data

Figure 11 shows a small subset of our data for ΔM^2_{xy}, plotted as a function of the meson mass, $r_1^2 M^2_{xy}$, before and after correction for FV effects. The subset consists of charge $\pm e$ unitary or nearly-unitary points, as described in more detail in the figure caption. Because the correction due to photon diagrams is proportional to M^2_{xy}, see Eq. (59), the absolute FV effect is larger for kaonlike points (right-hand half of the plot) than for pionlike points (left-hand half). The correction ranges from 0.0013 to 0.0021 for kaons and 0.0005 to 0.0009 for pions. Even the fractional correction is generally larger for kaons than for pions since the LO contribution to ΔM^2_{xy} itself is the mass-independent quantity Δ_{EM} [the Dashen term, Eq. (14)], which has no FV correction. The correction varies from 10% to 16% for kaons, and from 6% to 12% for pions.

Strictly speaking, the full FV correction to ΔM^2_{xy} depends on the chiral fit, because the FV effect of the meson tadpole $\delta^{\text{meson}}_{\text{FV}}$, Eq. (60), depends on the fit parameter Δ_{EM}. However, this dependence would not be visible in Fig. 11, because the exponentially suppressed meson tadpole FV corrections are very small compared with those from the photon diagrams, which are independent of the fit parameters.

Because the FV corrections depend, at least in principle, on the parameters of the fit, we fit uncorrected (raw) data for ΔM^2_{xy} to a chiral fit form that includes the FV NLO adjustment δ_{FV} in Eq. (58). However, we will always present the results of chiral fits after *a posteriori* correction to infinite volume of both the data and the fit lines. This allows us to present results obtained from different volumes in an accessible fashion, and also facilitates comparison to experiment.

C. Fits

We fit various subsets of the data to the chiral forms described in Sec. III B, with the FV corrections appropriate
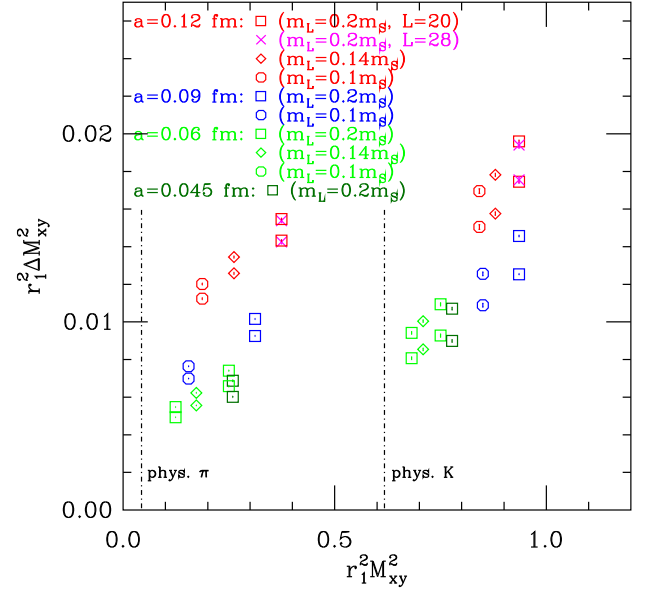

FIG. 11. Finite-volume corrections to ΔM^2_{xy} for a small subset of our data vs M^2_{xy} itself, where both quantities are expressed in r_1 units. For each pair of points with the same color and symbol, the lower point shows the raw datum, while the upper shows the result after correction for FV effects, i.e., in infinite volume. Colors and symbols identify the lattice spacing and light sea-quark mass, and (in one case) the spatial lattice size, as shown in the legend. Points in the left-hand cluster are pionlike and unitary ($m_x = m_y = m'_l$), while those in the right-hand cluster are kaonlike and almost always unitary ($m_x = m'_l$, $m_y = m'_s$). The exceptions are kaonlike points for $a \approx 0.12$ fm, which have $m_y = 0.8m'_s$, which is closer to the physical strange mass than m'_s itself. The locations of the physical pion and kaon masses are indicated by the vertical dot-dashed black lines. All points shown are for mesons with charge $\pm e$.

to each ensemble added on. The chiral forms include discretization effects, so from now on we will refer to the fits as chiral-discretization fits. The complete data set, which includes $a \approx 0.12$, 0.09, 0.06, and 0.045 fm ensembles, and quark charges 0, $\pm e/3$, $\pm 2e/3$, $\pm e$, $\pm 4e/3$, $\pm 2e$, is based on a total of 11,654 configurations and has 2978 data points for ΔM^2_{xy}. Without the $a \approx 0.12$ fm ensembles, which are often omitted from our fits, the data set has 2166 points based on 6040 configurations. Because points from the same ensemble but with different valence masses and/or quark charges are highly correlated, and because the number of points is not very much less than the number of configurations, the full covariance matrix is nearly singular and has many poorly determined low eigenvalues. Fits with acceptable p values to the whole data set are therefore not possible. However, once the data is thinned to a more reasonable number of points in comparison to the number of configurations (~ 250 to 450 points), acceptable fits are possible. For fits with up to about 350 points, we are able to include the complete covariance matrix, with no

modifications. For fits with more points than that, statistical and roundoff errors typically lead to a small number of negative eigenvalues (up to about 10) in the covariance matrix. We remove such eigenvalues with SVD when finding the inverse covariance matrix used in the fitting procedure. For every dropped eigenvalue, we reduce the number of d.o.f. by 1 in computing the p value of the fit. Our central fit, with 264 points, has no negative (and therefore no dropped) eigenvalues; the alternative fits used in estimating the errors of the chiral-continuum extrapolation do include some with dropped negative eigenvalues.

When determining the p value of a given fit, we take into account the fact that the sample covariance matrix is used, rather than the exact covariance matrix that would be computed from an infinite number of configurations in our ensembles. We make the leading corrections in $1/N$, where N is the total number of (independent) configurations in our sample [83].

We fit the thinned data to the LO + NLO $S\chi$PT form [6 parameters; Eqs. (18), (20) and (24)], plus generic discretization terms at LO and NLO [6 parameters; Eqs. (26) and (27)], and NNLO analytic terms [16 parameters; Eq. (25)]. The higher-order analytic terms, which include discretization terms, are necessary because our statistical errors in ΔM_{xy}^2 are as low as 0.2%, and always less than 3.3%.[10] In addition, as described at the end of Sec. III B, we must include at least the N³LO term λ_6, Eq. (28), to obtain chiral-discretization fits with acceptable p values. When we include data with charges greater than physical, other N³LO analytical terms are also necessary to obtain acceptable fits.

Our central fit includes data from the $a \approx 0.09$, 0.06, and 0.045 fm ensembles, and quark charges 0, $\pm e/3$, and $\pm 2e/3$. As explained following Eq. (24), we fix to zero the NLO analytic parameter κ_2, which describes sea-quark mass dependence at NLO, and leave the LO parameter Δ_{EM} unconstrained. The generic discretization parameter corresponding to κ_2, called ψ_2, is also fixed to zero. The fit thus has a total of 27 parameters.

Except for the NLO parameter κ_5, all NLO and NNLO parameters, as well as the N³LO term λ_6, are constrained in the central fit by Bayesian priors with a Gaussian width of 3 around 0. As discussed following Eq. (24), the usual χPT expectation would be that these parameters are $\mathcal{O}(1)$; we believe constraining them with a prior width of 3 is reasonable given that it is known that the size of the chiral corrections to ϵ are relatively large. The width for κ_5 is taken to be a factor of 10 larger still, in recognition of the

fact that it gets large unphysical contributions from EM quark-mass renormalization. The width for the generic discretization parameters ψ_i is 0.044, which implies a 1-σ deviation of 5.1% at $a \approx 0.09$ fm, 2.5% at $a \approx 0.06$ fm, and 1.4% at $a \approx 0.045$ fm.

The purpose of the Bayesian priors is to (loosely) enforce χPT behavior, as well as to stabilize the fit to lattice-spacing dependence, for which there are many parameters and several directions in parameter space not well constrained by the data. For the generic lattice spacing dependence, we can write the errors as $(a\Lambda)^2$, where Λ is a discretization scale, $\Lambda \approx 540$ MeV,[11] which we judge is large enough to be conservative. In any case, the effects of increasing the prior widths by factors of 3 or 10 (or in many cases, removing the Bayesian constraints entirely) is included in the systematic errors, as discussed in Sec. VII.

The central fit includes points with meson masses up to about 635 MeV. When masses significantly higher than that are included, it is difficult to fit the data to χPT forms, even with the NNLO analytic terms in the fit function. Some alternative chiral-discretization fits that are used to estimate systematic errors include data up to about 660 MeV, but their p values are rather poor (10^{-4} to 10^{-3}). Other alternatives reduce the maximum meson mass included; the lowest maximum is about 540 MeV. We do not go below this because, in order to be able to interpolate to the physical kaon with controlled errors, it is necessary to include the meson made from one valence quark with mass near m_s and the other the lightest valence quark. We always include the lightest mesons available, which are "pions" with mass of about 250 MeV at $a \approx 0.09$ fm and about 225 MeV at $a \approx 0.06$ and 0.045 fm.

We emphasize that the masses mentioned in the previous paragraph all refer to taste-ξ_5 (Goldstone) mesons, which are the only mesons for which we have a significant amount of data. Mesons with other tastes can appear at one loop in the chiral expansion. The minimum RMS mass of such mesons is about 330 MeV at $a \approx 0.09$ fm, about 260 MeV at $a \approx 0.06$, and about 240 MeV at $a \approx 0.045$ fm. The taste splittings have less effect on the maximum masses; for the central fit the maximum RMS mass goes from about 670 MeV at $a \approx 0.09$ fm to about 650 MeV at $a \approx 0.045$ fm.

Figure 12 shows the same subset of our data as in Fig. 11 (charge $\pm e$, unitary or approximately unitary) after correction for FV effects, along with the central chiral-discretization fit and its extrapolations. The unitary (or approximately unitary) points from the same fit for neutral mesons made out of d- or s-type quarks (charges $\pm e/3$) are shown in Fig. 13. The fit has 264 data points, 237 d.o.f., $\chi^2 = 248.0$, and $p = 0.47$. (Without correction for the use

---

[10]The smallest errors tend to occur either when both valence quarks have masses near the heavier end of our range, or when just one quark is light, and it is uncharged. Typically, relative errors for mesons with net charge are less than those for neutral mesons, because the total EM effect is smaller for the neutrals. Approximately 80% of the data points have errors of 1% or less.

[11]Note that the actual lattice scales used, which we obtain from r_1/a values in Table I and $r_1 = 0.3117(22)$ fm [46], are somewhat smaller than their nominal values.
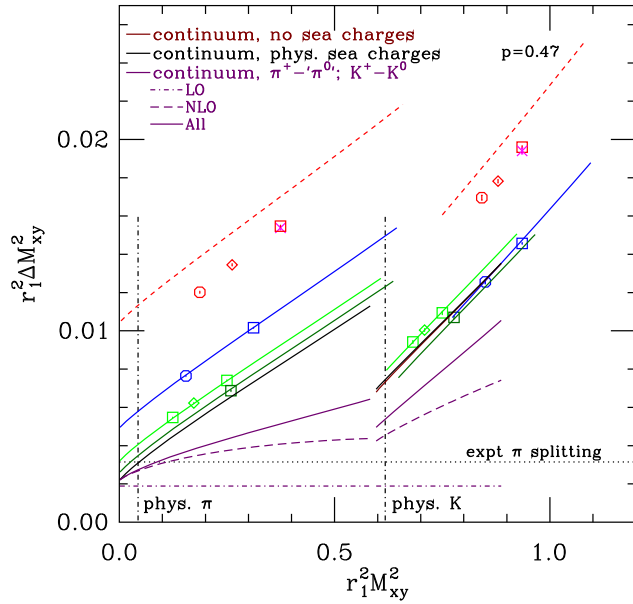
FIG. 12. Central fit to the EM splitting ΔM_{xy}^2 vs the sum of the valence-quark masses, after correction for FV effects. The same small subset of the data as in Fig. 11 is shown. The blue, light green, and dark green lines show the fit to the $a \approx 0.09$, 0.06, and 0.045 fm data, respectively. The largest lattice spacing (red and magenta points, $a \approx 0.12$ fm) is not included in the fit; the dashed red lines show how the fit does in "predicting" these points. The horizontal dotted line shows the experimental value of the $\pi^+ - \pi^0$ splitting; the vertical dashed-dotted lines show the quark mass values for physical π and K mesons. The black and brown curves are extrapolations of ΔM_{xy}^2 to the continuum, with and without the NLO effects of sea quark charges, respectively. (The brown curve is barely visible under the right hand black curve; the curves are identical for the pions at left, and only the black curve is visible.) The solid purple curves are obtained from the black ones by subtracting ΔM_{xy}^2 for the corresponding neutral mesons, K^0 and "π^0." The dashed-dotted line and the dashed purple curves show the LO, and LO + NLO contributions to the total solid purple lines, respectively.

of the sample covariance matrix, the p value of this fit would have been 0.30.) Here and below, when we give χ^2 or p values without further qualification, they are the standard ones, where χ^2 comes only from the difference of the data and the fit, and the d.o.f. are equal to the total number of data points minus the number of parameters, without regard to whether those parameters are constrained by Bayesian priors. We will specify when we actually mean the augmented values, where the priors are treated as additional data, contributing to χ^2 as well as to the d.o.f. Because the priors are loose, in the sense that the parameters are to a great extent constrained by the data and not the priors, we expect that the augmented p values will be larger than the standard p values. In the case of the central fit, $(\chi^2/\text{d.o.f.})_{\text{aug}} = 255.3/263$, giving $p_{\text{aug}} = 0.79$.

