

| $V_{us}$ | from  $K_{\ell 3}$  decay and four-flavor lattice QCD

A. Bazavov,<sup>1</sup> C. Bernard,<sup>2</sup> C. DeTar,<sup>3</sup> Daping Du,<sup>4</sup> A. X. El-Khadra,<sup>5,6</sup> E. D. Freeland,<sup>7</sup> E. Gámiz,<sup>8,\*</sup> Steven Gottlieb,<sup>9</sup> U. M. Heller,<sup>10</sup> J. Komijani,<sup>11,12,13</sup> A. S. Kronfeld,<sup>6,13</sup> J. Laiho,<sup>4</sup> P. B. Mackenzie,<sup>6</sup> E. T. Neil,<sup>14,15</sup> T. Primer,<sup>16</sup> J. N. Simone,<sup>6</sup> R. Sugar,<sup>17</sup> D. Toussaint,<sup>16</sup> and R. S. Van de Water<sup>6</sup>

(Fermilab Lattice and MILC Collaborations)

<sup>1</sup>Department of Computational Mathematics, Science and Engineering, and Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48823, USA

<sup>2</sup>Department of Physics, Washington University, St. Louis, Missouri 63130, USA

<sup>3</sup>Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah 84112, USA

<sup>4</sup>Department of Physics, Syracuse University, Syracuse, New York 13244, USA

<sup>5</sup>Department of Physics, University of Illinois, Urbana, Illinois 61801, USA

<sup>6</sup>Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA

<sup>7</sup>School of the Art Institute of Chicago, Chicago, Illinois 60603, USA

<sup>8</sup>CAFPE and Departamento de Física Teórica y del Cosmos, Universidad de Granada, 18071 Granada, Spain

<sup>9</sup>Department of Physics, Indiana University, Bloomington, Indiana 47405, USA

<sup>10</sup>American Physical Society, Ridge, New York 11961, USA

<sup>11</sup>School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom

<sup>12</sup>Physik-Department, Technische Universität München, 85748 Garching, Germany

<sup>13</sup>Institute for Advanced Study, Technische Universität München, 85748 Garching, Germany

<sup>14</sup>Department of Physics, University of Colorado, Boulder, Colorado 80309, USA

<sup>15</sup>RIKEN-BNL Research Center, Brookhaven National Laboratory, Upton, New York 11973, USA

<sup>16</sup>Department of Physics, University of Arizona, Tucson, Arizona 85721, USA

<sup>17</sup>Department of Physics, University of California, Santa Barbara, California 93106, USA



(Received 4 January 2019; published 24 June 2019)

Using highly improved staggered quark (HISQ)  $N_f = 2 + 1 + 1$  MILC ensembles with five different values of the lattice spacing, including four ensembles with physical quark masses, we perform the most precise computation to date of the  $K \rightarrow \pi \ell \nu$  vector form factor at zero momentum transfer,  $f_+^{K^0 \pi^-}(0) = 0.9696(15)_{\text{stat}}(12)_{\text{syst}}$ . This is the first calculation that includes the dominant finite-volume effects, as calculated in chiral perturbation theory at next-to-leading order. Our result for the form factor provides a direct determination of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element  $|V_{us}| = 0.22333(44)_{f_+(0)}(42)_{\text{exp}}$ , with a theory error that is, for the first time, at the same level as the experimental error. The uncertainty of the semileptonic determination is now similar to that from leptonic decays and the ratio  $f_{K^+}/f_{\pi^+}$ , which uses  $|V_{ud}|$  as input. Our value of  $|V_{us}|$  is in tension at the  $2\text{--}2.6\sigma$  level both with the determinations from leptonic decays and with the unitarity of the CKM matrix. In the test of CKM unitarity in the first row, the current limiting factor is the error in  $|V_{ud}|$ , although a recent determination of the nucleus-independent radiative corrections to superallowed nuclear  $\beta$  decays could reduce the  $|V_{ud}|^2$  uncertainty nearly to that of  $|V_{us}|^2$ . Alternative unitarity tests using only kaon decays, for which improvements in the theory and experimental inputs are likely in the next few years, reveal similar tensions and could be further improved by taking correlations between the theory inputs. As part of our analysis, we calculated the correction to  $f_+^{K\pi}(0)$  due to nonequilibrated topological charge at leading order in chiral perturbation theory, for both the full-QCD and the partially quenched cases. We also obtain the combination of low-energy constants in the chiral effective Lagrangian  $[C_{12}^r + C_{34}^r - (L_5^r)^2](M_\rho) = (2.92 \pm 0.31) \times 10^{-6}$ .

DOI: 10.1103/PhysRevD.99.114509

\*megamiz@ugr.es

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

## I. INTRODUCTION

High-precision tests of the unitarity of the Cabibbo-Kobayashi-Maskawa (CKM) matrix, as predicted by the Standard Model (SM), are at the forefront of the current flavor physics program. Any violation of the unitarity of the CKM matrix, which describes flavor-changing interactions, would be evidence of the existence of physics beyond the Standard Model (BSM).

In particular, first-row unitarity, which requires that

$$\Delta_u \equiv |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 - 1 \quad (1.1)$$

vanish, is currently the most precisely tested condition. Even in the absence of deviations, high-precision determinations of the CKM matrix elements involved in the test in Eq. (1.1) put important constraints on the scale of the allowed new physics [1].

At the current level of precision one can neglect  $|V_{ub}|^2$  in Eq. (1.1). Of the other two CKM matrix elements involved,  $|V_{ud}|$  is precisely determined from superallowed nuclear  $\beta$  decays [2]. It can also be extracted from measurements of the neutron lifetime [3] and pion  $\beta$  decay [4], albeit with much larger errors [5]. Improved experimental measurements of these processes would be interesting because they are theoretically cleaner.

The best determinations of  $|V_{us}|$  are from kaon decays [6]. The extraction of  $|V_{us}|$  from semileptonic kaon ( $K_{\ell 3}$ ) decay requires knowledge of the form factor at zero momentum transfer,  $f_+^{K\pi}(0)$ , which is still the largest source of uncertainty on  $|V_{us}|$ . On the experimental side, it is expected that the ongoing and forthcoming experiments (NA62, OKA, KLOE-2, LHCb and TREK E36) could reduce the experimental error to  $\sim 0.12\%$  within 5 years [7]. Reducing the theoretical error in the vector form-factor calculation is therefore a crucial task: it is this task that we take up in this paper.

Determinations of  $|V_{us}|$  from leptonic kaon and pion decays ( $K_{\ell 2}$  and  $\pi_{\ell 2}$ ), combined with  $f_K/f_\pi$  from lattice QCD, currently have somewhat smaller errors than those from  $K_{\ell 3}$ . The total error in  $|V_{us}|$  from leptonic decays is 0.25% [8–16], while from semileptonic decays it is 0.34% [6]. These leptonic determinations are indirect, however, because they require an external input for  $|V_{ud}|$ , namely Ref. [2]. The direct extraction of  $|V_{us}|$  from only kaon leptonic decays using  $f_K$  as nonperturbative input gives a larger error of 0.46%.<sup>1</sup>

Currently, the value of  $|V_{us}|$  obtained from leptonic kaon decay is  $\sim 2\sigma$  larger than the value from semileptonic kaon decay [16]. The leptonic decay is mediated by the axial-vector current while the semileptonic decay by the vector current. According to the SM, both approaches should give

<sup>1</sup>This error is based on the  $N_f = 2 + 1$  FLAG average for  $f_K$  [6], which includes only calculations which do not use  $f_\pi$ , and thus  $|V_{ud}|$ , to set the lattice scale [10,13,15].

the same  $|V_{us}|$ , because the  $W$  boson current has a  $V - A$  structure. Thus, any significant difference should be carefully analyzed.

In addition, if  $|V_{ud}|$  is taken from Ref. [2], the leptonic value of  $|V_{us}|$  is consistent with unitarity, Eq. (1.1), but the semileptonic value of  $|V_{us}|$  leads to a  $\sim 2\sigma$  disagreement with unitarity. As we were finishing this work, a paper appeared with a new calculation of the nucleus-independent electroweak radiative corrections involved in the extraction of  $|V_{ud}|$  from superallowed  $\beta$  decays with a new approach based on dispersion relations [17]. If this calculation is confirmed, the resulting value of  $|V_{ud}|$  would increase the present tension with unitarity. Investigating the origin of these tensions and performing even more stringent tests is crucial for the internal consistency of the Standard Model. It is thus necessary to reduce the error on both the experimental and the lattice-QCD inputs entering determinations of  $|V_{us}|$ .

In this paper, we focus on semileptonic kaon decay. The (photon-inclusive) decay rate for  $K^0$  can be written [18]

$$\Gamma(K^0 \rightarrow \pi^- \ell^+ \nu_\ell(\gamma)) = \frac{G_F^2 m_K^5}{128\pi^3} S_{\text{EW}} |V_{us} f_+^{K^0\pi^-}(0)|^2 I_{K^0\ell}^{(0)} (1 + \delta_{\text{EM}}^{K^0\ell} + \delta_{\text{SU}(2)}^{K^0\pi^-}), \quad (1.2)$$

where  $G_F$  is the Fermi constant as determined by muon decay,  $S_{\text{EW}} = 1.0232(3)$  is the universal short-distance electroweak correction [19–21],<sup>2</sup> and  $I_{K^0\ell}^{(0)}$  is a phase-space integral which depends on the shape of the  $f_+^{K^0\pi^+}(q^2)$  form factors given in Eq. (1.4) below. The long-distance electromagnetic corrections are parametrized by  $\delta_{\text{EM}}^{K^0\ell}$ . The strong isospin-breaking parameter  $\delta_{\text{SU}(2)}^{K\pi}$  is defined as a correction with respect to the  $K^0$  decay:

$$\delta_{\text{SU}(2)}^{K\pi} = \left( \frac{f_+^{K\pi}(0)}{f_+^{K^0\pi^-}(0)} \right)^2 - 1, \quad (1.3)$$

so that  $\delta_{\text{SU}(2)}^{K^0\pi^-} \equiv 0$ . The  $K^+$  decay rate,  $\Gamma(K^+ \rightarrow \pi^0 \ell^+ \nu_\ell(\gamma))$ , can be obtained by multiplying the right-hand side of Eq. (1.2) with the Clebsch-Gordan coefficient  $C_{K^+}^2 = 1/2$  and replacing  $I_{K^0\ell}^{(0)}$ ,  $\delta_{\text{EM}}^{K^0\ell}$ , and  $\delta_{\text{SU}(2)}^{K^0\pi^-}$  with the analogous  $I_{K^+\ell}^{(0)}$ ,  $\delta_{\text{EM}}^{K^+\ell}$ , and  $\delta_{\text{SU}(2)}^{K^+\pi^0}$ . The long-distance electromagnetic corrections, which are mode dependent, were calculated to  $\mathcal{O}(e^2 p^2)$  in Ref. [22] and are incorporated into the experimental average for  $|V_{us}| f_+^{K^0\pi^-}(0)$ , adding a 0.11% uncertainty to the experimental errors.

<sup>2</sup>This value of  $S_{\text{EW}}$  is from Ref. [21]. We use it because it is the value used for the experimental average in Ref. [7].