In Fig. 12, the blue, light green, and dark green curves show the quality of the fit to the $a \approx 0.09$, 0.06, and
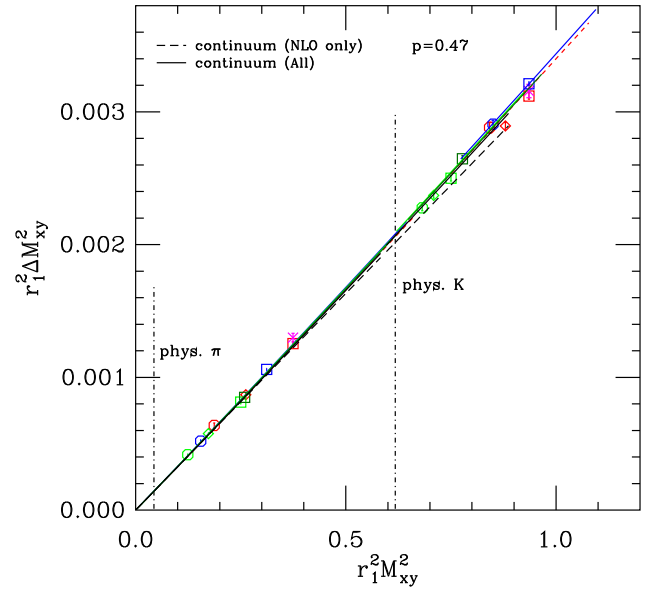

FIG. 13. Same fit as Fig. 12, but showing the neutral mesons made out of quarks with charges $\pm e/3$. The meaning of the blue, light green, dark green, and dashed red curves is the same as in Fig. 12; note the difference in vertical scale between the two plots. The solid black curves are extrapolations of ΔM_{xy}^2 to the continuum. The dashed black lines show the NLO contribution to the solid black curves; there are no LO contributions. The discretization errors and sea-mass dependence for neutral mesons are very small, as are the nonlinear contributions to the valence-mass dependence.

0.045 fm data, respectively. The points at $a \approx 0.12$ fm (red and magenta) are not included in the fit, but the dashed red curves show that the fit does reasonably well in predicting the data at this lattice spacing. It is more difficult to extrapolate to larger lattice spacing than to smaller lattice spacing, since larger lattice spacing may be sensitive to higher-order terms that are either not included in the fit or not well determined on finer lattices.

For the neutral mesons (Fig. 13), the discretization errors, as well as the sea-mass dependence, are quite small, since points from different lattice spacings and sea-mass values line up very well. Further, as required by chiral symmetry, ΔM_{xy}^2 vanishes in the chiral limit. It is also noteworthy that the curvature in the fit lines is small, as may be deduced from the small difference between the curves and the dashed black line, which is straight. There are no chiral logarithms for neutral particles at NLO, and the NNLO logs are not included in our fit function. There is a contribution from the NNLO analytic term that is quadratic in valence masses and can contribute to neutral mesons [the ρ_{11} term in Eq. (25)], but it is rather small: $\rho_{11} = 0.52(3)$. All our alternative chiral-discretization fits preserve these simple features, which are enforced by the lattice data. Therefore we only show the charged-meson plots for the alternative fits below.

The black curves in Fig. 12 show the fit after setting valence and sea masses equal, adjusting m_s to its physical
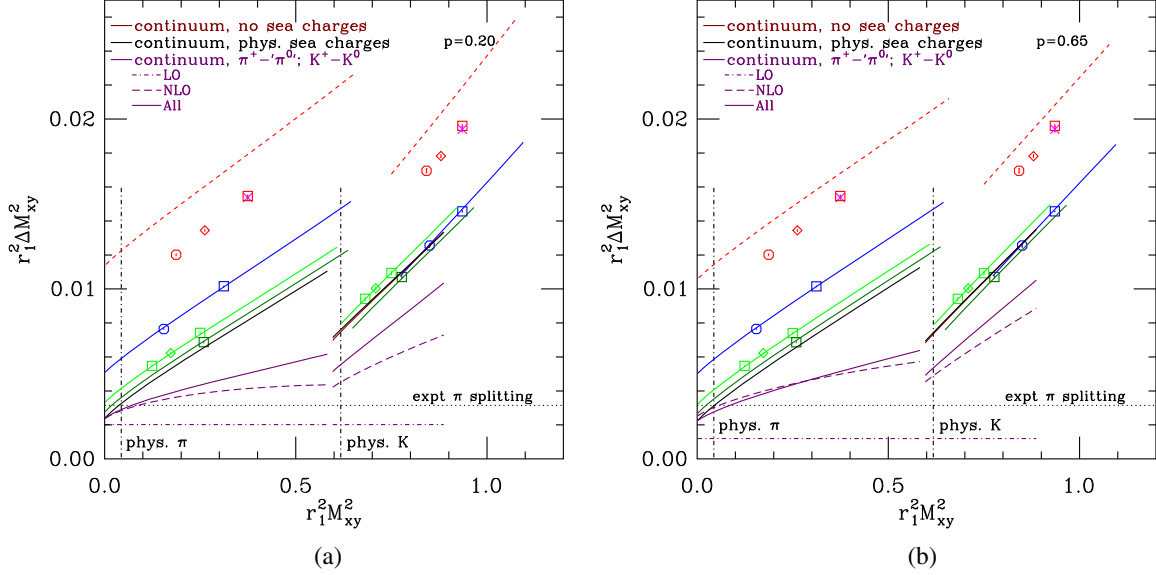
FIG. 14. Two alternative chiral-discretization fits. The data included in the fits, as well as the meaning of symbols and curves, is the same as for the central fit, Fig. 12. Fit (a) sets parameters ρ_6 and ρ_7 to zero, as well as κ_2. The differences with Fig. 12 are small: all curves are slightly higher in the chiral limit, and the predictions for the $a \approx 0.12$ fm points (dashed red curves) are noticeably higher. Fit (b) does not fix the parameter κ_2 (or ρ_6 and ρ_7) to zero. The main difference from Fig. 12 that is apparent in fit (b) is the relative size of the contributions to the continuum result of the LO contribution (horizontal, purple dashed-dotted line), and the LO + NLO contribution (dashed purple curves). The full continuum-extrapolated results (solid purple curves) are however very close to those in Fig. 12.

value, extrapolating to the continuum, and adjusting the sea charges to their physical values using NLO χPT. The last adjustment vanishes identically for pions and is very small for kaons. The brown kaon curve (barely visible under the black kaon curve) shows the value before adjustment, i.e., with vanishing sea-quark charges. From the black curves for the π^+ and K^+, we subtract the corresponding black curves for the neutral mesons "π^0" and K^0 shown in Fig. 13,[12] giving the solid purple curves, whose values at the physical point for each meson (indicated by the vertical dashed-dotted lines) are the physical results.

The solid purple curve in Fig. 12 includes all chiral terms through NNLO (and with the N^{3}LO term λ_6). We also show the LO contribution alone (the mass-independent horizontal dashed-dotted purple line) and the LO + NLO contributions (the dashed purple curves). In this fit, the LO contribution has the value $r_1^2 e^2 \Delta_{\rm EM} = 0.00189(12)$; this is about 60% of the value 0.00315 that would be necessary to give the full experimental pion splitting at LO. As expected from the fact that ϵ, which measures higher-order contributions to $(M^2)^\gamma$, is of order 1, the NLO contributions are relatively large, especially for the kaons or heavier-than-physical pions. The NNLO contributions are clearly much smaller than the NLO ones for physical kaons, and negligible, or nearly so, for physical pions. Thus, after

an anomalously large NLO contribution, SU(3) χPT appears to converge reasonably well.

One may wonder whether this picture of the convergence of χPT is strongly influenced by the Bayesian priors that constrain NLO and NNLO LECs in the fit. In fact, the priors on physical LECs (those whose contributions do not vanish in the continuum limit) have almost no effect on the convergence or the results. If we remove all prior constraint on these physical[13] LECs (κ_2, κ_3, κ_4, κ_5, ρ_6, ρ_7, ρ_8, ρ_9, ρ_{10}, ρ_{11}, ρ_{12}, ρ_{13}, ρ_{14}, ρ_1', ρ_2', λ_5), the fit and results are almost unaffected: ϵ changes by only 0.03%. This is however dependent on our setting parameter $\kappa_2 = 0$ as discussed above; the separation between LO and NLO contributions can be drastically altered if κ_2 is allowed to vary.

Figure 14 shows two alternative chiral-discretization fits to the same set of data points as the central fit. Both of these fits have reasonable p values. In (a), the NNLO parameters ρ_6 and ρ_7 are set to zero, in addition to the NLO parameter κ_2. These NNLO parameters play a role that is similar to κ_2: the corresponding analytic terms depend on the sea-quark masses and are nonzero in the chiral limit. No major changes from the central fit are visible, but the LO continuum contribution (dashed-dotted purple curve) is slightly higher than for the central fit [here,

---

[12]More precisely Fig. 13 only shows the $d\bar{d}$ contribution to the "π^0."

[13]The LO physical LEC $\Delta_{\rm EM}$ is never constrained unless nonzero κ_2 is allowed. We also include κ_5 and λ_6 among the "physical" LECs even though they get unphysical contributions from EM renormalization, because they do not vanish in the continuum limit.

$r_1^2 e^2 \Delta_{\rm EM} = 0.00202(4)$], and all the curves (fit lines as well as extrapolations) are correspondingly higher in the chiral limit. The value of ϵ, however, is only 0.002 below that of the central fit. One somewhat more obvious change is that the predictions for the $a \approx 0.12$ fm points (dashed red curves) are somewhat worse than in the central fit.

In Fig. 14(b), κ_2 is not fixed, but is constrained by our standard prior for physical LECs, 0 ± 3. The LO parameter is now also constrained by priors with central value $r_1^2 e^2 \Delta_{\rm EM} = 0.003$ and width 0.001. The posterior value is almost two sigma lower, $r_1^2 e^2 \Delta_{\rm EM} = 0.0012(3)$, now less than 40% of the value that would be necessary to give the experimental pion splitting at LO. Nevertheless, the final results (sum of all chiral orders) from this fit are quite close to those of the central fit, as can be seen by comparing the solid purple curves in Figs. 12 and 14(b). Indeed, the value of ϵ coming from this alternative fit is just 0.02 below that in the central fit. The fit lines to the data at fixed lattice spacings and sea masses are also very similar in Figs. 12 and 14(b).

In Fig. 15, we show a third alternative fit to the same data points as the central fit. Here, we have put a very tight prior on $\Delta_{\rm EM}$, $r_1^2 e^2 \Delta_{\rm EM} = 0.0031 \pm 0.0001$) to force the LO χPT contribution to be close to the experimental pion splitting. The posterior value is about two sigma below this, $r_1^2 e^2 \Delta_{\rm EM} = 0.0029(1)$. The chiral LECs all have priors 0 ± 3, including κ_2, which is allowed to vary, but now has a negative posterior value in order to reduce the pion chiral limit of the fit to something that is better tolerated by the data. The p value of the fit ($p = 0.035$) is significantly less than the other fits that we have considered so far, but is still acceptable. The resulting value of ϵ is 0.024 higher than that of our central fit, and is in fact the largest positive deviation from the central value of all the alternative fits we have considered.

The relative contributions in the continuum of various orders in the chiral expansion as predicted by the fits are also sensitive to the parameters that control lattice spacing dependence (κ_1, ρ_1, ρ_2, ρ_3, ρ_4, ρ_5, ψ_0, and ψ_i). As mentioned above, the fit becomes unstable if these parameters are completely unconstrained. If the prior widths are widened but not eliminated, the effects on the results are controlled (and included in the systematic error estimates), but the division between LO and NLO contributions can again be significantly changed. Thus the division between orders in χPT shown in Fig. 12 is at best very rough. The final results are nevertheless much more stable than the individual χPT orders, as we have already seen in comparing the fits in Figs. 12, 14, and 15. This remains true even for the more extreme divisions between orders considered below.

It is not surprising that inclusion of the $a \approx 0.12$ fm ensembles leads to difficulties with the fits. The taste-breaking effects at this lattice spacing produce large discretization errors, and the fact that the physical strange
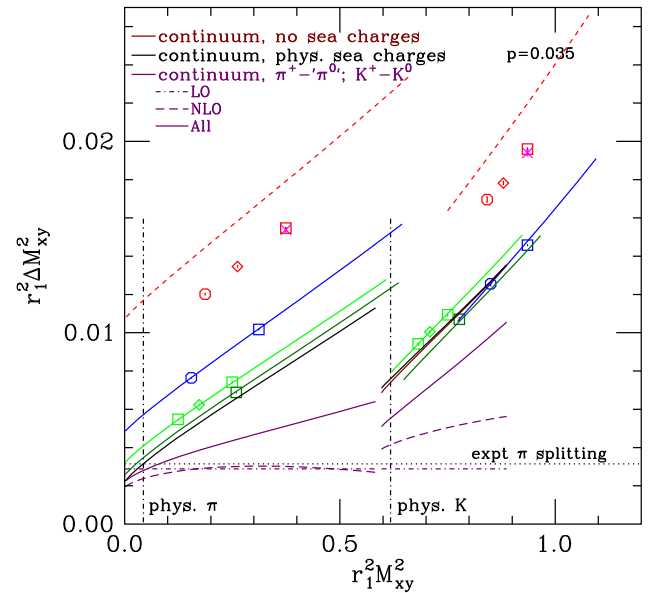

FIG. 15. An alternative chiral-discretization fit to the same data as for the central fit. The meaning of symbols and curves is the same as for the Fig. 12. This fit puts a strict prior on $\Delta_{\rm EM}$ to force it to be very near to the value that would give the physical pion splitting at LO. Note that the higher chiral orders reduce the pion chiral limit significantly below the LO contribution.

quark is about 35% below the simulated strange mass gives further problems. The smallest meson-mass maximum that allows us to interpolate to the kaon is approximately 645 MeV for the Goldstone meson, and about 750 MeV for the RMS taste meson. For low masses, the taste effects are even worse: while the lowest available Goldstone mass is about 275 MeV, this corresponds to an RMS taste mass of about 465 MeV. Thus, when we add in the $a \approx 0.12$ fm ensembles, the chiral-discretization fits have various undesirable features. Figure 16 shows two examples of such fits. Figure 16(a) is rather similar to the central fit: $\Delta_{\rm EM}$ is unconstrained but κ_2 is fixed to zero. Despite the fact that we have increased the prior widths of the LECs controlling lattice spacing dependence (κ_1, ρ_1, ρ_2, ρ_3, ρ_4, ρ_5) to 40, and the width of the generic variation parameters to 0.11, the fit is poor ($p = 0.0005$). Nevertheless, ϵ is only 0.01 below that of the central fit.

Fits with reasonable p values that include the $a \approx 0.12$ fm ensembles are possible. In Fig. 16(b), we allow κ_2 to vary, and put a relatively loose prior on $\Delta_{\rm EM}$ ($r_1^2 e^2 \Delta_{\rm EM} = 0.003 \pm 0.001$), as well as dramatically increasing the prior widths of the parameters that control lattice spacing dependence. We now obtain $p = 0.098$. However, this fit has a negative value of $\Delta_{\rm EM}$, which implies an extreme breakdown of χPT, as well as very large discretization LECs ($\kappa_1 \approx -34$, $\rho_1 \approx -70$, $\rho_2 \approx 33$, $\lambda_2 \approx 16$). It may very well be justified to drop this fit on these grounds. To be conservative, however, we keep it in estimating the systematic errors. Indeed, it is the fit that gives a value of ϵ
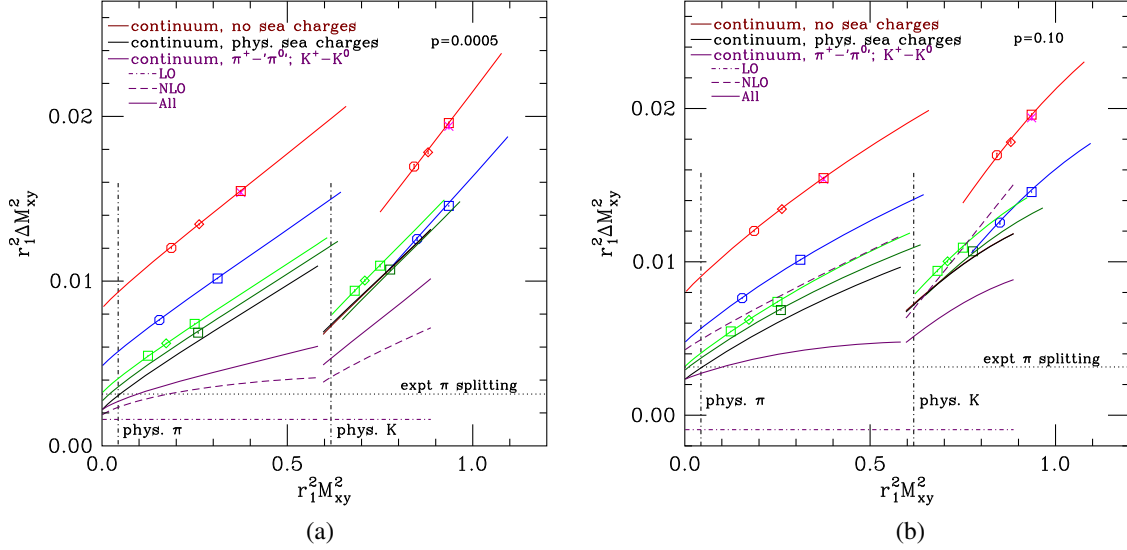
FIG. 16. Two examples of chiral-discretization fit and extrapolation like the central fit in Fig. 12, but including the points from the $a \approx 0.12$ fm ensembles. Fit (a) is most similar to the central fit, in that κ_2 is fixed to zero and Δ_{EM} is unconstrained. Fit (b) allows κ_2 to vary (with a prior 0 ± 40), and imposes a prior of 0.003 ± 0.001 on $r_1^2 e^2 \Delta_{EM}$. Nevertheless the fit value of Δ_{EM} is negative.

that is furthest away from our central value (0.082 lower) out of all the chiral-discretization alternatives we consider.

Adding in points with quark charges that are greater than the physical ones leads to problems with the chiral-discretization fits that are similar to those we have when adding in the $a \approx 0.12$ fm ensembles. This may be because the higher charges bring in greater discretization errors. Indeed, there is evidence [12] that taste violations from photon exchange begin to be important when the charges increase above their physical values. For data with quark charges $0, \pm 1/3, \pm 2/3, \pm 1$, and $\pm 4/3$, a fit like the central fit [but including all the $e^4 p^2$ LECs λ_i in Eq. (28)] and with somewhat larger priors (0 ± 5 on LECs) has $p = 0.005$ and an ϵ that is 0.03 below that of the central fit. A fit with κ_2 not fixed to zero, and very loose priors on LECs and generic discretization parameters, has $p = 0.21$. However Δ_{EM} is negative, and discretization terms are very large. Both features are quite similar to those seen in Fig. 16(b). In this case, ϵ is 0.065 below that of the central fit.

VII. SYSTEMATIC ERRORS AND RESULTS OF EM CALCULATION

Our calculation has the following significant sources of systematic errors: (1) chiral-continuum uncertainties from the extrapolations to the physical light quark mass and to $a = 0$, (2) finite-volume (FV) effects, (3) systematic issues involved in the EM renormalization, (4) effects of using the "π^0," which does not include quark-line disconnected contributions, instead of the true π^0, (5) errors in the physical value of r_1, the quantity we use to set the scale, (6) uncertainties in the physical values of the quark masses after extrapolation to the continuum, and (7) effects of EM

quenching. In the following subsections, we discuss each source of error in turn. For the two EM quantities we calculate, ϵ and ϵ_{K^0}, Table VI lists central values and statistical errors from the fit shown in Fig. 12, and each systematic error.

The separation between EM and isospin-violating effects is dependent on the scheme, which enters through the EM renormalization of quark masses. Our calculation is performed in the BMW scheme [71], described in Sec. III C. For some purposes, it may be useful to gauge the effects of changing to another reasonable scheme. In Sec. VII H, we estimate such scheme dependence for ϵ and ϵ_{K^0}.

A. Chiral-continuum error

We determine this error by considering a wide range of alternative chiral-discretization fits, with various priors and/or parameters set to zero, and apply them to various subsets of the data: different maximum meson masses included, different thinning, omitting or including the coarsest

TABLE VI. Central values and errors for ϵ and ϵ_{K^0}.

Source	ϵ	ϵ_{K^0}
Central value	0.776	0.035
Statistics	0.012	0.003
Chiral-continuum	$^{+0.024}_{-0.082}$	0.002
Finite volume	0.056	⋯
Renormalization	0.002	0.012
"π^0"	0.034	⋯
Absolute scale	0.001	0.000
Quark masses	0.009	0.011
EM quenching	0.040	0.012

($a \approx 0.12$ fm) ensembles, and omitting or including quark charges greater than the physical charges. Several of these fits have been presented in Sec. VI C. To be conservative we include fits that have p values as small as 10^{-5}, as well as ones that have very large discretization terms and/or exhibit extreme breakdown of χPT (e.g., negative Δ_{EM}). Altogether, 89 fits are included. We take the largest positive and negative differences from the value in the central fit as the error. For ϵ, this gives a positive error of $+0.024$, coming from the fit in Fig. 15, and a negative error of -0.082, coming from the fit in Fig. 16(b). For ϵ_{K^0}, the maximum positive and negative differences are comparable, so we average them and quote a symmetric error of 0.002.

B. Finite-volume error

To estimate the systematic error associated with the FV correction we use (a *residual* FV error), we examine the deviations of the fit lines from the data in Fig. 9. By far the largest deviation occurs for the "pion" (blue) curve at $L = 16$. Let x be the difference between the predicted infinite-volume value of the pion mass from the $L = 16$ point alone and the value from all the other points. We take x as the presumed absolute value of the error of our FV estimates, and divide it by the size of the estimated FV correction at $L = 20$, to find a fractional residual FV error of approximately 20%. We use the $L = 20$ point because most of the data used in the central fit is from physical volumes of approximately that size or slightly larger; using the $L = 28$ point instead would make a negligible difference. An error of 20% is also reasonable because usual corrections from higher orders in SU(3) χPT are of this size. For ϵ the net effect of the FV corrections we have made is 0.28, which we find simply by comparing our central value with the value obtained by refitting the data with FV corrections turned off. Our estimate for the residual FV error is then 20% of 0.28, or 0.056.

For neutral mesons, there are no chiral logarithms at NLO, and hence no FV effects at this order. There will be FV effects at higher orders, but they are very likely to be much smaller than our other systematic errors in ϵ_{K^0}, which are quite large. We therefore do not include a residual FV error for ϵ_{K^0} in Table VI.

C. EM renormalization error

We use the BMW scheme, as defined by Eq. (37) and as implemented by Eq. (40), to perform nonperturbative EM renormalization of the u- and d-quark masses. For ϵ, this is sufficient, since the renormalization of the s-quark mass cancels in the difference $(M_{K^+}^2 - M_{K^0}^2)^{\gamma}$. However s-quark mass renormalization is crucial for obtaining ϵ_{K^0}. We extend the renormalization to the s quark using Eq. (42).