The input needed from lattice QCD in Eq. (1.2) is the vector form factor at zero momentum transfer,  $f_+^{K^0\pi^-}(q^2 = 0)$ , defined by

$$\begin{aligned} \langle \pi^+ | V^\mu | K^0 \rangle &= f_+^{K^0\pi^-}(q^2) [p_K^\mu + p_\pi^\mu] + f_-^{K^0\pi^-}(q^2) [p_K^\mu - p_\pi^\mu] \\ &= f_+^{K^0\pi^-}(q^2) \left[ p_K^\mu + p_\pi^\mu - \frac{m_K^2 - m_\pi^2}{q^2} q^\mu \right] \\ &\quad + f_0^{K^0\pi^-}(q^2) \frac{m_K^2 - m_\pi^2}{q^2} q^\mu, \end{aligned} \quad (1.4)$$

where  $V^\mu = \bar{s}\gamma^\mu u$  and  $q \equiv p_K - p_\pi$ .

The most precise value for  $f_+^{K^0\pi^-}(q^2 = 0)$  to date is provided by the  $N_f = 2 + 1 + 1$  Fermilab Lattice/MILC calculation in Refs. [23,24],  $f_+^{K^0\pi^-}(0) = 0.9704(\pm 0.33\%)$ . More recent lattice-QCD calculations by the RBC/UKQCD ( $N_f = 2 + 1$ ) [25] and ETMC ( $N_f = 2 + 1 + 1$ ) [26] collaborations agree very well with the Fermilab Lattice/MILC central value but with larger errors. Earlier  $N_f = 2 + 1$  calculations with unphysically heavy pions by the Fermilab Lattice/MILC [27] and RBC/UKQCD [28,29] collaborations, as well as the more recent JLQCD calculation in Ref. [30], yielded smaller values for  $f_+^{K^0\pi^-}(0)$ , but with larger errors. With the exception of the earliest calculation [29], these  $N_f = 2 + 1$  results are compatible with the newer  $N_f = 2 + 1 + 1$  ones. For comparison, the average of the relevant experimental input [7],  $|V_{us}|f_+^{K^0\pi^-}(0) = 0.21654(41)$ , has a 0.19% error. This average includes the strong isospin and electromagnetic corrections in Eq. (1.2) for each decay mode.

In this work we reduce the main sources of uncertainty in our previous calculation of  $f_+^{K^0\pi^-}(q^2 = 0)$  to reach a total error of 0.19%, obtaining the most precise calculation to date, and matching the current experimental uncertainty, for the first time. The main improvements over our previous calculation [23,24] are increased statistics in some key ensembles, the addition of a new (smaller) lattice spacing, and the correction of finite-volume effects at next-to-leading order (NLO) in chiral perturbation theory (ChPT). Preliminary results were presented in Refs. [31,32].

There are other ways to determine  $|V_{us}|$ . Semileptonic hyperon decays unfortunately lack sufficiently precise knowledge of the SU(3)-breaking corrections, which precludes a competitive determination. A conservative estimate of such effects yields an uncertainty of  $\sim 2\%$  [33]. Inclusive hadronic  $\tau$  decays have, in the past, yielded values of  $|V_{us}|$  smaller than the semileptonic kaon determination and, thus, were in even more disagreement with unitarity [34]. A more recent analysis [35] uses lattice QCD to compute dimension-larger-than-4 condensates and, more importantly, employs a dispersive technique to obtain the  $K\pi$  branching fractions. It points to an inclusive- $\tau$  value of  $|V_{us}|$  compatible with unitarity [35], although it still remains on the low side. An even more promising approach

also based on inclusive strange hadronic  $\tau$  decay data is presented in Ref. [36]. Its basic ingredients are replacing the operator-product expansion in the relevant sum rules by lattice hadronic vacuum polarization functions and optimizing the weight functions to suppress contributions from the high-energy region, where the experimental data have poor precision. Preliminary results in Ref. [36] are compatible with both semileptonic and leptonic determinations (and thus with unitarity), but have larger errors than either. Because the total errors on  $|V_{us}|$  in Refs. [35–37] are dominated by experimental uncertainties, it is expected that these determinations will be significantly improved with new data from the Belle II experiment [38]. Determinations of  $|V_{us}|$  from exclusive  $\tau$  decays, which use the same nonperturbative inputs as the leptonic kaon decay determinations, namely  $f_{K^\pm}$  and  $f_{K^\pm}/f_{\pi^\pm}$ , are still in tension with unitarity [39], but this could also change with future experimental measurements.

This paper is organized as follows. In Sec. II we describe the methodology of the numerical lattice-QCD simulations and the details of the ensembles, actions, and correlation functions used. Section III shows how, following the ChPT approaches of Refs. [40] and [41], one can correct for leading-order finite-volume effects and for the effects of nonequilibrated topological charge on the ensemble with finest lattice spacing ( $a \approx 0.042$  fm), respectively. We discuss the joint chiral interpolation and continuum extrapolation of our data to the physical point in Sec. IV. Section V analyzes the statistical and systematic uncertainties. Final results for the form factor  $f_+^{K^0\pi^-}$ , as well as for the relevant  $O(p^6)$  low-energy constants, are presented in Sec. VI. In Sec. VII, we use our form-factor result to extract a value of  $|V_{us}|$  from kaon semileptonic experimental data and discuss the implications of this value for phenomenology. Finally, we present our conclusions and the prospects for further improvement in Sec. VIII.

## II. LATTICE SETUP AND ANALYSIS

The methodology in this work largely follows that of our previous work in Refs. [23,24,27]. The approach, pioneered by HPQCD [42], is based on the Ward-Takahashi identity relating the matrix elements of a vector current to that of the corresponding scalar density:

$$q^\mu \langle \pi | V_\mu^{\text{lat}} | K \rangle Z_V = (m_s - m_u) \langle \pi | S^{\text{lat}} | K \rangle Z_S, \quad (2.1)$$

with  $Z_V$  and  $Z_S$  the lattice renormalization factors for the vector current and scalar density, respectively, where the scalar density is defined as the product of the scalar current and the quark masses ( $m_s - m_u$ ). Working with staggered fermions, and choosing  $V_\mu^{\text{lat}}$  to be the partially conserved, taste singlet, vector current, and  $S^{\text{lat}}$  to be its divergence, we have  $Z_V = Z_S = 1$ . Thus,  $S^{\text{lat}}$  is a local, taste-singlet density, with the same flavor content as the vector current,

TABLE I. Parameters of the  $N_f = 2 + 1 + 1$  gauge-field ensembles used in this work, and details of the correlation functions generated. A dagger at the end of a row indicates an ensemble that is new since our work in Ref. [23]; an asterisk indicates that the statistics have been increased.  $N_{\text{conf}}$  is the number of configurations included in the analysis;  $N_{\text{src}}$  the number of time sources used on each configuration; and  $L$  and  $L_t$  the spatial and temporal sizes of the lattice, respectively. The column labeled  $T$  lists the source-sink separations for the three-point functions generated on each ensemble. The  $m_\pi$  values are in MeV, with  $m_{\pi,P}$  the Goldstone (pseudoscalar taste)  $\pi$  mass, and  $m_\pi^{\text{RMS}}$  the root-mean-squared (over all tastes)  $\pi$  mass. The ensemble with  $a \approx 0.12$  fm,  $m_l/m_s = 0.1$  and  $m_{\pi,P}L = 3.2$  is used solely for the study of finite-volume effects.

$\approx a(\text{fm})$	$m_l/m_s^{\text{sea}}$	$m_{\pi,P}L$	$L^3 \times L_t$	$N_{\text{conf}} \times N_{\text{src}}$	$T$	$am_s^{\text{sea}}$	$am_s^{\text{val}}$	$m_{\pi,P}$	$m_\pi^{\text{RMS}}$
0.15	0.035	3.2	$32^3 \times 48$	$1000 \times 4$	12,13,15,16,17,18	0.0647	0.069 05	130	314
0.12	0.2	4.5	$24^3 \times 64$	$1053 \times 8$	15,18,20,21,22	0.0509	0.0535	299	364
	0.1	3.2	$24^3 \times 64$	$1020 \times 8$	15,18,20,21,22	0.0507	0.053	221	303
	0.1	4.3	$32^3 \times 64$	$993 \times 4$	15,18,20,21,22	0.0507	0.053	216	299
	0.1	5.4	$40^3 \times 64$	$1029 \times 8$	15,18,20,21,22	0.0507	0.053	214	298
	0.035	3.9	$48^3 \times 64$	$945 \times 8$	15,18,20,21,22	0.0507	0.0531	133	246
	0.2	4.5	$32^3 \times 96$	$773 \times 4$	23,27,32,33,34	0.037	0.038	301	323
0.09	0.1	4.7	$48^3 \times 96$	$853 \times 4$	23,27,32,33,34	0.0363	0.038	215	221
	0.035	3.7	$64^3 \times 96$	$950 \times 8$	23,27,32,33,34	0.0363	0.0363	130	176
	0.06	0.2	$4.5$	$48^3 \times 144$	$1000 \times 8$	34,41,48,49,50	0.024	0.024	304
0.06	0.035	3.7	$96^3 \times 192$	$692 \times 6$	31,39,40,48,49	0.022	0.022	135	144
	0.042	0.2	$64^3 \times 192$	$432 \times 12$	40,52,53,64,65	0.0158	0.0158	294	296
									†

$S = \bar{s}u$ . With the above identity and the definition of the form factors in Eq. (1.4), one can extract the scalar form factor  $f_0(q^2)$  at any value of the momentum transfer  $q^2$  by using

$$f_0^{K\pi}(q^2) = \frac{m_s - m_u}{m_K^2 - m_\pi^2} \langle \pi | S | K \rangle_{q^2}. \quad (2.2)$$

In addition, a kinematic constraint requires  $f_+^{K\pi}(0) = f_0^{K\pi}(0)$ , so this relation can be employed to calculate  $f_+^{K\pi}(0)$  from 3-point correlation functions with a scalar insertion. As already discussed in our first work [27], the use of a local scalar density instead of a vector current has two main advantages: avoiding the use of a renormalization factor and avoiding the use of noisier correlation functions with either a nonlocal vector current or external non-Goldstone mesons [42,43].

### A. Lattice actions, parameters, and correlation functions

We perform our calculation on the highly improved staggered quark (HISQ)  $N_f = 2 + 1 + 1$  MILC configurations [44–46] with sea quarks simulated with the HISQ action [47]. We also employ the HISQ action for the valence quarks. We have already seen in our previous work that the use of the HISQ action greatly reduces discretization effects [23,24]. The charm-quark and strange-quark masses on the  $N_f = 2 + 1 + 1$  MILC configurations are always tuned to values close to the physical ones, while the light-quark masses vary between  $0.2m_s$  and  $m_s/27$ , with the latter approximately the physical value. In this work, we include data generated at five different values of the lattice spacing down to  $a \approx 0.042$  fm, with sea pion masses

ranging from 319 to 134 MeV. Table I lists the key parameters of the ensembles analyzed here and the correlation functions calculated on them. The ensembles include four with physical quark masses and  $a \approx 0.15, 0.12, 0.09, 0.06$  fm. Ensembles that are new since our analysis in Refs. [23,24] are marked with a dagger in the last column; those where we have increased the statistics are marked with an asterisk. Table I also lists the pseudoscalar-taste (physical) pion mass  $m_{\pi,P}$  and the root-mean-squared pion mass  $m_\pi^{\text{RMS}}$  for each ensemble. The difference is a measure of the dominant discretization effects, which arise from taste-changing interactions. As expected, they decrease rapidly as the lattice spacing is reduced. The data included in this analysis are graphically depicted in Fig. 1.