We can implement Eq. (40) to high accuracy from our chiral fits, so the only significant systematic errors in the scheme come from the errors in our values of the derivatives of the squared meson masses with respect to quark mass: $B \equiv \partial M_{\pi}^2 / \partial m_l$, $B_l = \partial M_K^2 / \partial m_l$, and $B_s = \partial M_K^2 / \partial m_s$. We only need these quantities for physical quark masses and in the continuum limit, since we perform the renormalization after the chiral-discretization fit and its extrapolations. For B, we have the SU(2) χPT result, Eq. (41), which is quite precise: the error from $\bar{\ell}_3 = 2.81(64)$ [2] results in a 0.4% error in B. Corrections from NNLO should be even smaller, since the NLO correction is already only 2%. For B_l and B_s, SU(3) χPT would be needed, but the higher-order corrections, as well as the uncertainty in the relevant LEC, are large. Instead, we extract these quantities from our lattice data, and make a simple extrapolation (linear in a^2) to the continuum. We find $B_s/B = 0.974(15)$, $B_l/B = 0.946(19)$, where we give the results in terms of the central value of B (errors in B should not be added on to these values). The error in B_l is small enough that the resulting error in ϵ is small compared to other systematic errors; ϵ is independent of B_s. The total renormalization error on ϵ is 0.002.

For ϵ_{K^0}, we find a renormalization error of 0.012. The error is dominated by the uncertainty coming from B_s, and would therefore benefit from increased precision in this quantity. A significant improvement in B_s could be obtained from a dedicated pure QCD calculation with several closely spaced strange-quark masses around the physical value at each lattice spacing. However, the fact that ϵ_{K^0} has an uncontrolled quenched-EM error means that one cannot decrease the overall error very much without also going to dynamical QED simulations (or equivalent approaches to include the effects of sea-quark charges at order α_{EM}).

D. Error from dropping disconnected π^0 diagrams

As described in Sec. III D, we simulate a "π^0" in which quark-line disconnected diagrams are dropped, rather than the physical π^0. The difference is $\mathcal{O}(\alpha_{\text{EM}} M_{\pi}^2)$.

We may estimate the size of this effect by noting that the disconnected contributions are solely responsible for the chiral logarithm term found in Ref. [22]. The connected contributions, which are equal to $(M_{uu'}^2)^{\gamma}$ or $(M_{dd'}^2)^{\gamma}$, have no chiral logarithms at NLO. Indeed, there are no NLO chiral logarithms for any neutral meson that has only connected contributions, such as the neutral kaon. In Sec. III D, we estimated the chiral logarithm term in $(M_{\pi^0}^2)^{\gamma}$ as approximately 30(MeV)^2. Using the result from our central fit for $e^2 \Delta_{\text{EM}}$ instead of the value from Ref. [21], gives a smaller result of 25(MeV)^2.

In estimating the error on ϵ, we also need the pion splitting, which appears in the denominator. The experimental value is 1261 (MeV)^2, but we can put the denominator on the same footing as the error in the numerator by using instead the value obtained from LO, namely $e^2 \Delta_{\text{EM}}$.

With the value of Δ_{EM} from Ref. [21], we get about 900 $(MeV)^2$ for the LO pion splitting. Taking this smaller value for the denominator and the larger estimate of the error in numerator, we get a conservative estimate of 0.034 for the error in ϵ. This value is independent of what is assumed for Δ_{EM}, since it cancels between the numerator and denominator.

Another approach to estimating the "π^0" error would be to compare ϵ with ϵ', defined by Eq. (3). In ϵ', the experimental value of the pion splitting appears in the numerator instead of the computed value of the pion EM splitting, so ϵ' is independent of how we treat the π^0. From the discussion in Sec. I, we expect that, in the absence of statistical or other systematic errors, $\epsilon = \epsilon' + \epsilon_m$, where $\epsilon_m = 0.04(2)$ [2]. However, chiral-discretization errors play a significant role here, since there is a partial cancelation of errors in ϵ when we subtract $(M_{\pi^+}^2 - M_{"\pi^0"}^2)^\gamma$ from $(M_{K^+}^2 - M_{K^0}^2)^\gamma$. Indeed, the chiral-discretization error for ϵ' is a factor of about 4 larger than for ϵ. If we ignore this problem and just focus on the central fit, $\epsilon - (\epsilon' + 0.04) = 0.089$. This is slightly smaller than the expected error (0.091) from the addition in quadrature of the chiral-discretization and "π^0" errors on ϵ, and the ϵ_m error. Because there are also likely to be some residual FV errors in the difference $\epsilon - (\epsilon' + 0.04)$, the ϵ' result suggests that the errors we have already included are reasonable and do not need to be increased further.

The calculation of the EM effect for the K^0 is independent of the treatment of the π^0, so there is no corresponding error in ϵ_{K^0}.

E. Scale error

The absolute scale of our ensembles is set by $r_1 = 0.3117(22)$ fm [46]. To find the induced error in our results, we rerun the analysis with r_1 changed by 1 σ. In doing so, it is necessary to include the changes, caused by the scale, in the physical quark masses in the continuum limit. The scale error in these masses is given in Table V. Note that the estimates of the quark masses move in the same direction as r_1 because the quark masses are adjusted to reproduce the experimental values of the meson masses multiplied by r_1.

The resulting scale errors are very small: 0.001 in ϵ and 0.0002 in ϵ_{K^0}. In each case, the effect of changing the scale itself is largely cancelled by the scale changes in the quark masses. Only the denominators, which come from the squared experimental splitting multiplied by r_1^2, are affected by the change in the scale itself, and only the numerators are affected by the changes in quark masses.

F. Quark mass error

To find the errors coming from our values of the physical quark masses, we rerun the analysis with the continuum physical mass values[14] given in Table V changed by 1 systematic σ (not including scale errors). Because the nonscale errors arise largely from variations over the same set of pure QCD chiral fits, m_l and m_s are highly, positively correlated, and we change both in the same direction. We find quark mass errors of 0.009 in ϵ and 0.011 in ϵ_{K^0}. Assuming instead that the errors on m_l and m_s were uncorrelated would change the resulting errors only slightly because changes in one of the masses always dominate: m_s dominates for ϵ, while m_l dominates for ϵ_{K^0}.

Errors arising from the other inputs in Table V are negligible and are not included in Table VI. Because only the products $B_0 m_l$ and $B_0 m_s$ enter our results, it is clear that the errors in B_0 will have a negligible effect compared to the effect of the quark mass errors. We bound the effects of the errors in the splittings by rerunning the analysis with all splittings at a lattice spacing changed by 1σ in the same direction, but with the direction varied randomly at different lattice spacings. Because splitting errors are uncorrelated for different ensembles, and only somewhat correlated for different splittings on the same ensemble, changes of this type provide an upper limit on the changes we find if we change individual spacings randomly within their errors. The maximum differences we find are 0.007 in ϵ and 0.0004 in ϵ_{K^0}, which are in each case smaller than statistical errors, and much smaller than the dominant systematic errors.

G. Quenched EM error

For ϵ, the effect of having quenched the EM interactions is controlled at NLO in SU(3) χPT, per the argument of Ref. [21]. This is because effects that depend on the sea-quark charges and unknown LECS are independent of valence-quark charges and therefore cancel in $(M_{K^+}^2 - M_{K^0}^2)^\gamma$ and in $(M_{\pi^+}^2 - M_{"\pi^0"}^2)^\gamma$—see Eq. (23). Errors arise at NNLO, in which cross terms between valence and sea charges can first appear in analytic terms, which have unknown LECs. Examples of such terms are ones proportional to $q_{xy}(\mu_x - \mu_y)(q_u \mu_u + q_d \mu_d + q_s \mu_s)$ or $(q_x + q_y)(\mu_x + \mu_y)(q_u \mu_u + q_d \mu_d + q_s \mu_s)$. From our central fit, the calculated effect of turning on the sea quark charges is 0.040, or 8.2% of the 0.486 NLO contribution for neutral sea quarks. Assuming the quenching effects on NNLO would be of this same size, an estimate of the electroquenching error is 8.2% of the 0.250 NNLO contribution, or 0.020. It may be, however, that the electroquenching effect at NLO is anomalously small. In particular, there is no effect on "pions" (mesons with degenerate quarks) at this order. We therefore follow a more conservative

---

[14]The values of physical quark masses at nonzero lattice spacings affect our results only through the values of r_1/a, which are extrapolated to these quark masses in our mass-independent scale-setting scheme. The effects on the final results are negligible.

approach and take the full value of the NLO sea-quark charge effect, namely 0.040, as the error estimate for ϵ.

As explained at the end of Sec. III A, the electroquenching error in ϵ_{K^0} is uncontrolled, in the sense that it is not computable at lowest nontrivial χPT order. We can get a rough handle on this error by $1/N_c$ counting. At $\mathcal{O}(\alpha_{EM})$, the electroquenching effects come from diagrams with either (1) a photon that connects a sea-quark loop to a valence line, (2) a photon that spans a single sea-quark loop, or (3) a photon that connects two sea-quark loops. In all three cases, each loop must also have attached gluons,[15] so $1/N_c$ counting applies. Diagrams (1) and (2) are then suppressed by $1/N_c$, while diagram (3) is suppressed by $1/N_c^2$. Diagram (1) is further suppressed by SU(3) flavor [71], since the sum of sea-quark charges vanishes, and quark mass factors must be included to get a nonzero result. Since diagram (2) cancels for ϵ, the double suppression of electroquenching effects may explain why the contribution of sea-quark charges is only 0.04 at NLO. However, diagram (2) does not cancel for ϵ_{K^0}, so we have only the $1/N_c$ suppression. We therefore take 1/3 of the central value, namely 0.012, as the electroquenching error in ϵ_{K^0}.

H. Scheme dependence

It may be helpful for some purposes to estimate the changes that would be induced in our results if we changed to a different scheme for EM renormalization. For example, in a pure-QCD calculation that relies on our results to remove EM effects from physical quantities that are used to set the quark masses or scale, it would be useful to know how much the results could change in a different scheme for separating EM from isospin-violating effects.

In addition to the BMW scheme, we have tried renormalizing the quark masses using the $\overline{\text{MS}}$ scheme at scale $\mu = 2$ GeV. Unfortunately, we have only a one-loop determination of the renormalization, and this may suffer from rather large perturbative errors, as we remarked in Sec. III C. Nevertheless, comparison of the $\overline{\text{MS}}$ scheme at one loop, Eq. (33), with that of the BMW scheme, Eq. (42), gives at least a rough estimate of how much the results may change over various choices of "reasonable" schemes.

With the $\overline{\text{MS}}$ scheme and the central fit, we obtain $\epsilon = 0.814(12)$, where the error is statistical only. This suggests a scheme dependence of $\sim$0.04 in ϵ. The small dependence is in accordance with the general discussion at the beginning of Sec. III C. The corresponding result for the neutral kaon is $\epsilon_{K^0} = 0.365(2)$ giving a scheme dependence of $\sim$0.330. Note that, if the two-loop corrections from QCD make a comparable contribution to the one-loop EM renormalization as they do in the pure QCD, asqtad

---

[15]This follows from Furry's theorem, which forbids loops with only one photon vertex, as well as the usual cancellation of vacuum bubbles that are completely unconnected to the rest of the diagram.

case [70], the value of ϵ_{K^0} in the $\overline{\text{MS}}$ scheme at 2 GeV would be reduced by a factor of order 3.

The large dependence on scheme for ϵ_{K^0} is not surprising, since ϵ_{K^0} is very sensitive to the renormalization of the strange quark mass. The fractional shift in the continuum of the strange mass under EM renormalization in the BMW scheme is 0.32%, while in one-loop $\overline{\text{MS}}$, it is only 0.12%. Neither of these shifts is of an unreasonable size for an $\mathcal{O}(\alpha_{EM})$ effect. The fractional difference of 0.2% in the strange mass would correspond to a change in $(M_{K^0}^2)^\gamma$ of roughly $0.002M_K^2/(1 + 1/27)$, giving a change of ϵ_{K^0} of approximately 0.37, where we take $M_K \approx 495$ MeV. The factor of $1/(1 + 27)$ comes from the fact that the light quark mass, which is not changing, is about $1/27$ of the strange quark mass.

Adding the systematic errors in Table VI in quadrature, we find

$$\epsilon = 0.78(1)_{\text{stat}} \binom{+8}{-11}_{\text{syst}}. \tag{74}$$

$$\epsilon_{K^0} = 0.035(3)_{\text{stat}}(20)_{\text{syst}}. \tag{75}$$

The result for ϵ_{K^0} implies $(M_{K^0}^2)^\gamma = 44(3)_{\text{stat}}(25)_{\text{syst}}(\text{MeV})^2$. A preliminary value for $(M_{K^0}^2)^\gamma$ was reported in Ref. [13]. That result did not yet take into account EM quark-mass renormalization and is thus not reliable.

VIII. CALCULATION OF m_u/m_d

Using the values of ϵ and ϵ_{K^0} given in Eqs. (74) and (75), we can use the dependence of the kaon mass on the light quark mass to find the quark mass ratio m_u/m_d. Because ϵ and ϵ_{K^0} are physical parameters (albeit in a fixed scheme for separating EM from strong isospin-violating effects), we need not use the same set of simulations for this step. Here we use the MILC HISQ $(2 + 1 + 1)$-flavor QCD ensembles, since these have smaller lattice artifacts than the asqtad ensembles and contain ensembles with light quark masses near their physical values. Table VII shows the $2 + 1 + 1$ flavor ensembles with approximately physical light sea quark masses, which are used in this section.

The procedure for finding m_u/m_d is described in detail in Ref. [26]. Very briefly, the essential steps are

(1) Use the pion mass and decay constant to fix the lattice spacing and average light quark mass, $m_l = (m_u + m_d)/2$, on each ensemble. Here we use the physical π^0 mass, since this has small electromagnetic contributions. This mass is also adjusted for QCD finite size effects.

(2) Find the tuned strange quark mass on each ensemble by matching $2M_K^2 - M_\pi^2$. In this step, the lattice masses use the average light quark mass computed in the first step, and the input M_K is the average of

TABLE VII.   Ensembles used in the calculation of m_u/m_d. The first column in this table is the approximate lattice spacing in fm. The second column is the gauge coupling $\beta = 10/g^2$, and the next three columns are the sea-quark masses in lattice units. The primes on the masses indicate that they are the values used in the runs, and in general differ slightly from the physical values because of tuning errors.

Key	β	am'_l	am'_s	am'_c	$(L/a)^3 \times (T/a)$	N_{lats}	a (fm)	L (fm)	$M_\pi L$	M_π (MeV)
0.15	5.80	0.002 35	0.0647	0.831	$32^3 \times 48$	1000	0.150 79(17)	4.83	3.2	130
0.12	6.00	0.001 84	0.0507	0.628	$48^3 \times 64$	999	0.121 11(10)	5.82	3.9	133
0.09	6.30	0.0012	0.0363	0.432	$64^3 \times 96$	1031	0.087 72(12)	5.62	3.7	130
0.06	6.72	0.0008	0.022	0.260	$96^3 \times 192$	895	0.056 73(5)	5.44	3.7	135
0.04	7.00	0.000 569	0.01555	0.1827	$144^3 \times 288$	470	0.042 54(5)	6.12	4.17	134

the K^0 and K^+ masses after subtracting the electromagnetic contributions parametrized by ϵ and ϵ_{K^0}.

(3) Use the derivative of the lattice M_K^2 with respect to the light valence quark mass and the difference between the K^0 and K^+ masses after removing electromagnetic contributions to find $m_d - m_u$ and hence m_u/m_d on each ensemble.

(4) Fit the values of m_u/m_d on each ensemble to a smooth function of the lattice spacing, and evaluate the fit at $a = 0$. For our central fit we fit the points with $a \leq 0.12$ fm to a quadratic in $\alpha_s a^2$, using the strong coupling constant α_s determined from the plaquette. Alternative fits to estimate systematic errors from the continuum extrapolation include a quadratic fit including the 0.15 fm data, a linear fit excluding the 0.15 fm ensemble, and a linear fit excluding both the 0.12 and 0.15 fm ensembles.

The most important differences between this analysis and that of Ref. [26] are the extension of the 0.06 fm ensemble to 895 lattices and the addition of an ensemble with $a \approx 0.04$ fm. Figure 17 shows the values of m_u/m_d for each ensemble, and the continuum extrapolation. With the addition of this data at small lattice spacings, we now choose to omit the 0.15 fm data from our central fit, and use the fit including this ensemble as one of our alternative fits for estimating the systematic error due to the choice of continuum extrapolation. We take the range of all of these continuum extrapolations as our estimate of the systematic error coming from the choices made in our continuum extrapolation.

Using the MILC HISQ $(2 + 1 + 1)$-flavor QCD ensembles and the values of ϵ and ϵ_{K^0} given in Eqs. (74) and (75), and following the approach described in Ref. [26], we obtain

$$m_u/m_d = 0.4529(48)_{\text{stat}} \binom{+118}{-0}_{\text{cont}}$$
$$\times \binom{+91}{-66}_\epsilon (0)_{\epsilon_{K^0}} (4)_{\text{FV}_{\text{QCD}}} (13)_{\Delta M_K(\text{exp})}. \quad (76)$$

The errors on the quantity are, in order, the statistical error and the errors from choices in the continuum extrapolation, from ϵ, from ϵ_{K^0}, from finite volume in the pure QCD

calculation, and from the error in the experimental value of $M_{K^0} - M_{K^+}$ [84]. The finite-volume effects are taken to be the difference between a NLO staggered chiral perturbation theory calculation and a nonstaggered calculation at NNLO for M_π and F_π and NLO for M_K and F_K. We note that the result in Eq. (76) should be considered an update to the result quoted in Ref. [85], $m_u/m_d = 0.4556(55)_{\text{stat}} \binom{+114}{-67}_{\text{syst}} (13)_{\Delta M_K}$. The current result includes newly generated $2 + 1 + 1$ HISQ configurations at $a \approx 0.06$ fm and 0.04 fm, as well as all our configurations at $a \approx 0.09$ fm. Reference [85], which focused on physics for quarks heavier than m_c, included only the subset of configurations at $a \approx 0.09$ fm for which we have generated propagators for those heavy quarks. The smaller statistical error of the current result reflects the larger data set used. Our procedures for estimating systematic errors, however,
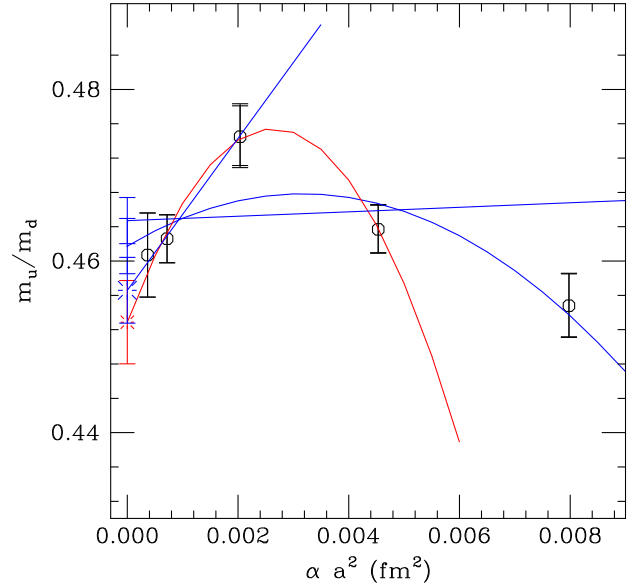

FIG. 17. m_u/m_d on the physical quark mass HISQ ensembles, and the continuum extrapolation. The red line is the fit used for our central value, and the blue lines three of the alternative fits used for estimating systematic error from the continuum extrapolation. These alternate fits are a quadratic fit including all the data points, a linear fit omitting the 0.15 fm. data, and a linear fit omitting both the 0.15 and 0.12 fm data.

actually give slightly larger values in the current analsyis than in Ref. [85].

To this level of precision, and within the scheme we are using, our EM errors in m_u/m_d come only from ϵ and not from ϵ_{K^0}, despite the large relative error in the latter quantity. The errors in ϵ_{K^0} do, however, have an effect on the errors in m_s and in ratios such as m_s/m_l [85,86].

IX. CONCLUSIONS AND OUTLOOK

Using the three-flavor MILC asqtad configurations, we have computed the EM quantities ϵ and ϵ_{K^0}, which parametrize the EM contribution to the K^+–K^0 mass splitting, and to the K^0 mass itself, respectively. Our results are given in Eqs. (74) and (75). A comparison of our result for ϵ (labeled as MILC 18) with those of other groups (and our preliminary result, labeled as MILC 16) is shown in Fig. 18. We note that different groups in general use different schemes for separating electromagnetic and strong isospin-violating effects. Nevertheless, the scheme-dependence of ϵ is likely to be small—see the discussion in Secs. III C and VII H. With the exceptions of the early result in RBC 07, which quotes statistical errors only, and the result from QCDSF 15, from which we differ by about 2 sigma, the agreement with the work of other groups is good.

With the EM contributions in hand, we have proceeded to compute the quark mass ratio m_u/m_d in QCD, using the four-flavor MILC HISQ configurations. Figure 19 compares our work with that of other lattice groups. In general, we only show results that employ a lattice evaluation of the EM effects; however we have included for comparison the MILC 09 [4] result (shown with an open symbol), which relies on a phenomenological estimate of ϵ. With our new results of the EM effects, our estimate for the EM uncertainty in m_u/m_d has been reduced by more than a factor of 5 from our error in MILC 09. Other systematic errors are comparable between MILC 09 and MILC 18, so the total error is reduced by a factor of about 3.5.

Note that our current value for m_u/m_d is plotted in Fig. 19 with the u, d, s, c sea results. The pure QCD HISQ ensembles that are used in finding m_u/m_d indeed have $2+1+1$ dynamical flavors. On the other hand, our EM calculation giving ϵ and ϵ_{K^0} employs the asqtad $2+1$ ensembles. The error from omitting the dynamical charm quark, however, is expected to be at most a few percent. An error of that size would be small compared to the other errors in the EM calculation, so should not effect the final value for m_u/m_d significantly.

Our result for m_u/m_d is consistent with those from most other groups, but lies on the low side of the range of results. From Fig. 17 one can see that the low continuum value from our data set is due to the results from the two finest lattice spacings, $a \approx 0.06$ fm and $a \approx 0.04$ fm. The latter is finer than the finest of the ensembles used by the other groups, which has $a \approx 0.054$ fm. Because discretization errors depend on the lattice action, however, it is unclear at
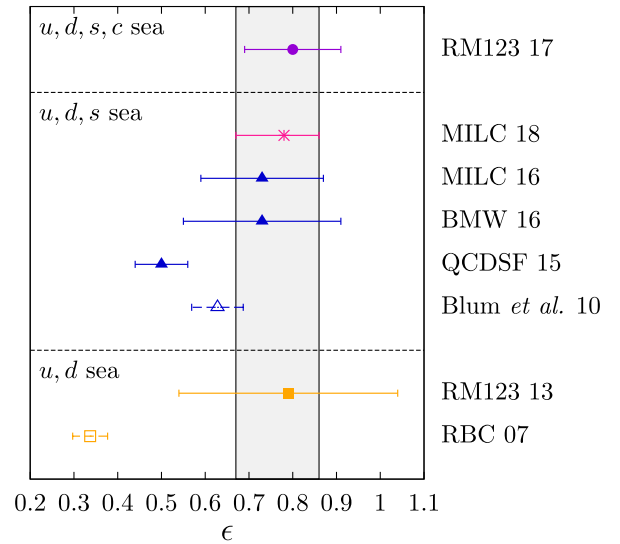

FIG. 18. Comparison of ϵ in Eq. (74) (magenta burst) with previous unquenched lattice-QCD calculations. The open symbols with dashed error bars represent early work, with only statistical errors quoted. The references are RM123 17 [19], MILC 16 [15] (a preliminary result), BMW 16 [17], QCDSF 15 [16], Blum *et al.* 10 [10], RM123 13 [18], and RBC 07 [9].
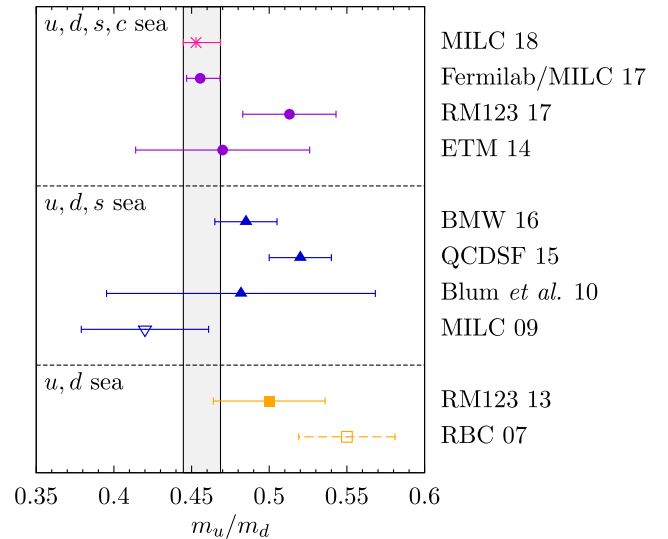

FIG. 19. Comparison of m_u/m_d in Eq. (76) (magenta burst) with previous unquenched lattice-QCD calculations that include a lattice evaluation of the EM effects. For comparison, we also show, with an open triangle, the MILC 09 result, which uses a phenomenological estimate of the EM effects. An early result, RBC 07, just quotes statistical errors and is shown with dashed error bars and an open symbol. The current (MILC 18) result should be considered an update of the Fermilab/MILC 17 result [see discussion following Eq. (76)]. The references are Fermilab/MILC 17 [85], RM123 17 [19], ETM 14 [87], BMW 16 [17], QCDSF 15 [16], Blum *et al.* 10 [10], MILC 09 [4], RM123 13 [18], and RBC 07 [9].

this point whether the difference in available lattice spacings is relevant to the apparent differences seen in Fig. 19.