The structure of the three-point function with a scalar insertion that we generate to access the matrix element in Eq. (2.2) is the same as in our previous work [23,24,27]. We generate light quarks at a time slice  $t_{\text{src}}$  and extended strange propagators at a fixed distance  $T$  from the source. For each configuration we have  $N_{\text{src}}$  time sources placed at  $t_{\text{src}} = t_0, t_0 + L_t/N_{\text{src}}, t_0 + 2L_t/N_{\text{src}}, \dots$ , where  $L_t$  is the temporal length of the lattice. The time  $t_0$  varies randomly from configuration to configuration in an interval  $[0, L_t/N_{\text{src}}]$  to reduce autocorrelations. Roughly following Ref. [48] we use random-wall sources at the pion source time  $t_{\text{src}}$ . On that spatial time slice, we choose four stochastic color-vector fields from a Gaussian distribution, with support on all three colors, and compute light-quark propagators from each of the four sources. Between the source and the kaon sink at time  $t_{\text{src}} + T$ , we contract the extended strange propagator with a light propagator to form the scalar density. We then study the  $t$  dependence to isolate the desired matrix element.

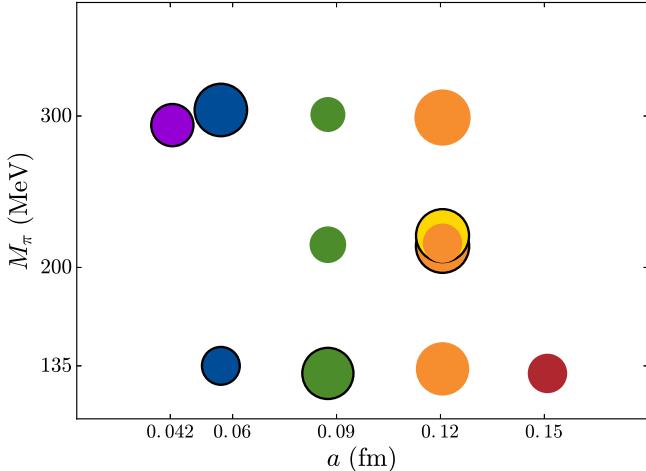


FIG. 1. Gauge-field ensembles analyzed in this work (parameters of these gauge-field ensembles are listed in Table I). The area of each disk is proportional to the statistical sample size  $N_{\text{conf}} \times N_{\text{src}}$ . Ensembles on which we have increased the statistics or we have added since our earlier work in Ref. [24] are indicated with black outlines. The three disks with  $a \approx 0.12$  fm and  $m_\pi \approx 200$  MeV correspond to the three ensembles with  $m_l/m_s^{\text{sea}} = 0.1$  and different volumes (smaller to larger from top to bottom) in Table I. The yellow disk (smallest volume) is not included in the final analysis but used only to study finite-volume effects.

The light-quark masses are always the same in the sea and valence sectors, while the sea and valence strange-quark masses are slightly different in some of the ensembles<sup>3</sup>; see Table I. Table I also lists the number of configurations and time sources on each ensemble. We compute three-point functions as described above for 5 or 6 different values of the source-sink separation  $T$ , listed in Table I, which correspond to approximately the same physical distances across ensembles. We include both even and odd values of  $T$  to disentangle the effects from oscillating states in the correlation functions.

We simulate directly at zero momentum transfer,  $q^2 \approx 0$ , by tuning the external momentum of the pion using partially twisted boundary conditions. In particular, we tune

$$|\vec{\theta}_2| = \frac{L}{\pi} \sqrt{\left( \frac{m_K^2 + m_\pi^2}{2m_K} \right)^2 - m_\pi^2}, \quad (2.3)$$

with  $\vec{\theta}_2$  the twist angle of the daughter propagator going from the pion to the current. The rest of the propagators are generated with periodic boundary conditions, the same as in the sea sector. We always have diagonal twist angles,

<sup>3</sup>At the time the analysis began, the physical value of  $am_s$  on those ensembles had been determined more accurately than when the ensembles were generated. We used the more accurate values for the valence strange-quark mass to be closer to the physical point.

TABLE II. Twisting angles and external momenta injected in the three-point functions. The quark masses  $am_l$  and  $am_s^{\text{sea}}$  are the same as in Table I, and  $\vec{\theta}_2$  is the twisting angle for the light daughter propagator in the pion, defined in Eq. (2.3). The superscript  $P$  in the pion masses refers to the pseudoscalar taste.

$0.25 \approx a(\text{fm})$	$m_l/m_s^{\text{sea}}$	$m_{\pi,P}L$	$ \vec{\theta}_2 $	$ a\vec{p}_P $
0.15	0.035	3.2	1.80966	0.17766
	0.2	4.5	0.84749	0.11094
	0.1	3.2	0.98192	0.12853
	0.1	4.3	1.30923	0.12853
	0.1	5.4	1.63653	0.12853
	0.035	3.9	2.16464	0.14168
0.09	0.2	4.5	0.82675	0.08117
	0.1	4.7	1.45024	0.09492
	0.035	3.7	2.08413	0.10230
0.06	0.2	4.5	0.81673	0.05345
	0.035	3.7	2.01756	0.06602
0.042	0.2	4.3	0.78006	0.03829

$\vec{\theta}_2 = |\vec{\theta}_2|(1, 1, 1)/\sqrt{3}$ , which turn out to give smaller finite-volume effects than twisting in only one direction [40]. The values of  $|\vec{\theta}_2|$  for each ensemble, as well as the corresponding momentum of the pion, are given in Table II.

For each ensemble, we generate zero-momentum two-point  $\pi$  and  $K$  correlation functions and two-point  $\pi$  correlation functions with external momentum given by the twist angle  $\vec{\theta}_2$ , defined in Eq. (2.3). These correlators are given by

$$C_{2\text{pt}}^P(\vec{p}; t) = \frac{1}{L^3} \sum_{\vec{x}} \sum_{\vec{y}} \langle \Phi_P^{\vec{p}}(\vec{y}, t + t_{\text{src}}) \Phi_P^{\vec{p}\dagger}(\vec{x}, t_{\text{src}}) \rangle, \quad (2.4)$$

where the interpolating operator  $\Phi_P^{\vec{p}\dagger}(\vec{x}, t)$  creates a meson  $P = \pi, K$  at time  $t$  with momentum  $\vec{p}$ . The random wall sources automatically implement the sum over  $\vec{x}$ . We also generate three-point correlation functions with the kaon at rest:

$$C_{3\text{pt}}^{K \rightarrow \pi}(\vec{p}_\pi, \vec{p}_K = 0; t, t_{\text{src}}, T) = \frac{1}{L^3} \sum_{\vec{x}, \vec{y}, \vec{z}} \langle \Phi_K^{\vec{p}_K=0}(\vec{x}, t_{\text{src}} + T) S(\vec{z}, t) \Phi_\pi^{\vec{p}_\pi\dagger}(\vec{y}, t_{\text{src}}) \rangle, \quad (2.5)$$

where the pion recoil momentum  $\vec{p}_\pi$  is either equal to zero or to the values listed in Table II. The scalar density is a local taste-singlet.

## B. Fit methods and statistical analysis

The fitting strategy we follow to extract the physical quantities from our correlation functions has already been discussed in Refs. [27,49]. We fit the two-point correlation functions for a pseudoscalar meson  $P$  to the expression

$$C_{2\text{pt}}^P(\vec{p}_P; t) = \sum_{m=0}^{N_{\text{exp}}} (-1)^{m(t+1)} (Z_m^P)^2 (e^{-E_P^m t} + e^{-E_P^m (L_t - t)}), \quad (2.6)$$

using Bayesian techniques. In Eq. (2.6),  $L_t$  is the temporal size of the lattice. The oscillating terms with  $(-1)^{m(t+1)}$  do not appear for a zero-momentum  $\pi$ . We fit the three-point correlation functions to

$$C_{3\text{pt}}^{K \rightarrow \pi}(\vec{p}_\pi, \vec{p}_K; t, T) = \sum_{m,n=0}^{N_{\text{exp}}^{\text{3pt}}} (-1)^{m(t+1)} (-1)^{n(T-t+1)} A^{mn}(q^2) Z_m^\pi Z_n^K \times (e^{-E_\pi^m t} + e^{-E_\pi^m (L_t - t)}) (e^{-E_K^n (T-t)} + e^{-E_K^n (L_t - T+t)}), \quad (2.7)$$

where the pion and kaon energies and amplitudes,  $E_\pi^n$ ,  $E_K^n$ ,  $Z_n^\pi$  and  $Z_n^K$ , are the same as those appearing in the two-point fit functions.

We first fit the two-point functions one by one on each ensemble and check the stability of the ground state masses/energies and amplitudes under the choice of fitting range,  $t \in [t_{\text{min}}, t_{\text{max}}]$ , and the number of exponentials included in the fit function [ $N_{\text{exp}}$  in Eq. (2.6)]. We always include the same number  $N$  of regular and oscillating states in those fits, i.e.,  $N_{\text{exp}} = 2N$ . To evaluate the relative quality of the fits we use the  $\chi^2/\text{dof}$  and the  $Q$  value defined in Ref. [50], a quality of fit statistic adapted for fits with Bayesian priors that is similar to the standard  $p$  value [3]. By construction,  $Q \in [0, 1]$  with larger  $Q$  values indicating greater compatibility between the data and fit function given the prior constraints—see Ref. [50] for details and explicit formulas. In particular, we disregard any fit with  $Q < 0.1$ . We also disregard fits with  $\chi^2/\text{dof} \lesssim 0.05$ , since those low  $\chi^2/\text{dof}$  are generally an indication of a bad identification of the ground state and the corresponding fits tend to be unstable with the variation of  $t_{\text{min}}$ , number of exponentials, and/or bootstrap resampling.

We observe that for most of the choices of time range and for all ensembles, fits stabilize when including 2 + 2 or 3 + 3 states. From that parameter-scanning procedure, we select an optimal set of fit parameters for the two-point functions, with a common  $t_{\text{min}}$  for all the functions on the same ensemble. Fixing  $N = 3$ , the chosen  $t_{\text{min}}$  is the smallest value for which the ground state parameters for all relevant two-point functions reach a plateau and, in addition, for which the fit results (central values, errors, and quality) are stable under variations of the number of exponentials and bootstrap resampling. For a fixed  $t_{\text{min}}$ ,  $t_{\text{max}}$  is chosen, in general, as the value for which fit results are insensitive to the addition of late-time data for which statistical errors are larger.

We then use those  $[t_{\text{min}}, t_{\text{max}}]$  ranges to perform a fully correlated combined Bayesian fit including the two- and

three-point functions needed to extract  $f_0^{K\pi}(0)$ :  $\pi$  two-point correlation functions with and without momentum;  $K$  two-point correlation functions without momentum; and  $N_T$  three-point correlation functions with  $q^2 \approx 0$ , where  $N_T$  is the number of source-sink separations  $T$  included in the combined fit. In general, we include three-point functions only at  $N_T = 3$  or 4 different values of  $T$  out of the 5 or 6, for which we have data. However, the  $T$  values included in the combined fits generally cover most of the available range, corresponding to a physical range of  $\approx 0.5$ –1.0 fm. This allows us to resolve excited states while at the same time including data with good ground state contributions. We find that the resulting fits are stable under variations of time range, number of exponentials, and bootstrap resampling. We also find that adding more  $T$  values does not improve the quality (error and stability) of the fits. Table III lists our parameter choices for the combined three-point function fits.

In general, we use three-point data in the combined fits with  $t \in [t_{\text{min}}, T - t_{\text{min}}]$ , where  $t_{\text{min}}$  is the value optimized for the two-point functions. However, on some ensembles, especially the largest ones, we need to either shorten the three-point fit range or thin the three-point data in order to obtain an acceptable fit, as measured by the  $\chi^2/\text{dof}$  and  $Q$  value. A comparison of fit results to data for the  $a \approx 0.09$  fm ensemble with physical quark masses, one of the most relevant in our analysis, is given in Fig. 2. This is a typical case, the comparisons of the fits and data on the other ensembles are similar. The figure plots the rescaled three-point functions, in which the time-dependent contributions of the kaon and pion ground states are removed:

TABLE III. Values of the source-sink separation  $T$  and  $t_{\text{min}}$  in our preferred fits; results for the vector form factor at zero momentum transfer; and one-loop finite-volume corrections,  $\Delta^V f_+(0) = f_+^V(0) - f_+^\infty(0)$ , on each ensemble—see Sec. III A for details of the calculation of  $\Delta^V f_+(0)$ . The errors in  $f_+(0)$  are statistical only. They are generated with a 500-bootstrap distribution.

$\approx a(\text{fm})$	$m_l/m_s^{\text{sea}}$	$m_{\pi,P}L$	$T$	$t_{\text{min}}$	$f_+(0)$	$\Delta^V f_+(0)$
0.15	0.035	3.2	12,16,17	4	0.9744(24)	-0.0007
0.12	0.2	4.5	15,21,22	5	0.9874(24)	0.0002
	0.1	3.2	15,18,21	4	0.9830(31)	-0.0003
	0.1	4.3	15,18,21	4	0.9808(22)	-0.0001
	0.1	5.4	15,18,21	4	0.9809(17)	$-4 \times 10^{-5}$
	0.035	3.9	18,21,22	6	0.9707(18)	-0.0003
0.09	0.2	4.5	27,32,33	3	0.9868(18)	0.0006
	0.1	4.7	23,27,32	6	0.9807(22)	0.0002
	0.035	3.7	23,27,32	8	0.9709(27)	-0.0001
0.06	0.2	4.5	34,41,49,50	8	0.9862(16)	0.0008
	0.035	3.7	31,40,49	10	0.9697(33)	0.0005
0.042	0.2	4.3	40,52,53	12	0.9856(37)	0.0010

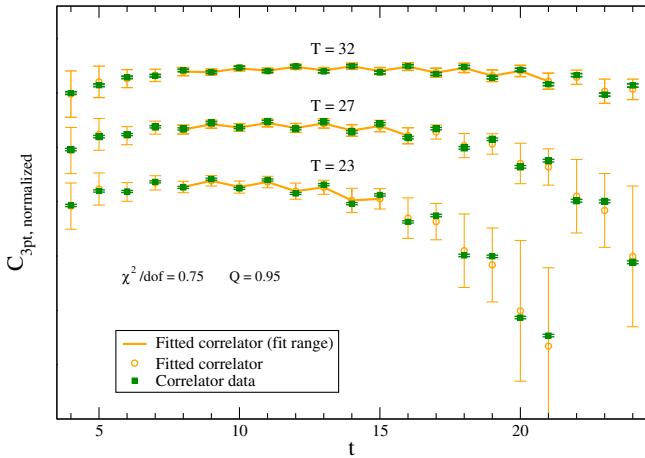


FIG. 2. Comparison of data and fit results for the rescaled three-point functions defined in Eq. (2.8) on the  $a \approx 0.09$  fm ensemble with physical quark masses. Green squares are the data points and orange circles are obtained from the fit posteriors. The fit includes the three three-point functions shown, with  $T = 23, 27, 32$ , in the fit ranges shown by the orange lines. Correlators with  $T = 27, 23$  are given a vertical offset, different for each  $T$ , so results for the three correlators do not lie on top of each other. Errors are statistical only.

$$C_{3pt, \text{rescaled}} = \frac{C_{3pt}^{K \rightarrow \pi}(\vec{p}_\pi, 0; t, T)}{Z_0^\pi Z_0^K (e^{-E_\pi^0 t} + e^{-E_\pi^0 (L_t - t)})(e^{-E_K^0 (T - t)} + e^{-E_K^0 (L_t - T + t)})}. \quad (2.8)$$

In the absence of excited state contributions  $C_{3pt, \text{rescaled}}$  would be time independent. Figure 2 shows the

comparison of the rescaled correlation functions included in the fit (filled green points) with the results from the fit (open orange circles). We see that the  $C_{3pt, \text{rescaled}}$  exhibit plateaus with a mild oscillation that is more pronounced for the smaller  $T$  value, but which can be accounted for almost entirely by the first oscillating state included in the fit. The agreement between data and the fit is excellent, especially for the time ranges included in the fit, marked by the orange lines. For times closer to the source or the sink, the large errors on the orange points indicate a substantial contribution from the excited states included in the fit that the data cannot constrain accurately. Nevertheless, the falloff of the correlators is well described by the fit functions.

From the combined fits, we extract the scalar form factor at zero momentum transfer via

$$f_0^{K\pi}(0) = 2A^{00}(0)\sqrt{E_\pi m_K} \frac{m_s - m_l}{m_K^2 - m_\pi^2}, \quad (2.9)$$

where  $A^{00}(0)$  is the ground state three-point parameter in Eq. (2.7), the meson masses and energies are the values extracted from the combined fits and  $m_s$  and  $m_l$  are the valence strange and light-quark masses simulated.

We check the stability of the combined fit results under the variation of fit ranges, number of states, and number and values of source-sink separations included, and choose a preferred fit for each ensemble so the shift on the central value with those variations is well under the statistical error, and the error is also stable. Examples of these stability studies are shown in Fig. 3. On a few ensembles, the stability tests lead to slight adjustments of our chosen value of  $t_{\min}$ , from the  $t_{\min}$  determined in the two-point only fits

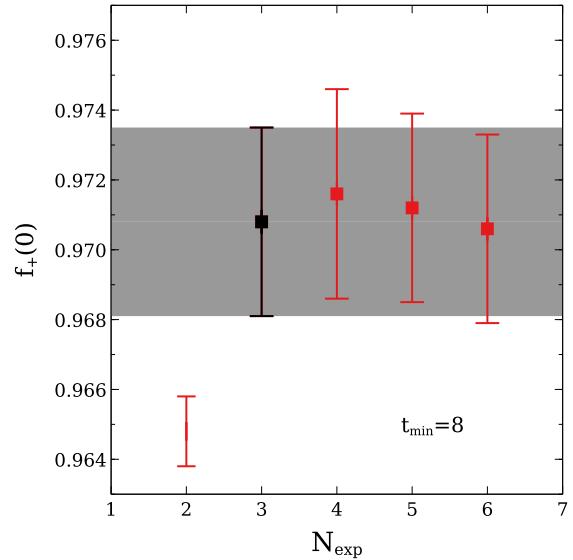
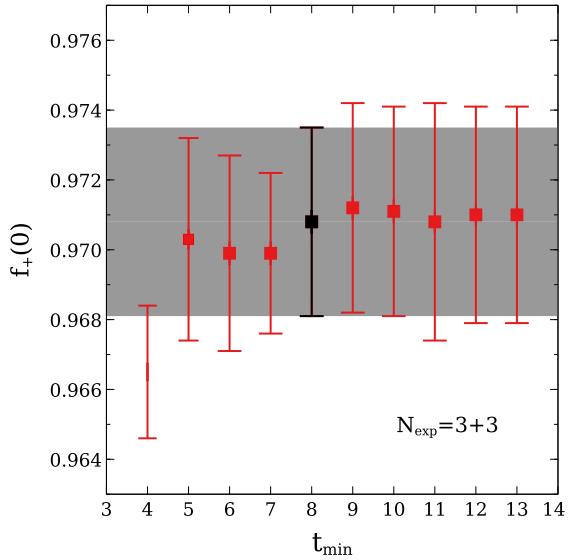


FIG. 3. Variation of the fit result for  $f_+(0)$  with  $t_{\min}$  (left panel) and  $N_{\text{exp}}$  (right panel) for the ensemble with  $a \approx 0.09$  fm and physical light-quark masses. The errors are statistical, generated with a 500-bootstrap distribution. The black point on each figure corresponds to our preferred fit with  $t_{\min} = 8$  and  $N_{\text{exp}} = 3 + 3$ .

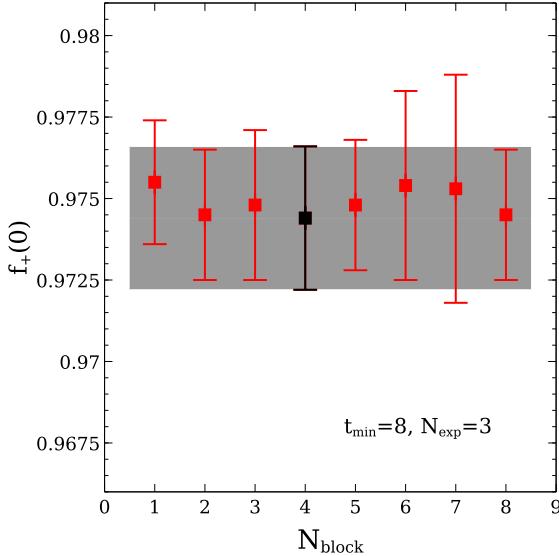


FIG. 4. Variation of  $f_+(0)$  with the block size for the ensemble with  $a \approx 0.15$  fm and physical quark masses. The errors are statistical, generated with a 500-bootstrap distribution. The black point corresponds to our preferred fit with  $N_{\text{block}} = 4$ .

discussed above. Our final choices of  $t_{\min}$  correspond to very similar physical distances, approximately 0.6–0.7 fm, on each ensemble. The number of exponentials is always chosen to be  $3 + 3$ , since also for these combined fits adding more exponentials does not change the fit results and also does not improve fit stability. The parameter values used in the preferred combined fits are listed in Table III.

We study the effect of autocorrelations by blocking the data by increasing numbers of successive configurations and redoing the analysis. We do not see evidence of significant autocorrelations on all ensembles, but for those where we do see significant changes in central value and error, stability is reached with a block size of 4. For ensembles where we observe significant changes in central value and error for the form factor with blocking, those effects stabilize when blocking by 4. An example for the ensemble with  $a \approx 0.15$  fm and physical quark masses is given in Fig. 4. Similar results are obtained for the other ensembles. An alternate estimate of autocorrelation effects can be obtained by calculating the integrated autocorrelation time. We find that the integrated autocorrelation times in the two-point correlation functions included in our analysis are all smaller than 1.4, suggesting that a reasonable block size would be 3 or less. We thus choose to account for autocorrelation effects and block the data in all ensembles by 4. In another test, we construct the covariance matrix from the correlation matrix obtained with the unblocked data together with the variances obtained from the blocked data [46]. Using the same fit setup and parameter choices as before, we find results that are essentially the same as those obtained with our preferred fit method.

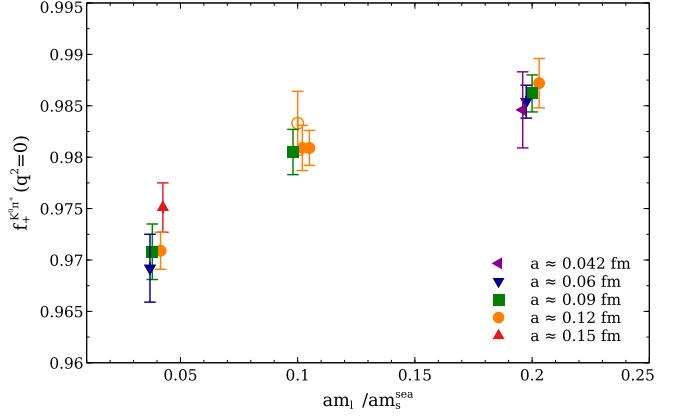


FIG. 5. Form factor  $f_+^{K^0\pi^-}(0)$  vs light-quark mass. The data points are the raw results listed in Table III before applying the corrections described in Sec. III. Errors shown are statistical only, obtained from 500 bootstrap resamples. Different symbols and colors denote different lattice spacings. Data at the same light-quark mass but different lattice spacing are offset horizontally. The open orange circle corresponds to the smallest volume ensemble with  $a \approx 0.12$  fm and  $m_l/m_s = 0.1$ .