While the electroquenching errors for ϵ are under control, these errors are uncontrolled for most quantities, e.g., ϵ_{K^0}. To move beyond the electroquenched approximation, we have developed a dynamical EM code [88] and are beginning to generate unquenched QCD + QED ensembles. These ensembles will be crucial to our efforts to obtain precise results for the hadronic contributions to $(g - 2)_\mu$, as well as for calculations such as the proton-neutron mass difference and improvements in the result for ϵ_{K^0}.

ACKNOWLEDGMENTS

APPENDIX: OBTAINING THE MASS FROM THE SELF-ENERGY AT FINITE T

For infinite T, the standard procedure to get the correction to the squared mass is to evaluate the Euclidean self energy $\sigma(p_0)$ at $p_0 = im$. In this Appendix, we show that method does not in general give the right answer at finite T. In particular, $\sigma(im)$ is dependent on the routing of the loop

momentum through the diagram. Nevertheless, we show that the particular momentum routing chosen in Sec. V does allow us to extract the mass correction from $\sigma(im)$ because the natural continuation of $\sigma(p_0)$ away from the Matsubara frequencies $2\pi n/T$ happens to be particularly simple.

To introduce our notation and approach, we first review the usual procedure when the time extent T is infinite. The momentum-space Euclidean propagator has the form

$$\tilde{G}_\infty(p_0) = \frac{1}{p_0^2 + m^2 + \sigma_\infty(p_0)}, \qquad (A1)$$

where m is the Lagrangian mass, σ_∞ is the self energy, the subscript ∞ indicates that T is infinite, and we have taken the case of vanishing spatial momentum, $p = (p_0, \vec{0})$, for simplicity.

To find the physical mass, we Fourier transform to position space

$$G_\infty(t) = \int \frac{dp_0}{2\pi} e^{ip_0 t} \tilde{G}_\infty(p_0) \qquad (A2)$$

$$= C e^{-Mt} + \cdots \qquad [t > 0], \qquad (A3)$$

where C is a constant, and $p_0^2 = -M^2$ is the location of the single-particle pole

$$M^2 = m^2 + \sigma_\infty(iM) \approx m^2 + \sigma_\infty(im), \qquad (A4)$$

and $\cdots$ in Eq. (A3) represents the contributions of excited and multiparticle states. Equation (A3) follows from Eq. (A2) by completing the contour in the upper half plane using Jordan's lemma, which requires only that

$$\lim_{|p_0| \to \infty} \frac{1}{p_0^2 + m^2 + \sigma_\infty(p_0)} = 0 \qquad (A5)$$

in the upper half plane. From Eq. (A4), we read off the standard answer: the first-order correction to the squared mass is simply the self-energy evaluated at $p_0 = im$.

When T is finite, the calculation of the mass correction changes in two crucial ways. First of all, the integral over p_0 in Eq. (A2) becomes a sum over $p_0 = 2\pi\ell/T$, where ℓ runs over the integers. The self energy $\sigma(p_0)$ and hence $\tilde{G}(p_0)$ are moreover only well defined for these discrete values of p_0. We may continue these functions to $\sigma_{\text{cont}}(p_0)$ and $\tilde{G}_{\text{cont}}(p_0)$, defined on the full complex p_0 plane, but the continued functions are not unique. Second, the internal loop energy [e.g., k_0 in Eq. (63)] in the determination of $\sigma(p_0)$ is itself discrete, so that $\sigma(p_0)$ is not the same function of p_0 as $\sigma_\infty(p_0)$, even on the discrete points $p_0 = 2\pi\ell/T$. These two changes interact in interesting ways, with the result that the procedure to obtain the squared-mass correction by evaluating $\sigma_{\text{cont}}(p_0)$ at $p_0 = im$ is not valid in general.

We discuss the discrete sum over p_0 first. To extract the mass, we need to compute

$$G(t) = \frac{1}{T} \sum_{p_0 = 2\pi\ell/T} e^{ip_0 t} \tilde{G}(p_0). \qquad (A6)$$

$G(t)$ is a periodic function of t with period T. The standard technique is to use the Poisson summation formula to rewrite $G(t)$ as a sum of nonperiodic propagators to each periodic image of the fundamental domain $0 \leq t \leq T$. Usually these nonperiodic propagators are just the known $T = \infty$ propagators, but here that is not the case, since we are keeping T finite for the internal energy sums in σ. We instead simply use a continuation $\tilde{G}_{\text{cont}}(p_0)$ of $\tilde{G}(p_0)$, which defines $G_{\text{cont}}(t)$ by Fourier transformation. The Poisson formula then gives

$$G(t) = \sum_{n=-\infty}^{+\infty} G_{\text{cont}}(t + nT), \qquad (A7)$$

$$G_{\text{cont}}(\tau) = \int \frac{dp_0}{2\pi} e^{ip_0 t} \tilde{G}_{\text{cont}}(p_0). \qquad (A8)$$

Although $\tilde{G}_{\text{cont}}(p_0)$ is not unique, it is straightforward to check that another continuation constructed by adding a function that vanishes at $p_0 = 2\pi\ell/T$, such as $\sin(p_0 T/2)$, will not change $G(t)$, although it does of course change $G_{\text{cont}}(\tau)$. This still leaves open the question of how $\tilde{G}_{\text{cont}}(p_0)$ should be chosen. For now, we simply state that we should choose $\tilde{G}_{\text{cont}}(p_0)$ so that $G_{\text{cont}}(\tau)$ is strongly damped for large τ. By a standard theorem of Fourier transformations, we can accomplish this if $\tilde{G}_{\text{cont}}(p_0)$ and all its derivatives are continuous and absolutely integrable over the real p_0 line [89].

If $G_{\text{cont}}(t)$ is exponentially damped for $mt \gg 1$, we can, in practical situations, neglect most or all of the $n \neq 0$ terms in Eq. (A7). A standard approach is just to include $n = -1$ in addition to $n = 0$, so that we include a backward propagating meson in our fit Ansatz for $G(t)$:

$$G(t) \sim C(e^{-Mt} + e^{-M(T-t)}). \qquad (A9)$$

The fit for $mt \gg 1$ will then effectively isolate the first contribution, from $G_{\text{cont}}(t)$, and extract the corrected mass M from its exponential decay. If Jordan's lemma applies to the Fourier transform and if the only single-particle pole in $\tilde{G}(p_0)$ is the usual one near $p_0 = im$, then the correction to the squared mass is indeed just $\sigma_{\text{cont}}(im)$. We will see below, however, that this will not be true in general.

So we are led to consideration of the finite-T self energy in momentum space $\sigma(p_0)$, and how it may be continued away from the special values $p_0 = 2\pi\ell/T$ to $\sigma_{\text{cont}}(p_0)$. A natural choice for $\sigma_{\text{cont}}(p_0)$ is simply the result of doing the loop energy/momenta sums for arbitrary external p_0, instead of only for the special values. For example, we can perform the sum in the first term on the right-hand side

of Eq. (63) for any p_0. Because the resulting self-energy function and its derivatives obey the continuity and integrability conditions mentioned above, $G_{\text{cont}}(t)$ will automatically be exponentially damped[16] as desired.

An undesirable, but unavoidable, feature of this continuation $\sigma_{\text{cont}}(p_0)$ is that it depends on the routing of the external momentum p_0 through the diagram. The dependence on routing vanishes when $p_0 = 2\pi\ell/T$ because the loop energy may be shifted by this amount. But away from these special points, there is no reason for $\sigma_{\text{cont}}(p_0)$ to be independent of the routing; we have checked this dependence numerically for

$$\sigma_{\text{cont}}(p_0) = \frac{1}{L^3 T} \sum_{k_0, \vec{k}}' \mathcal{I}_{\hat{s}} - \int \frac{d^4 k}{(2\pi)^4} \mathcal{I}_{\hat{s}}, \qquad (A10)$$

with $\mathcal{I}_{\hat{s}}$ the photon-sunset integrand given by Eq. (62). Here we have considered the difference between the sum and the integral, rather than the sum itself, to avoid having to cut off the sum over $\vec{k}$, which is irrelevant to the current discussion.[17] Further, the dependence on momentum routing persists when σ_{cont} is evaluated at $p_0 = im$, which indicates that the rule relating the mass correction to $\sigma_{\text{cont}}(im)$ cannot be true in general. We emphasize that this is a problem with the rule, rather than some fundamental problem with the definition of the mass correction itself: The finite-T propagator $G(t)$ is of course completely independent of the routing.

To examine this issue further, we consider the two obvious possible momentum routings in the sunset diagram. Routing A, which we used in Sec. V, has $p - k$ on the photon line and k on the internal meson line. Routing B has k on the photon line and $p - k$ on the internal meson line. With $p = (p_0, \vec{0})$ and $\vec{k} \neq 0$,

$$\mathcal{I}_{\hat{s}}^A = \frac{\vec{k}^2 + m^2 - p_0^2}{\vec{k}^2 (k_0^2 + \vec{k}^2 + m^2)}, \qquad (A11)$$

$$\mathcal{I}_{\hat{s}}^B = \frac{\vec{k}^2 + m^2 + 2p_0 k_0 - 3p_0^2}{\vec{k}^2 ((k_0 - p_0)^2 + \vec{k}^2 + m^2)}, \qquad (A12)$$

where Eq. (A11) is copied from Eq. (62). In both cases we have added on the 00 component of the tadpole, which is independent of the external momentum. The linear term in k_0 in the numerator of $\mathcal{I}_{\hat{s}}^B$ cannot be dropped since the denominator is not symmetric under $k_0 \to -k_0$. When $p_0 = 2\pi\ell/T$, $\mathcal{I}_{\hat{s}}^A$ and $\mathcal{I}_{\hat{s}}^B$ clearly give the same result for

---

[16]More precisely, it will decrease faster than any power of $1/t$ for large t [89].

[17]From now on we use the term "mass correction" to mean the finite-L and finite-T contribution to the mass correction. The additional correction when T and L are infinite will not affect any of the following discussion, as long as that correction is small enough that it does not violate the perturbative expansion.
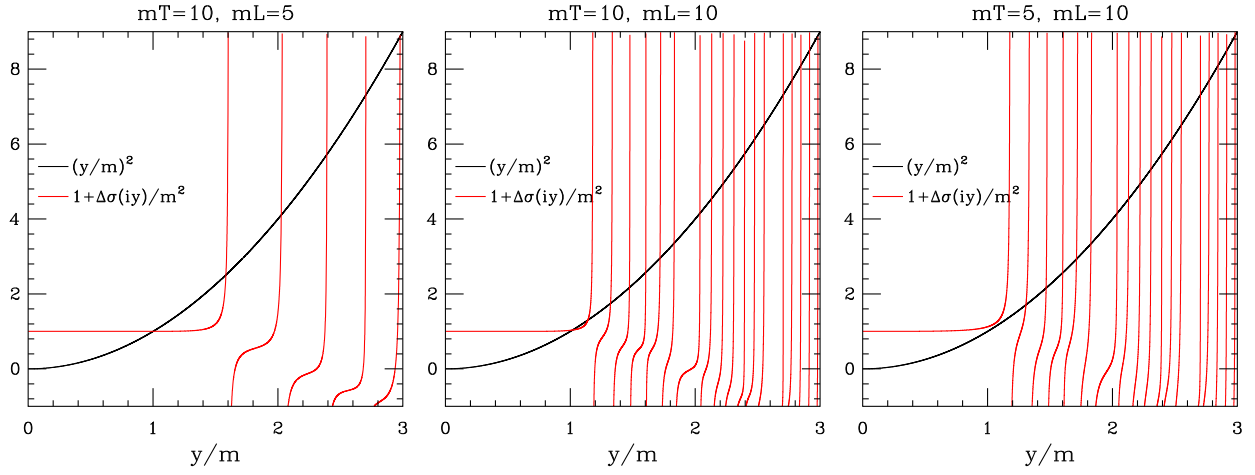
FIG. 20. Location of the poles of the momentum space propagator $\tilde{G}_{\text{cont}}(p_0)$ for routing B, for three different values of mT and mL. The quantities $(y/m)^2$ (black lines) and $1 + \Delta\sigma(iy)/m^2$ (red lines) are shown as a function of y/m, where y is the imaginary part of the Euclidean energy p_0. The y/m values of the poles are given by the locations of the crossings of the two curves.