In Table III and Fig. 5, we show the raw results for the vector form factor at zero momentum transfer from the combined fits described above. The statistical errors shown in the table and the figure as a function of  $am_l/am_s^{\text{physical}}$  come from 500 bootstrap resamples and range from 0.16% to 0.38%. The fully correlated covariance matrix is recalculated on each bootstrap resample. In the figure, one can see that for a fixed value of the light-quark mass,  $m_l$ , the points with different shapes, which correspond to different values of the lattice spacing, lie on top of each other, with the exception of the data point for the 0.15 fm ensemble with physical quark masses (in the leftmost cluster of points). This is the only ensemble where we observe statistically significant discretization effects.

### III. FORM-FACTOR CORRECTIONS

Before performing the chiral-continuum fit, we correct the form-factor results listed in Table III and shown in Fig. 5 for the leading-order finite-volume effects and the nonequilibrated topological charge in our finest ensemble. These corrections are described in the following two subsections.

#### A. Finite volume

In this work we use NLO ChPT to correct our form-factor results for finite-volume effects, whereas in our previous calculation [23,24] we simply estimated the associated systematic error from a comparison of the lattice data at two different spatial volumes, with other parameters held fixed. The partially twisted boundary conditions used in our calculation introduce several complications in the analysis. In particular, an extra form factor  $h_\mu$  is required to

parametrize the weak-current matrix element in finite volume:

$$\langle \pi^-(p') | V_\mu | K^0(p) \rangle = f_+(p_\mu + p'_\mu) + f_- q_\mu + h_\mu. \quad (3.1)$$

The three form factors depend on the choice of twisting angles, as well as the value of  $q^2$ .

We apply the one-loop formulas in Ref. [40] in the staggered partially twisted partially quenched case to all ensembles included in our calculation for the choices of twist angles in Table III. Because we are calculating the vector form factor at zero momentum transfer via the relation in Eq. (2.2), and the quantity we obtain at finite volume on the lattice is  $\langle \pi | S | K \rangle (m_s - m_d) / (m_K^2 - m_\pi^2)$ , the FV correction to our results is given by

$$\begin{aligned} \Delta^V f_+(0) &\equiv f_+^V(0) - f_+^\infty(0) \\ &= \frac{(m_s - m_d) \Delta^V \langle \pi | S | K \rangle}{(m_K^V)^2 - (m_\pi^V)^2} \\ &\quad - \frac{(m_s - m_d) \langle \pi | S | K \rangle^V (\Delta^V m_K^2 - \Delta^V m_\pi^2)}{[(m_K^V)^2 - (m_\pi^V)^2]^2}, \end{aligned} \quad (3.2)$$

where the meson masses in the denominators are the ones from the simulations, and quantities that are second order in the finite-volume corrections have been neglected. In Eq. (3.2), the FV correction for a given quantity  $X$ ,  $\Delta^V X$ , is defined as  $\Delta^V X \equiv X^V - X^\infty$ . Notice that, since we extract the meson masses from correlation functions where all the propagators have zero momentum, the FV corrections to the meson masses should be calculated from the formulas in Ref. [40] with zero twisting angles.

The resulting FV corrections are listed in the last column of Table III. We find that they are  $\leq 0.1\%$  on all ensembles. Some of the values for  $\Delta^V f_+(0)$  are particularly small due to the cancellation between the two contributions in Eq. (3.2). We subtract the  $\Delta^V f_+(0)$  from the finite-volume  $f_+^{K\pi}(0)$  (listed in the next-to-last column in Table III) before performing the chiral-continuum fit discussed in Sec. IV.

## B. Nonequilibrated topological charge

The HISQ  $N_f = 2 + 1 + 1$  MILC simulations with smallest lattice spacings have reached a regime where the distribution of the topological charge  $Q$  is not properly sampled [41,51], which affects the physical observables calculated on those ensembles. The issue is relevant here for the ensemble with the finest lattice spacing,  $a \approx 0.042$  fm. On the other hand, the topological charge is reasonably well equilibrated on the other ensembles, which have  $a \gtrsim 0.06$  fm.

In order to correct for this systematic effect, one can use ChPT to study the  $Q$ -dependence of a given observable [41,52,53]. The recent ChPT study in Ref. [41] has already been applied to the calculation of heavy-light

meson decay constants and masses in Refs. [46,54]. Here, we extend the analysis of Ref. [41] to  $f_+^{K\pi}(0)$ .

The three-point correlation functions relevant for this study, as well as any meson mass calculated in a finite volume  $V$  and at fixed  $Q$ , satisfy [52,53]

$$B|_{Q,V} = B + \frac{1}{2\chi_T V} B'' \left( 1 - \frac{Q^2}{\chi_T V} \right) + \mathcal{O} \left( \frac{1}{(\chi_T V)^2} \right), \quad (3.3)$$

where  $B$  on the right-hand side is the infinite-volume value of the quantity of interest averaged over  $Q$ ;  $B''$  is its second derivative with respect to the vacuum angle  $\theta$ , evaluated at  $\theta = 0$ ; and  $\chi_T = \lim_{V \rightarrow \infty} \langle Q^2 \rangle / V$  is the infinite-volume topological susceptibility. Knowing the dependence on  $Q$  or, equivalently, on  $\theta$ , one can calculate the appropriate correction to  $B$  to account for the difference between the correct  $\langle Q^2 \rangle$  and the simulation  $\langle Q^2 \rangle_{\text{sample}}$ .

With Eq. (3.3), we follow Ref. [41] to calculate the correction as

$$\begin{aligned} \Delta_Q f_+^{K\pi}(0) &\equiv f_+^{K\pi}(0)_{\text{sample}} - f_+^{K\pi}(0)_{\text{equil}} \\ &= \frac{1}{2\chi_T V} (f_+^{K\pi}(0))'' \left( 1 - \frac{\langle Q^2 \rangle_{\text{sample}}}{\chi_T V} \right), \end{aligned} \quad (3.4)$$

where  $f_+^{K\pi}(0)_{\text{sample}}$  is the simulation value.

Although we extract  $f_+^{K\pi}(0)$  from the scalar-density matrix element in Eq. (2.2) it is simpler to first calculate the  $\theta$  dependence of the vector-current matrix element directly. In ChPT, the vector current with the relevant flavor is

$$V^\mu = \frac{f^2}{4} (\partial^\mu \Sigma \Sigma^\dagger - \Sigma^\dagger \partial^\mu \Sigma)_{13}, \quad (3.5)$$

where  $\Sigma$  is the SU(3) chiral matrix. In the presence of  $\theta$ , and for the  $m_u = m_d = m_l$  and full QCD case (the case relevant for this work), the  $\mathcal{O}(p^2)$  ChPT Lagrangian is

$$\mathcal{L}_\chi = \frac{f^2}{8} \text{tr}(\partial_\mu \Sigma \partial_\mu \Sigma^\dagger) - \frac{\mu f^2}{4} \text{tr}(\mathcal{M}_A^* \Sigma + \mathcal{M}_A \Sigma^\dagger), \quad (3.6)$$

where  $f$  is the chiral-limit value of the meson decay constant, and  $\mu$  the low-energy constant that relates meson and quark masses at leading order (LO)—see Eq. (4.3). Here  $\mathcal{M}_A \equiv e^{i\theta/3} \mathcal{M}$ , with  $\mathcal{M}$  the usual quark mass matrix in the absence of  $\theta$ .

When  $\theta \neq 0$ ,  $\Sigma$  gets the vacuum expectation value

$$\langle \Sigma \rangle = \begin{pmatrix} e^{i\alpha} & 0 & 0 \\ 0 & e^{i\alpha} & 0 \\ 0 & 0 & e^{-2i\alpha} \end{pmatrix}. \quad (3.7)$$

The parameter  $\alpha$  encodes the dependence on  $\theta$ , with  $\alpha(\theta = 0) = 0$ . The relation between  $\alpha$  and  $\theta$  is obtained

by minimizing the potential energy term in the Lagrangian, which gives the condition

$$m_l \sin\left(\alpha - \frac{\theta}{3}\right) + m_s \sin\left(2\alpha + \frac{\theta}{3}\right) = 0. \quad (3.8)$$

For the expansion of the relevant observables, one needs  $\alpha'$ , the first derivative of  $\alpha$  with respect to  $\theta$  evaluated at  $\theta = 0$ . Equation (3.8) implies [41]

$$\alpha' = \frac{m_l - m_s}{3(m_l + 2m_s)}. \quad (3.9)$$

One may expand  $\Sigma$  around its vacuum expectation value via

$$\Sigma = \sqrt{\langle \Sigma \rangle} e^{2i\Phi/f} \sqrt{\langle \Sigma \rangle}, \quad (3.10)$$

with  $\Phi$  the  $3 \times 3$  matrix of meson fields. With this result substituted into Eq. (3.5), at tree level there are two possible diagrams, shown in Fig. 6, that may contribute to the matrix element  $\langle \pi | V^\mu | K \rangle$ . The strong three-point vertex in the right-hand diagram is forbidden by parity when  $\theta = 0$ , but here comes from the mass term in Eq. (3.6), which violates parity symmetry unless one also takes  $\theta \rightarrow -\theta$  (which is called “extended parity”). The weak vertex in the right-hand diagram generates a factor of  $q^\mu$ , implying that diagram contributes only to the form factor  $f_-$ . From the left-hand diagram, one finds

$$f_+^{K\pi}(0) = \cos\left(\frac{3}{2}\alpha\right). \quad (3.11)$$

Finally, from Eq. (3.9), the result needed to adjust the form factor via Eq. (3.4) is

$$f_+^{K\pi}(0)'' = -\frac{1}{4} \frac{(m_l - m_s)^2}{(m_l + 2m_s)^2}. \quad (3.12)$$

Because we actually use Eq. (2.2) to calculate  $f_+^{K\pi}(0)$ , it is important to check that we can reproduce Eq. (3.12) by

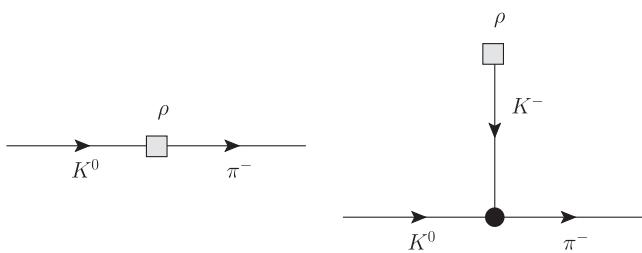


FIG. 6. Diagrams contributing to  $\langle \pi | \rho | K \rangle$  at tree level, where  $\rho$  is either the vector current  $V^\mu$  or the scalar density  $\tilde{S}$ . The squares are weak vertices with the insertion of the current or density and the black dot is a strong vertex.

calculating the matrix element of the scalar density that appears in the  $\theta \neq 0$  Ward identity at  $q^2 = 0$ ,

$$f_+^{K\pi}(0) = \frac{1}{m_K^2 - m_\pi^2} \langle \pi | \tilde{S} | K \rangle_{q^2=0}, \quad (3.13)$$

with  $\tilde{S} = \frac{1}{2} \bar{\psi} [\lambda^f, M] \psi$ , and  $\lambda_f \in SU(N_f)$  the appropriate flavor matrix to select the  $\bar{s}u$  current. Note that it is no longer convenient to take out a factor of  $m_s - m_l$  from  $\tilde{S}$ , as we do for  $S$  in Eq. (2.2), because the quark masses now carry factors of  $\exp(\pm i\theta/3)$ . In ChPT,

$$\tilde{S} = -\frac{f^2}{4} \mu (\Sigma \mathcal{M}_A^* + \Sigma^\dagger \mathcal{M}_A - \mathcal{M}_A^* \Sigma - \mathcal{M}_A \Sigma^\dagger)_{13}. \quad (3.14)$$

Evaluating the diagrams in Fig. 6, we find

$$\begin{aligned} \langle \pi | \tilde{S} | K \rangle_{q^2=0} &= \mu(m_s - m_l) \cos\left(\frac{\alpha}{2} + \frac{\theta}{3}\right) \\ &+ \frac{2}{3} \frac{m_s - m_l}{m_K^2} \mu^2 \sin\left(\frac{\alpha}{2} + \frac{\theta}{3}\right) \\ &\times \left[ m_s \sin\left(2\alpha + \frac{\theta}{3}\right) - 2m_l \sin\left(\alpha - \frac{\theta}{3}\right) \right], \end{aligned} \quad (3.15)$$

where the contributions in the first and second lines come from the propagator and vertex diagrams in Fig. 6, respectively.

The Ward identity in Eq. (3.13) is then satisfied trivially at LO for  $\theta = 0$ . For  $\theta \neq 0$ , we may calculate  $f_+^{K\pi}(0)''$  from Eq. (3.13) using Eq. (3.9), the fact that  $\alpha'' = 0$  (which follows from extended parity), and the second derivatives

$$\begin{aligned} m_\pi'' &= -m_\pi(0) \frac{m_s^2}{2(m_l + 2m_s)^2}, \\ m_K'' &= -m_K(0) \frac{m_l m_s}{2(m_l + 2m_s)^2} \end{aligned} \quad (3.16)$$

from [41]. After some algebra, we find that the result agrees with Eq. (3.12), as expected. We have also checked analytically that the Ward identity holds for arbitrary  $q$  and  $\theta$ .

Following the procedure in Ref. [41] for the partially quenched case, we may generalize Eq. (3.12) to

$$f_+^{K\pi}(0)'' = -\frac{1}{4} \frac{m_l^2 m_s^2}{(m_l + 2m_s)^2} \frac{(m_x - m_y)^2}{m_x^2 m_y^2}, \quad (3.17)$$

where  $x$  and  $y$  are the active valence quarks [the valence up and strange for  $f_+^{K^0\pi^-}(0)$ ], and  $m_l$  and  $m_s$  are the light and strange sea quark masses. In deriving Eq. (3.17), we have set the spectator quark mass (the  $d$  quark mass for  $f_+^{K^0\pi^-}$ ) equal to the light sea mass  $m_l$ ; in other words, the spectator quark is unitary, not partially quenched. This has allowed

us to avoid analyzing the case of three partially quenched quarks, which was not treated in Ref. [41]. Since the mass of the spectator quark does not affect  $f_+^{K^0\pi^-}(0)$  to LO, we believe Eq. (3.17) will remain valid even when the spectator quark is partially quenched. As expected, Eq. (3.17) reduces to Eq. (3.12) when  $m_x = m_l$  and  $m_y = m_s$ .

As discussed at the beginning of this section, the correction is only needed on the finest ensemble included in this analysis, with  $a \approx 0.042$  fm. We calculate the correction of Eq. (3.4) using Eq. (3.12), the value of the average of the topological charge measured on that ensemble,  $\langle Q^2 \rangle_{\text{sample}} = 27.59$  [41], and the correct  $\langle Q^2 \rangle$  as estimated by the LO ChPT expression for the topological susceptibility [55,56]

$$\chi_T = \frac{f_\pi^2}{4} \left( \frac{1}{2m_{ll,I}^{-2} + m_{ss,I}^{-2}} \right), \quad (3.18)$$

where the singlet meson sea masses are defined in Eq. (4.3), below. The resulting correction  $\Delta_Q f_+^{K\pi}(0) = 0.00018$  is subtracted from the  $f_+^{K\pi}(0)_{\text{sample}}$  value listed in the last row of Table III before performing the chiral-continuum fit.

#### IV. CHIRAL-CONTINUUM INTERPOLATION/EXTRAPOLATION

We follow a methodology very similar to that in our previous analyses [23,24] in order to combine our simulation data into physical results in the continuum limit and with the correct quark/meson masses. Here we summarize the main ingredients and then discuss in more detail the new features added in order to accurately account for finite-volume and isospin-breaking corrections. Accounting for these effects turns out to be essential, given the improvements in the simulation data.

Our methodology is developed in the framework of chiral perturbation theory, which allows us to incorporate effects due to mass dependence, discretization, finite volume, and isospin breaking in a systematic way. In particular, in the isospin limit, we can write  $f_+^{K\pi}(0)$  as a chiral expansion

$$f_+^{K\pi}(0) = 1 + f_2 + f_4 + f_6 + \dots, \quad (4.1)$$

where the functions  $f_i$  are chiral corrections of  $\mathcal{O}(p^i)$ . The Ademollo-Gatto (AG) theorem [57] ensures that the vector form factor goes to 1 in the limit  $m_s \rightarrow m_u$ , and that corrections to this limit are second order. That means that the functions  $f_i$  are proportional to  $(m_s - m_u)^2$  or, equivalently,  $(m_K^2 - m_\pi^2)^2$ . The theorem thus implies that, in the continuum, the  $\mathcal{O}(p^2)$  (one-loop) contribution,  $f_2$ , is completely fixed in terms of experimental quantities: the decay constant  $f_\pi$  and meson masses.

The specific fit function we employ for the extrapolation to the continuum and interpolation to the physical quark

masses is the same as in Ref. [23]. It consists of a NLO partially quenched staggered ChPT (PQSChPT) expression [58]  $f_2^{\text{PQSChPT}}(a)$ , plus NNLO continuum ChPT terms [59]  $f_4^{\text{cont}}$ , plus extra analytic terms to parametrize higher-order discretization and chiral effects. Schematically, it can be written

$$f_+^{K\pi}(0) = 1 + f_2^{\text{PQSChPT}}(a) + f_4^{\text{cont}} + g_{1,a} + r_1^4(m_\pi^2 - m_K^2)^2[\tilde{C}_4 + g_{2,a} + h_{m_\pi}], \quad (4.2)$$

where the functions  $g_{1,a}$  and  $g_{2,a}$  account for higher-order discretization effects, and the function  $h_{m_\pi}$  includes analytical terms that parametrize higher-order chiral effects. We have taken the pure counterterm contribution at two loops out of  $f_4^{\text{cont}}$  and written it separately. This contribution corresponds to the term proportional to  $\tilde{C}_4$ , which is given by the combination of low-energy constants (LECs)  $C_{12} + C_{34} - L_5^2$ . The  $\mathcal{O}(p^4)$  LEC  $L_5$  can be extracted from global fits or from lattice-QCD calculations of light-light quantities, but the  $\mathcal{O}(p^6)$  LECs  $C_{12}$  and  $C_{34}$  [60,61] are not known. (Only model-based estimates and imprecise global fit values exist.) We therefore take  $\tilde{C}_4$  as a constrained fit parameter. All dimensionful quantities entering in the fit function in Eq. (4.2) are converted into  $r_1$  units by using the values of  $r_1/a$  in Table IV.

Since our simulations are performed in the isospin limit,  $m_u = m_d$ ,  $f_2$  and  $f_4$  are evaluated for degenerate up and down quarks. The explicit NLO PQSChPT function  $f_2^{\text{PQSChPT}}(a)$  can be found in Ref. [58]. It incorporates the dominant discretization effects coming from the taste-symmetry breaking of staggered fermions. The function  $f_2^{\text{PQSChPT}}(a)$  depends on the HISQ taste splittings  $\Delta_\Xi$  through the sea meson masses

$$m_{ij,\Xi}^2 = \mu(m_i + m_j) + a^2 \Delta_\Xi, \quad (4.3)$$

with  $m_i, m_j$  sea quark masses, the slope  $\mu$  to be determined by fits of the ChPT expressions to experimentally measured meson masses, and  $\Xi$  labeling the meson taste. Values of  $\Delta_\Xi$  for each ensemble are given in Table IV. The function  $f_2^{\text{PQSChPT}}(a)$  also depends on the taste-violating hairpin parameters,  $\delta'_V$  and  $\delta'_A$ , which come from ChPT disconnected diagrams. We fix the taste splittings in the fit function to their values in Table IV since they are precisely enough known that the corresponding errors do not affect our results significantly. The values are from Ref. [45], as well as unpublished updates with better statistics and the inclusion of new ensembles not previously analyzed. The uncertainty in the hairpin parameters is, however, quite large. We therefore treat them as constrained fit parameters with central values and widths equal to those in Table V, determined from fits to light-light meson quantities [66]. Their uncertainty is thus propagated to the final fit errors.

TABLE IV. Inputs for the parameters taken as fixed in the fit function. The  $r_1/a$  values are obtained from a mass-independent scale setting [45,62]. The absolute scale  $r_1$  is from Ref. [63]. The value of the decay constant  $f_\pi$  is taken from Ref. [3]; its error, though shown, is negligible in our calculation. Taste splittings  $r_1^2 a^2 \Delta_\Xi$  are taken from Ref. [45] and more recent updates; slopes  $a\mu$  come from the analysis presented in Ref. [64], although they were not published there. We do not consider errors either on the taste splittings or on the slopes because they also have a negligible effect on the final results. Notice that taste splittings for the  $a \approx 0.042$  fm ensemble are not measured but obtained from the 0.06 fm results, by applying the expected scaling factor  $a_s^2 a^2$ . The LECs  $L_7$  and  $L_8$ , both central values and errors, are taken from fit BE14 in Ref. [65].

$\approx a$ (fm)	0.15	0.12	0.09	0.06	0.042	Continuum
$r_1$						$0.3117 \pm 0.0022$ fm
$f_\pi$						$130.50 \pm 0.13$ MeV
$\Lambda_\chi r_1 = M_\rho r_1$						1.2163
$a\mu$	2.0565	1.6994	1.2820	0.8873	0.6986	
$r_1/a$	2.090(6)	2.608(4)	3.588(7)	5.442(10)	7.143(24)	
$r_1^2 a^2 \Delta_P$	0	0	0	0	0	
$r_1^2 a^2 \Delta_V$	0.301197	0.167563	0.052723	0.009542	0.004794	
$r_1^2 a^2 \Delta_T$	0.204127	0.103326	0.034894	0.006974	0.003504	
$r_1^2 a^2 \Delta_A$	0.106046	0.053983	0.018187	0.003588	0.001803	
$r_1^2 a^2 \Delta_I$	0.399862	0.209269	0.066393	0.012493	0.006276	
$L_7^r(\Lambda_\chi) \times 10^3$						-0.34 ± 0.09
$L_8^r(\Lambda_\chi) \times 10^3$						0.47 ± 0.10

TABLE V. Priors for the fit parameters entering in Eq. (4.2), as well as the posterior values obtained for those parameters in our preferred fit. The dimensionless  $\chi$ PT parameter  $s$  is given by the quantity  $1/(8\pi^2(r_1 f_\pi)^2) \approx 0.3$ . The priors listed for the hairpin parameters are for the  $a \approx 0.12$  fm ensembles, and those for the other lattice spacings are obtained by rescaling these numbers, assuming that the hairpin parameters scale like the average of the  $\Delta_\Xi$ . These values are obtained from fits to light-light quantities using two-loop PQChPT [66]. The uncertainty includes statistical and systematic errors. The prior central values for the NLO LECs are from fit BE14 in Ref. [65] with  $\Lambda_\chi = 0.77$  GeV, while the prior widths are twice the errors in Ref. [65]. We fix the LECs  $L_7$  and  $L_8$  and give their values in Table IV, as explained in the text. The entries “-0.000” denote small negative numbers that round to zero.