$\sigma(p_0)$, as can be seen by shifting the summation variable $k_0 \rightarrow k_0 + p_0$ in $\mathcal{I}_{\hat{s}}^B$ (and then dropping a linear term in k_0 in the numerator).

Extracting the mass correction is easy for routing A. We can see from Eq. (A11) that $\sigma_{\text{cont}}^A(p_0)$ has the simple form $\alpha + \beta p_0^2$, where α and β are independent of p_0. Therefore, $\tilde{G}_{\text{cont}}^A(p_0)$ has a simple pole close to $p_0 = im$, and Jordan's lemma allows us to close the contour as usual in the upper half plane (for $t > 0$) for the Fourier transform of $\tilde{G}_{\text{cont}}^A(p_0)$. This determines the squared-mass correction to be $\sigma_{\text{cont}}^A(im)$, as was assumed in Sec. V.

Extracting the mass correction in the case of routing B is more subtle. To see the relation between the self energy from $\mathcal{I}_{\hat{s}}^B$ and $\mathcal{I}_{\hat{s}}^A$ when p_0 is not at a special point, we use the Poisson summation formula to write

$$\sigma_{\text{cont}}^B(p_0) = \frac{1}{L^3} \sum_{\vec{k}}{}' \sum_n \int \frac{dk_0}{2\pi} e^{inTk_0} \mathcal{I}_{\hat{s}}^B - \int \frac{d^4k}{(2\pi)^4} \mathcal{I}_{\hat{s}}^B,$$

(A13)

where n runs over the integers. We can now make the shift $k_0 \rightarrow k_0 + p_0$ in both integrals, converting $\mathcal{I}_{\hat{s}}^B$ into $\mathcal{I}_{\hat{s}}^A$. Differences remain, however, from the resulting phase e^{inTp_0} and from the term in the numerator linear in k_0, which gives a nonvanishing contribution when $n \neq 0$. The difference between the self-energies is then

$$\Delta\sigma(p_0) = -\frac{2}{L^3} \sum_{n\geq 1} \sum_{\vec{k}}{}' \frac{e^{-\omega_k nT}}{\vec{k}^2}$$
$$\times \left[\sin^2(nTp_0/2) \frac{\vec{k}^2 + m^2 - p_0^2}{\omega_k} + p_0 \sin(nTp_0) \right],$$

(A14)

where $\Delta\sigma \equiv \sigma_{\text{cont}}^B - \sigma_{\text{cont}}^A$, and $\omega_k \equiv \sqrt{\vec{k}^2 + m^2}$.

Because of the additional factors of $\sin^2(nTp_0/2)$ and $\sin(nTp_0)$, which blow up for large imaginary p_0, the analytic properties of σ_{cont}^B are not standard, and we must reexamine the usual assumptions that go into finding the mass. To simplify the algebra we take $mT \gg 1$ and consider $p_0 = iy$ with $y \gtrsim m$. We can therefore neglect the exponentially falling terms $\exp(-ynT)$ from the sine functions and keep only the growing ones. The sum over n then immediately gives

$$\Delta\sigma(iy) = \frac{1}{2L^3} \sum_{\vec{k}}{}' \frac{(\omega_k + y)^2}{\vec{k}^2 \omega_k (e^{T(\omega_k - y)} - 1)}. \quad \text{(A15)}$$

The values of ω_k for each discrete value of $\vec{k}$ therefore determine singularities in $\Delta\sigma$. As y approaches a discrete value of ω_k from below, $\Delta\sigma$ goes to $+\infty$, and then comes up from $-\infty$ as y increases above ω_k. Because the self-energy varies over the full range $(-\infty, \infty)$, $\tilde{G}_{\text{cont}}^B$ will have a pole near each of the singularities in $\Delta\sigma$. Figure 20 shows how this occurs for three choices of mT and mL. The equation for the poles is $y^2 = m^2 + \sigma_{\text{cont}}^B(iy)$, which we find from the crossings of the curves $(y/m)^2$ and $1 + \Delta\sigma(iy)/m^2$, where we have neglected the difference between $\Delta\sigma$ and σ_{cont}^B. This difference is σ_{cont}^A, which just gives a relatively small correction to the terms $(y/m)^2$ and 1, and does not change the qualitative picture. In the plots, we have included all the $\vec{k}$ values in the sum in Eq. (A15) that contribute significantly in the region of y/m shown.

The left-hand plot ($mT = 10$, $mL = 5$) shows that, in addition to the "normal" pole close to $y = m$, there are anomalous poles close the singular values in $\Delta\sigma$ where $y = \omega_k$, for some $\vec{k}$. As L increases, the possible values of k get closer, and the poles get denser. We observe this feature in the middle plot ($mT = 10$, $mL = 10$). As $L \rightarrow \infty$, the

poles pile up at $y = m$. On the other hand, as mT gets smaller, the residues of the singularities in $\Delta\sigma$ increase like $1/(mT)$. This can lead to the particularly strange situation where the normal pole in $\tilde{G}_{\mathrm{cont}}$ close to $y = m$ disappears, as shown in the right-hand plot ($mT = 5$, $mL = 10$).

Because there are many poles in the propagator, and often many of them are close to $p_0 = im$, the quantity $\sigma_{\mathrm{cont}}^B(im)$ has no direct relation to the mass correction. Nevertheless, the finite-T propagator $G(t)$ is always well defined, and in principle one could always extract the mass from $G(t)$ numerically. The relation between the self-energy at $p_0 = im$ and the mass correction is however problematic, and it seems unlikely in most cases that the dependence of the self-energy on p_0 will be simple enough to relate the mass-correction to the self energy at $p_0 = im$, as we did for $\sigma_{\mathrm{cont}}^A(im)$.

It is worth making contact here with the argument given in Ref. [72] about the effect of finite T. They write the difference between finite and infinite T for arbitrary momentum routing as in Eq. (A13)

$$\delta\sigma_{\mathrm{cont}}(p_0) = \frac{1}{L^3}{\sum_{\vec{k}}}'{\sum_{n}}'\int\frac{dk_0}{2\pi}e^{inTk_0}\mathcal{I}(k_0,\vec{k},p_0), \quad \text{(A16)}$$

where the prime on the sum on n indicates that $n = 0$ should be omitted; it is cancelled by the infinite-T subtraction. They then argue that $\mathcal{I}(k_0,\vec{k},im)$ has no poles on the real k_0 axis and is infinitely differentiable, with all of its derivatives integrable, which implies that $\delta\sigma_{\mathrm{cont}}(im)$ vanishes faster than any power of $1/T$ as $T \to \infty$. This argument explains why the QED_L FV correction $\delta_{\mathrm{FV}}^{\gamma,\mathrm{QED}_L}$ shows negligible dependence on T for the values of mT relevant to Fig. 6. Indeed, using routing A, the unique single-particle pole in $\tilde{G}_{\mathrm{cont}}$ near $p_0 = im$ implies that the leading T-dependence in QED_L is suppressed by a factor of $\exp(-mT)$.[18]

However, the Borsanyi *et al.* argument does not apply in general for routings that generate complicated p_0 dependence away from the discrete points $2\pi\ell/T$. In particular, the argument cannot be use to conclude that routing-dependent differences in $\sigma_{\mathrm{cont}}(im)$ are similarly suppressed by $\exp(-mT)$ and therefore negligible for values of mT used in our computation. To see this, we look at a simple example with

$$\mathcal{I}(k_0,\vec{k},p_0) = \frac{1}{(k_0 + p_0)^2 + \vec{k}^2 + m^2}. \quad \text{(A17)}$$

With $p_0 = iy$ and $y \leq \omega_k$ (for fixed $\vec{k}$), there is a simple pole in the upper half plane at $k_0 = i(\omega_k - y)$. When

$y = m$, the $n \geq 1$ terms in the sum give contributions (after integration over k_0) proportional to $\exp(-nT(\omega_k - m))$, while the $n \leq -1$ contributions are more highly suppressed for large T since the pole in the lower half-plane is further from the axis. Because $\vec{k} \neq 0$, all the terms in the sum over $\vec{k}$ indeed decay exponentially with T. The rate of decay, however, can be very small for large L, because the lowest momenta have magnitude $2\pi/L$. Thus it is not obvious that the difference between finite T and infinite T can be neglected, even if one just focusses on $\sigma_{\mathrm{cont}}(im)$. More importantly, there are poles in $\sigma_{\mathrm{cont}}(iy)$ for $y = \omega_k$ (for some $\vec{k}$) as the k_0 pole in $\mathcal{I}$ moves down to the real axis. These poles mean that $-p_0^2 = y^2 = m^2 + \sigma_{\mathrm{cont}}(iy)$ can have multiple solutions, so there are multiple poles in the momentum space propagator $\tilde{G}(p_0)$, as we have seen in Fig. 20. If the higher poles are close to $y = m$ [as in Fig. 20 (middle)], or the $y \approx m$ pole is absent entirely [as in Fig. 20 (right)], $\sigma_{\mathrm{cont}}(im)$ will have little to do with the finite-T mass correction. Unfortunately, it is likely that the generic case will be like routing B rather than routing A—it seems to be an accident that with routing A no p_0 dependence appears in the denominator of our integrand, so that $\sigma_{\mathrm{cont}}^A(p_0)$ is a simple (quadratic) polynomial in p_0.

Finally, as an estimate of how important these effects are for the actual simulation data, we study the routing dependence of the self energy at $p_0 = im$, coming from
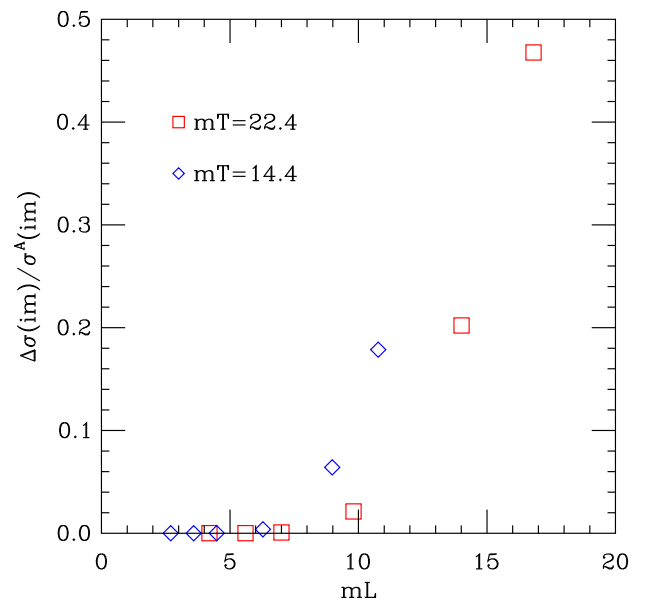

FIG. 21. Relative size of the routing dependence of the self-energy contribution at $p_0 = im$ of the sunset graph, for the same data at 0.12 fm that was presented in Fig. 9. The red squares are for "kaon" points; the blue diamonds, for "pion" points. For the three leftmost diamonds and the three leftmost squares, where no deviation from 0 is visible, the actual deviation varies between $\approx 5 \times 10^{-5}$ and $\approx 1 \times 10^{-4}$.

---

[18]Note that the extra term $\delta_{\mathrm{FV}}^{\mathrm{t,+}}(mL, mT)$ for QED_{TL}, given in Eq. (67), is not negligible for any of our data.

the photon sunset graph and 00 component of the photon tadpole. (As elsewhere in this Appendix, the spatial part of the photon tadpole is not included because it has no routing dependence.) In Fig. 21, we plot the ratio of $\Delta\sigma(im)/\sigma_{\mathrm{cont}}^{A}(im)$ vs mL for the data shown in Fig. 9

above. As expected from the above discussion, the dependence increases with mL for fixed mT, and decreases with mT for fixed mL. Note that, even though mT is large, the routing dependence is not negligible for much for our data, and approaches 50% for the largest values of mL.