Fit parameters	Gaussian priors (central value ± width)	ChPT fit posterior
$r_1^2 a^2 \delta'_V$	$0.050 \pm 0.024$	$0.050 \pm 0.024$
$r_1^2 a^2 \delta'_A$	$-0.0946 \pm 0.0094$	$-0.0958 \pm 0.0093$
$K_1$	$0 \pm 0.01$	$0.001 \pm 0.010$
$K_2$	$0 \pm 0.03$	$0.001 \pm 0.030$
$K'_2$	$0 \pm 0.81$	$0.083 \pm 0.063$
$K_3$	$0 \pm 0.015$	$-0.000 \pm 0.015$
$\tilde{C}_4$	$0 \pm s^2$	$-0.052 \pm 0.006$
$\tilde{C}_6$	$0 \pm s^3$	$0.006 \pm 0.022$
$\tilde{C}_8$	$0 \pm s^4$	$-0.000 \pm 0.008$
$L_1^r(\Lambda_\chi) \times 10^3$	$0.53 \pm 0.12$	$0.55 \pm 0.12$
$L_2^r(\Lambda_\chi) \times 10^3$	$0.81 \pm 0.08$	$0.81 \pm 0.08$
$L_3^r(\Lambda_\chi) \times 10^3$	$-3.07 \pm 0.40$	$-3.03 \pm 0.40$
$[2L_6^r - L_4^r(\Lambda_\chi)] \times 10^3$	$-0.02 \pm 0.10$	$-0.01 \pm 0.11$
$L_5^r(\Lambda_\chi) \times 10^3$	$1.01 \pm 0.12$	$1.00 \pm 0.12$
$L_6^r(\Lambda_\chi) \times 10^3$	$0.14 \pm 0.10$	$0.13 \pm 0.09$

For some of the meson masses that appear in  $f_2$  there are no experimental measurements or lattice results, as for example, for  $m_{ss}^{\text{valence}}$  or for the sea-valence meson masses involving strange quarks. Because we use values given by NLO ChPT for these masses, our  $f_2$  function has some dependence on the corresponding  $O(p^4)$  LECs  $L_i$ . This is the best approximation we have, and we find that different implementations of higher-order corrections result in changes to the central values that are significantly smaller than the statistical errors.

The continuum NNLO ChPT function  $f_4^{\text{cont}}$  also depends on the  $O(p^4)$  LECs. We take most of them as constrained fit parameters with prior central values equal to the posteriors obtained in the  $O(p^6)$  global fit BE14 in Ref. [65]. We take as an input parameter the combination  $2L_6 - L_4$  instead of  $L_4$  because the fit is more sensitive to that combination and because  $L_4$  is fixed in fit BE14. The prior widths are set to twice the errors in Ref. [65]. The chiral scale, at which the LECs and chiral logarithms in the ChPT expression of Eq. (4.2) are evaluated, is set equal to the mass of the  $\rho$  meson, i.e.,  $\Lambda_\chi = M_\rho$ . The  $O(p^4)$  LECs from Ref. [65], used as priors here, agree within errors with (but are more precise than) the only realistic lattice calculations available at the moment: the  $N_f = 2 + 1$  MILC [13] and the  $N_f = 2 + 1 + 1$  HPQCD [12] calculations. The prior central values and widths used in our chiral-continuum fit to Eq. (4.2) are listed in Table V.

The  $O(p^4)$  LECs  $L_7$  and  $L_8$  appear only in the isospin corrections and in the NLO expressions for some of the meson masses in  $f_2$  and  $f_4$ . Their effect on  $f_2$  and  $f_4$  in the isospin limit is, however, negligibly small, and their main impact is via the isospin-breaking corrections, which are

added *after* performing the chiral-continuum fit (see Sec. V F). We choose then to take  $L_7$  and  $L_8$  as fixed parameters in the chiral-continuum fit and include their uncertainties in the total error as described in Sec. V B.

Once the  $O(\alpha_s^2 a^2)$  taste-violating discretization errors for staggered fermions are removed through the explicit dependence on  $a$  of  $f_2^{\text{PQSChPT}}(a)$ , the dominant discretization errors at  $O(p^2)$  in ChPT are  $O(\alpha_s a^2)$  and  $O(a^4)$ . Since we are forced to use continuum ChPT at  $O(p^4)$ , the discretization errors there are  $O(\alpha_s a^2)$  and  $O(\alpha_s^2 a^2)$ . We take these errors into account through the functions  $g_{1,a}$  and  $g_{2,a}$  in Eq. (4.2):

$$g_{1,a} = K_1 \sqrt{r_1^2 a^2 \bar{\Delta} \left( \frac{a}{r_1} \right)^2} + K_3 \left( \frac{a}{r_1} \right)^4, \quad (4.4a)$$

$$g_{2,a} = K_2 \sqrt{r_1^2 a^2 \bar{\Delta} \left( \frac{a}{r_1} \right)^2} + K'_2 r_1^2 a^2 \bar{\Delta}, \quad (4.4b)$$

where the  $K_i$  are fit parameters,  $\bar{\Delta} = \frac{1}{16}(\Delta_P + 4\Delta_A + 6\Delta_T + 4\Delta_V + \Delta_I)$  is the average taste splitting, and  $r_1^2 a^2 \bar{\Delta}$  is a proxy for  $\alpha_s^2 a^2$ . Table V lists the priors employed for the  $K_i$ 's. The terms proportional to  $K_2$  and  $K'_2$  are generic terms parametrizing discretization effects of  $O(\alpha_s a^2)$  and  $O(\alpha_s^2 a^2)$ , respectively, obeying the AG theorem. We include the terms proportional to  $K_1$  and  $K_3$  to account for  $O(\alpha_s a^2)$  and  $a^4$  violations of the AG theorem at finite lattice spacing arising from symmetry-breaking discretization effects in the form-factor decomposition, Eq. (1.4), and in the continuum dispersion relation. We find that adding an  $O(a^4)$  term instead of the one proportional to  $K'_2$  yields fit results that are nearly identical.

As in Refs. [23,24], we also add generic analytical terms corresponding to higher orders in the chiral expansion until the error of the chiral-continuum fit saturates, i.e., until the central value, the error and the  $\chi^2/\text{dof}$  (and  $Q$ ) value do not change appreciably. That happens at N<sup>4</sup>LO [ $O(p^8)$ ]—see Sec. V. The function  $h_{m_\pi}$  in Eq. (4.2), which collects these effects, therefore takes the form

$$h_{m_\pi} = \tilde{C}_6 r_1^2 m_\pi^2 + \tilde{C}_8 r_1^4 m_\pi^4. \quad (4.5)$$

The terms proportional to  $\tilde{C}_6$  and  $\tilde{C}_8$  are  $O(p^6)$  and  $O(p^8)$ , respectively. The  $\tilde{C}_i$  are constrained fit parameters; the priors for them can be found in Table V. Further discussion of the fit function, priors used in the Bayesian approach, and tests performed can be found in Refs. [23,24].

### A. Fit results

We fit our finite-volume corrected form-factor data to the functional form in Eq. (4.2) with the functions  $g_{i,a}$  and  $h_{m_\pi}$  given in Eqs. (4.4) and (4.5), respectively. All the results

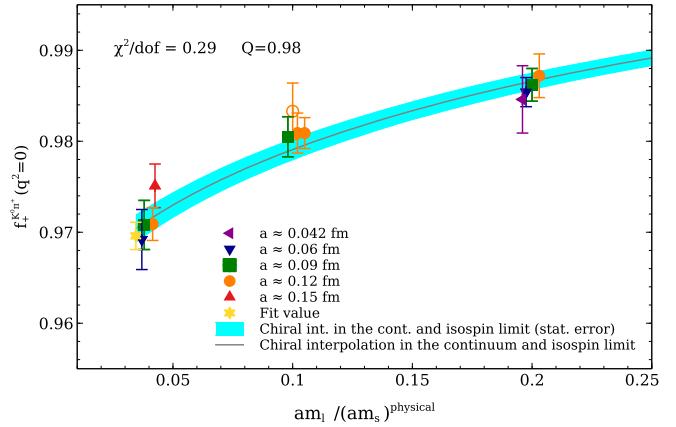


FIG. 7. Form factor  $f^{K^0\pi^-}(0)$  vs light-quark mass. The data points correspond to the results in Table III and are corrected for the one-loop finite-volume effects also listed in that table. Different symbols and colors denote different lattice spacings. The data point at the  $a \approx 0.042$  fm ensemble includes the correction given in Sec. III B. The error bars on the data points are statistical only, obtained from 500 bootstrap resamples. Data points at the same light-quark mass but different lattice spacings are offset horizontally. The gray continuous line shows the continuum extrapolation in the isospin limit as a function of the light-quark mass, and the yellow star is the continuum result interpolated to the physical light-quark masses. The cyan error band, as well as the error bar on the physical point, is the statistical chiral-continuum fit error (obtained using 500 bootstrap resamples), which includes discretization and higher-order chiral errors, as well as the uncertainty from some of the input parameters, as discussed in the text. The continuum extrapolation line is obtained by setting the valence and sea light-quark masses equal, setting  $m_s$  to its physical value, turning off all discretization effects, and considering  $m_u = m_d$ , i.e., without isospin-breaking effects. On the other hand, the yellow star is the interpolation to the physical masses and includes strong isospin-breaking effects at NNLO.

listed in Table III are included in our central-value fit, except for the ensemble with  $a \approx 0.12$  fm and  $m_{\pi,P} L = 3.2$ , which we use only to check finite-volume effects. We then extrapolate to the continuum limit and interpolate to the pure-QCD meson masses, i.e., with electromagnetic effects removed, using the parameters determined from the fit described above together with the continuum isospin-breaking NNLO ChPT expressions in Ref. [67], plus the N<sup>3</sup>LO and the N<sup>4</sup>LO chiral terms in Eq. (4.5), which do not vanish in the continuum limit. We take the pure-QCD masses from Ref. [46]<sup>4</sup>:  $m_{K^0}^{\text{QCD}} = 497.567$  MeV,  $m_{K^+}^{\text{QCD}} = 491.405$  MeV,  $m_{\pi^+}^{\text{QCD}} = 135.142$  MeV,  $m_{\pi^0}^{\text{QCD}} = 134.977$  MeV. For the  $K^0 \rightarrow \pi^- \ell \nu$  case we find

<sup>4</sup>The  $\pi^0$  QCD mass is just the experimental one, and the  $\pi^+$  QCD mass includes the estimate of the small isospin-breaking correction, which comes from Ref. [68].

$$f_+^{K^0\pi^-}(0) = 0.9696 \pm 0.0015, \quad (4.6)$$

where the error is from the fit only, and does not yet include all systematic effects. The statistical error is estimated by fitting to a set of 500 bootstrap samples for each ensemble. On each of those fits, we randomly change the central values of all the priors sampling over Gaussian distributions, keeping the same widths as in Table V. The plot in Fig. 7 shows, as a function of the light-quark mass, the central interpolation curve as well as its error band in the continuum and with the strange-quark mass adjusted to its physical value. In order to make the comparison to data clearer, the curve in Fig. 7 does not include any strong isospin-breaking effects, i.e.,  $m_l = m_u = m_d$ . The point at the physical masses (yellow star) in Fig. 7, however, is our central result in Eq. (4.6), which includes strong isospin-breaking effects at NNLO.