[1] R. F. Dashen, Phys. Rev. **183,** 1245 (1969).
[2] S. Aoki et al. (FLAG Collaboration), Eur. Phys. J. C **77,** 112 (2017).
[3] See, e.g., J. Bijnens and J. Prades, Nucl. Phys. **B490,** 239 (1997); J. F. Donoghue and A. F. Perez, Phys. Rev. D **55,** 7075 (1997); B. Moussallam, Nucl. Phys. **B504,** 381 (1997).
[4] A. Bazavov et al., Rev. Mod. Phys. **82,** 1349 (2010).
[5] A. Bazavov et al., Proc. Sci. CD09 (**2009**) 007 [arXiv:0910.2966]; A. Bazavov et al. (MILC Collaboration), Proc. Sci. LAT2009 (**2009**) 079 [arXiv:0910.3618].
[6] J. Laiho and R. S. Van de Water, Proc. Sci. LAT2011 (**2011**) 293 [arXiv:1112.4861].
[7] C. Aubin et al. (MILC Collaboration), Phys. Rev. D **70,** 114501 (2004).
[8] A. Duncan, E. Eichten, and H. Thacker, Phys. Rev. Lett. **76,** 3894 (1996).
[9] T. Blum, T. Doi, M. Hayakawa, T. Izubuchi, and N. Yamada, Phys. Rev. D **76,** 114508 (2007).
[10] T. Blum, R. Zhou, T. Doi, M. Hayakawa, T. Izubuchi, S. Uno, and N. Yamada, Phys. Rev. D **82,** 094508 (2010).
[11] A. Portelli et al. (BMW Collaboration), Proc. Sci. LAT2011 (**2011**) 136 [arXiv:1201.2787]; Proc. Sci. LAT2010 (**2010**) 121 [arXiv:1011.4189].
[12] S. Basak et al. (MILC Collaboration), Proc. Sci. LAT2012 (**2012**) 137 [arXiv:1210.8157].
[13] S. Basak et al. (MILC Collaboration), Proc. Sci. CD12 (**2013**) 030 [arXiv:1301.7137].
[14] S. Basak et al. (MILC Collaboration), Proc. Sci. LAT2014 (**2014**) 116 [arXiv:1409.7139]; J. Phys. Conf. Ser. **640,** 012052 (2015).
[15] S. Basak et al. (MILC Collaboration), Proc. Sci. LAT2015 (**2016**) 259 [arXiv:1606.01228].
[16] R. Horsley et al., J. Phys. G **43,** 10LT02 (2016); J. High Energy Phys. 04 (2016) 093.
[17] Z. Fodor, C. Hoelbling, S. Krieg, L. Lellouch, Th. Lippert, A. Portelli, A. Sastre, K. K. Szabo, and L. Varnhorst, Phys. Rev. Lett. **117,** 082001 (2016).
[18] G. M. de Divitiis, R. Frezzotti, V. Lubicz, G. Martinelli, R. Petronzio, G. C. Rossi, F. Sanfilippo, S. Simula, and N. Tantalo (RM123 Collaboration), Phys. Rev. D **87,** 114505 (2013).
[19] D. Giusti, V. Lubicz, C. Tarantino, G. Martinelli, F. Sanfilippo, S. Simula, and N. Tantalo, Phys. Rev. D **95,** 114504 (2017).
[20] A. Portelli, Proc. Sci. LAT2014 (**2015**) 013 [arXiv: 1505.07057].
[21] J. Bijnens and N. Danielsson, Phys. Rev. D **75,** 014505 (2007).
[22] R. Urech, Nucl. Phys. **B433,** 234 (1995).
[23] J. Gasser and H. Leutwyler, Nucl. Phys. **B250,** 465 (1985).
[24] A. Bazavov et al. (MILC Collaboration), Phys. Rev. D **87,** 054505 (2013).
[25] C. Bernard and E. D. Freeland, Proc. Sci. LAT2010 (**2010**) 084 [arXiv:1011.3994].
[26] A. Bazavov et al. (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **90,** 074509 (2014).
[27] S. Basak et al. (MILC Collaboration), Proc. Sci. LAT2008 (**2008**) 127 [arXiv:0812.4486]; A. Torok et al. (MILC Collaboration), Proc. Sci. LAT2010 (**2010**) 127.
[28] S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, K. K. Szabo, and G. Vulvert, Phys. Lett. B **701,** 265 (2011); J. High Energy Phys. 08 (2011) 148.
[29] S. A. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, Phys. Rev. D **35,** 2531 (1987).
[30] S. Duane, Nucl. Phys. **B257,** 652 (1985).
[31] S. Duane and J. B. Kogut, Nucl. Phys. **B275,** 398 (1986).
[32] M. A. Clark and A. D. Kennedy, Phys. Rev. Lett. **98,** 051601 (2007).
[33] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195,** 216 (1987).
[34] J. C. Sexton and D. H. Weingarten, Nucl. Phys. **B380,** 665 (1992).
[35] I. Horvath, A. D. Kennedy, and S. Sint, Nucl. Phys. B, Proc. Suppl. **73,** 834 (1999).
[36] M. Hasenbusch, Phys. Lett. B **519,** 177 (2001).
[37] I. P. Omelyan, I. M. Mryglod, and R. Folk, Phys. Rev. E **66,** 026701 (2002).
[38] T. Takaishi and P. de Forcrand, Phys. Rev. E **73,** 036706 (2006).
[39] M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi, Comput. Phys. Commun. **181,** 1517 (2010); R. Babich et al., in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC) (ACM, New York, 2011).
[40] G. Shi, S. Gottlieb, A. Torok, and V. Kindratenko, Design of MILC lattice QCD application for GPU clusters, in Proceedings of the 2011 IEEE Parallel and Distributed Processing Symposium, IPDPS '11, 2011 (IEEE, New York, 2011).
[41] G. Shi, R. Babich, M. Clark, B. Joó, S. Gottlieb, and V. Kindratenko, The fat-link computation on large GPU clusters for lattice QCD, in Proceedings of the 2012 IEEE Parallel and Distributed Processing Symposium, IPDPS '12, 2012 (IEEE, New York, 2012).
[42] R. Babich, M. A. Clark, B. Joo, G. Shi, R. C. Brower, and S. Gottlieb, Scaling lattice QCD beyond 100 GPUs, in

Proceedings of the SC11 International Conference for High Performance Computing, Networking, Storage and Analysis Seattle, Washington, 2011 (ACM, New York, 2011), DOI: 10.1145/2063384.2063478.

[43] R. Babich, R. Brower, M. Clark, S. Gottlieb, B. Joo, and G. Shi, Proc. Sci. LAT2011 (**2011**) 033.

[44] C. W. Bernard, T. Burch, K. Orginos, D. Toussaint, T. A. DeGrand, C. DeTar, S. Gottlieb, U. M. Heller, J. E. Hetrick, and B. Sugar, Phys. Rev. D **62**, 034503 (2000).

[45] R. Sommer, Nucl. Phys. **B411**, 839 (1994).

[46] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **85**, 114506 (2012).

[47] E. Marinari, G. Parisi, and C. Rebbi, Nucl. Phys. **B190**, 734 (1981).

[48] E. Follana, A. Hart, and C. T. H. Davies, Phys. Rev. Lett. **93**, 241601 (2004).

[49] S. Dürr, C. Hoelbling, and U. Wenger, Phys. Rev. D **70**, 094502 (2004).

[50] Y. Shamir, Phys. Rev. D **71**, 034509 (2005).

[51] S. Dürr and C. Hoelbling, Phys. Rev. D **71**, 054501 (2005).

[52] K. Y. Wong and R. M. Woloshyn, Phys. Rev. D **71**, 094508 (2005).

[53] S. Prelovsek, Phys. Rev. D **73**, 014506 (2006).

[54] C. Bernard, Phys. Rev. D **73**, 114503 (2006).

[55] C. Bernard, M. Golterman, and Y. Shamir, Phys. Rev. D **73**, 114511 (2006).

[56] S. Dürr and C. Hoelbling, Phys. Rev. D **74**, 014513 (2006).

[57] Y. Shamir, Phys. Rev. D **75**, 054503 (2007).

[58] C. Bernard, C. DeTar, Z. Fu, and S. Prelovsek, Phys. Rev. D **76**, 094504 (2007).

[59] G. Donald, C. T. H. Davies, E. Follana, and A. S. Kronfeld, Phys. Rev. D **84**, 054504 (2011).

[60] W. J. Lee and S. R. Sharpe, Phys. Rev. D **60**, 114503 (1999).

[61] C. Aubin and C. Bernard, Phys. Rev. D **68**, 034014 (2003).

[62] C. W. Bernard and M. F. L. Golterman, Phys. Rev. D **49**, 486 (1994).

[63] S. R. Sharpe and N. Shoresh, Phys. Rev. D **62**, 094503 (2000).

[64] S. R. Sharpe and N. Shoresh, Phys. Rev. D **64**, 114510 (2001).

[65] P. H. Damgaard and K. Splittorff, Phys. Rev. D **62**, 054509 (2000).

[66] C. Aubin and C. Bernard, Nucl. Phys. B, Proc. Suppl. **129–130**, 182 (2004).

[67] S. R. Sharpe, Phys. Rev. D **46**, 3146 (1992); Nucl. Phys. B, Proc. Suppl. **17**, 146 (1990).

[68] C. Aubin *et al.* (HPQCD, MILC, and UKQCD Collaborations), Phys. Rev. D **70**, 031504 (2004).

[69] G. P. Lepage and P. B. Mackenzie, Phys. Rev. D **48**, 2250 (1993).

[70] Q. Mason, H. D. Trottier, R. Horgan, C. T. H. Davies, and G. P. Lepage (HPQCD Collaboration), Phys. Rev. D **73**, 114501 (2006).

[71] S. Borsanyi *et al.* (Budapest-Marseille-Wuppertal Collaboration), Phys. Rev. Lett. **111**, 252001 (2013).

[72] S. Borsanyi *et al.*, Science **347**, 1452 (2015).

[73] S. Uno and M. Hayakawa, Prog. Theor. Phys. **120**, 413 (2008).

[74] Z. Davoudi and M. J. Savage, Phys. Rev. D **90**, 054503 (2014).

[75] Z. Fodor, C. Hoelbling, S. D. Katz, L. Lellouch, A. Portelli, K. K. Szabo, and B. C. Toth, Phys. Lett. B **755**, 245 (2016).

[76] J. W. Lee and B. C. Tiburzi, Phys. Rev. D **93**, 034012 (2016).

[77] C. Bernard (MILC Collaboration), Phys. Rev. D **65**, 054031 (2002).

[78] See, for instance, the Wikipedia article on the Basel problem, https://en.wikipedia.org/wiki/Basel_problem.

[79] G. P. Lepage, J. Comput. Phys. **27**, 192 (1978).

[80] See, for instance, the Wolfram MathWorld entry on Jacobi theta functions, http://mathworld.wolfram.com/JacobiTheta Functions.html.

[81] M. G. Endres, A. Shindler, B. C. Tiburzi, and A. Walker-Loud, Phys. Rev. Lett. **117**, 072002 (2016).

[82] C. Aubin *et al.* (MILC Collaboration), Nucl. Phys. B, Proc. Suppl. **140**, 231 (2005).

[83] See, e.g., D. F. Morrison, *Multivariate Statistical Methods* (McGraw-Hill, New York, 1967); R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis* (Prentice-Hall, New York and London, 1982); D. Toussaint and W. Freeman, arXiv:0808.2211.

[84] C. Patrignani *et al.* (Review of Particle Physics), Chin. Phys. C **40**, 100001 (2016).

[85] A. Bazavov *et al.*, Phys. Rev. D **98**, 074512 (2018).

[86] A. Bazavov *et al.* (Fermilab Lattice, MILC, and TUMQCD Collaborations), Phys. Rev. D **98**, 054517 (2018).

[87] N. Carrasco *et al.* (European Twisted Mass Collaboration), Nucl. Phys. **B887**, 19 (2014).

[88] R. Zhou *et al.* (MILC Collaboration), Proc. Sci. LAT2014 (**2014**) 024 [arXiv:1411.4115]; Y. Liu *et al.* (MILC Collaboration), https://indico.fnal.gov/event/15949/session/11/contribution/209/material/slides/0.pdf.

[89] See, for instance, C. Gasquet and P. Witomski, *Fourier Analysis and Applications*, translated by R. Ryan (Springer, New York, 1999), p. 172.