The result in Eq. (4.6) includes isospin corrections up to NNLO—see Sec. V F for more details. For  $K^0$  decays, isospin corrections enter only at NLO ( $f_2$ ) and beyond, and are small,  $< 0.15\%$ . It also includes corrections for the leading-order finite-volume effects as described in Sec. III A.

The second column in Table V shows the posteriors for the fit parameters of the chiral-continuum fit that leads to the result in Eq. (4.6). We cannot determine the coefficients  $K_i$  accurately since there is very little  $a^2$  dependence in our results. In fact, if we remove the  $a \approx 0.15$  fm point we could fit our data without including discretization effects at all. Our lattice data also provide little constraint on the

individual values of the  $O(p^4)$  LECs. As seen in Table V, the posterior fit values of the  $L'_i$  are generally the same as the priors.

## V. SYSTEMATIC ERROR ANALYSIS

The error in Eq. (4.6) includes statistical, chiral-extrapolation, and discretization errors, as well as the uncertainties associated with the inputs that are treated as constrained fit parameters:  $O(p^4)$  LECs (except  $L_{7,8}$ ) and taste-violating hairpin parameters. The uncertainties of the data and constrained input parameters are propagated through the fit via 500 bootstrap resamples.

In this section, we further study these sources of uncertainty, perform tests of the stability of our preferred fit strategies, and estimate the other sources of systematic error entering in our calculation of  $f_+^{K\pi}(0)$ : uncertainty in the inputs, scale error, partial-quenching effects, higher-order finite-volume effects, isospin-breaking corrections, and the effects of nonequilibrated topological charge.

### A. Fit function, discretization error and chiral interpolation

Because we have data at the physical light-quark masses, the chiral fit is an interpolation, and is largely independent of the precise form of the fit function and the values of the ChPT parameters. We have performed a number of tests to check this stability under variations in the fit function and to estimate the effect of higher-order terms in the chiral and Symanzik expansions. Figure 8 shows the tests performed

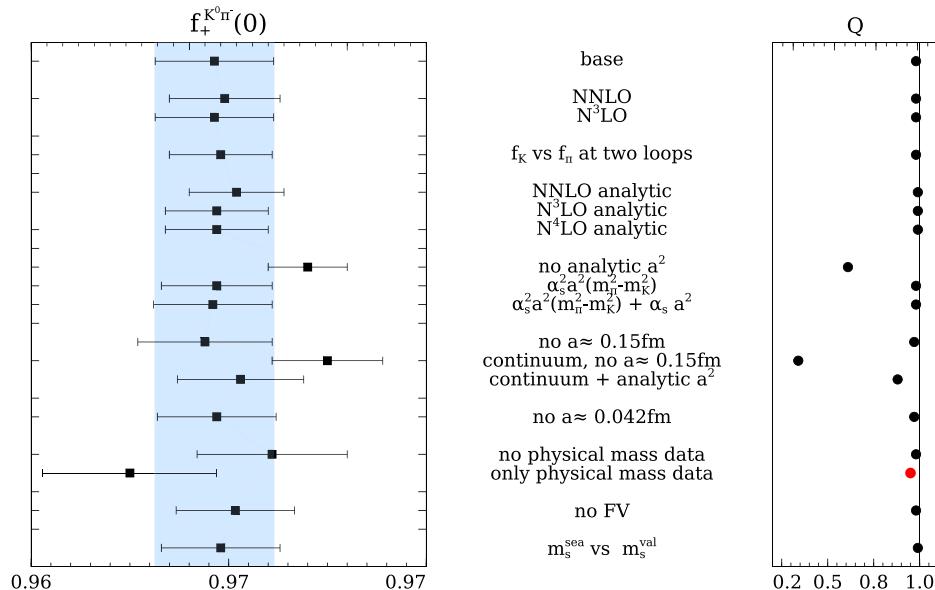


FIG. 8. Stability of the continuum extrapolation and chiral interpolation with respect to the choice of fit function. The blue band corresponds to our preferred fit function (labeled “base”). Notice that the analytical parametrization is applied only at NNLO and beyond. The PQSChPT expression is used at NLO, including isospin corrections. The  $Q$  value for each fit is shown in the right-hand-side plot. The red point on that plot corresponds to a fit with  $\chi^2/\text{dof} < 0.05$ . See the text for the explanation of the different tests performed.

and discussed in this subsection, together with the results from additional fits discussed in the following subsections, which we use to estimate several systematic uncertainties.

First of all, in order to be sure that effects due to higher-order terms in the chiral expansion are properly included in the Bayesian analysis that leads to the fit error shown in Eq. (4.6), we need to check that this error stabilizes as we add higher-order chiral terms. The point labeled NNLO in Fig. 8 includes only terms up to NNLO, i.e., without the  $h_{m_\pi}$  function in Eq. (4.2). Only minimal changes in the central value and errors are produced by the addition of the N<sup>3</sup>LO term  $\tilde{C}_6(m_K^2 - m_\pi^2)^2 m_\pi^2$  from Eqs. (4.5) and (4.2). The difference between this and the base fit is the N<sup>4</sup>LO term  $\tilde{C}_8(m_K^2 - m_\pi^2)^2 m_\pi^4$ , and we see that it has a negligibly small effect.

Different values of the decay constant used in the two-loop (NNLO) term  $f_4^{\text{cont}}$  are equivalent up to omitted higher-order terms, and therefore should have a negligible effect on our analysis. In contrast, the decay constant at one loop has to be set equal to the physical pion decay constant,  $f_\pi$ , in order to be consistent with the particular expression chosen for  $f_4^{\text{cont}}$ . In the base fit, we use  $f_\pi$  as the chiral expansion parameter in  $f_4^{\text{cont}}$ . We check that other possible choices, such as  $f_K = (155.6 \pm 0.4)$  MeV [16] (labeled “ $f_K$  vs  $f_\pi$  at two loops” in Fig. 8) and an estimate of the decay constant in the chiral limit,  $f_0 = (113.5 \pm 8.5)$  MeV [13], shift the central value in Eq. (4.6) by less than 0.06%, well under the statistical error.

As another test of the ChPT fit and errors, we replace the continuum two-loop ChPT expression in Eq. (4.2),  $f_4^{\text{cont}}$ , by an analytic function, and consecutively add N<sup>3</sup>LO and N<sup>4</sup>LO analytic terms—see results labeled “NNLO analyt.”, “N<sup>3</sup>LO analyt.”, and “N<sup>4</sup>LO analyt.”, respectively, in Fig. 8. All the results agree very well with our base fit within statistics, being nearly identical once the N<sup>3</sup>LO analytic term is included.

The three results labeled “no analyt.  $a^2$ ” (which corresponds to a fit without including  $g_{1,a}$  and  $g_{2,a}$  in the fit function), “ $\alpha_s^2 a^2 (m_\pi^2 - m_K^2)$ ”, and “ $\alpha_s^2 a^2 (m_\pi^2 - m_K^2) + \alpha_s a^2$ ”, in Fig. 8 represent a check that the discretization errors are properly included in the fit error of Eq. (4.6). Once we include the term of order  $\alpha_s^2 a^2 (m_K^2 - m_\pi^2)^2$  (proportional to  $K'_2$ ) in Eq. (4.4), which is required to get a fit of similar quality to our base fit (see Fig. 8), the central value and errors barely change with the addition of  $\alpha_s a^2$  corrections [the result labeled “ $\alpha_s^2 a^2 (m_\pi^2 - m_K^2) + \alpha_s a^2$ ”]. Adding the two remaining discretization terms in  $g_{1,a}$  and  $g_{2,a}$ , which returns us to the base fit, makes no noticeable difference. The rapid stabilization of the fit reflects the tiny lattice-spacing dependence of our data.

Of all the data in Fig. 5, only the point at  $a \approx 0.15$  fm shows what appear to be significant discretization effects. Dropping that data point has the effect of increasing the errors (see result labeled “no  $a \approx 0.15$  fm”), since the other

ensembles provide very little constraint of the analytical  $a^2$  fit parameters. In fact, after dropping that point, we can fit our remaining data with a continuum fit function, although we see from Fig. 8 (result labeled “continuum, no  $a \approx 0.15$  fm”) that the result is larger than our central result by about two standard deviations, measured in terms of the fit errors, and the quality of the fit significantly drops. Adding analytical discretization corrections via the functions  $g_{1,a}$  and  $g_{2,a}$  to the continuum fit function allows us to fit all our data, giving a result that is consistent with the base fit and with a similar  $Q$  value (see result labeled “continuum + analyt.  $a^2$ ”), although with a larger error.

In contrast to the noticeable effect of the coarsest ensemble on the total error, the effect of our finest lattice spacing,  $a \approx 0.042$  fm, on the central value and the error is very small since statistics in this ensemble is limited and, in addition, it has  $m_l = 0.2m_s$ , so it is relatively far from the physical point.

As shown in Fig. 8, both the ensembles with physical quark masses and those with unphysical masses are important in fixing the central value and reducing the fit error. The larger error of the fit including only physical-quark-mass ensembles reflects the weaker constraints on the higher-order discretization terms and the lack of constraints on the higher-order chiral terms, which can have an effect on the results from nominally “physical” ensembles due to mistunings of the strange and light-quark masses. On the other hand, the larger error of the fit including only the unphysical-quark-mass ensembles reflects primarily the error of the chiral extrapolation.

Finally, we test the robustness of our Bayesian error estimation strategy similarly to our previous work [23,24], by obtaining separate estimates of each source of error from central value variations observed with simpler fits with and without the corresponding higher-order terms—see Ref. [23] for details. Taking the total error as their quadrature sum, we find that this procedure yields smaller uncertainties than those in Eq. (4.6).

For the reasons discussed above, the statistical fit error shown in Eq. (4.6), which is obtained with our base fit using Eq. (4.2), together with the higher-order chiral and discretization terms in Eqs. (4.4) and (4.5), properly includes the errors from higher-order discretization effects and chiral corrections in addition to the statistical errors. The inclusion of the unphysical light-quark-mass data in our ChPT description gives us a handle on these higher-order effects and allows us to robustly correct for mass mistunings and estimate the error associated with the truncation of the corresponding series.

## B. Inputs for the fixed parameters in the chiral function

The values and errors of the fixed inputs we use in our chiral-continuum fit are listed in Table IV. The HISQ taste splittings are known precisely enough that their errors have

TABLE VI. Error budget for  $f_+^{K^0\pi^-}(0)$  in percent.

Source of uncertainty	Error $f_+^{K^0\pi^-}(0)$ (%)
Statistical + discretization + chiral interpolation	0.154
$L_{7,8}^r$	0.079
Scale $r_1$	0.080
$m_s^{\text{val}} \neq m_s^{\text{sea}}$	0.013
Higher-order finite-volume corrections	0.053
Higher-order isospin corrections	0.015
Isospin-breaking parameter $R$	0.002
Total error	0.199

no impact on the final uncertainty. Similarly, when we change the pion decay constant within its error and repeat the fit, results for the form factor are unchanged at the precision we quote. The uncertainty is small because the dependence on  $f_\pi$  enters through the coefficients and parameters in the ChPT fit function, which, as discussed above, already have little effect on the results. Finally, by varying  $\Lambda_\chi$  in the range  $M_\rho \pm 0.5$  GeV, we have checked that our results are independent of the chiral scale, as they should be. We therefore do not need to add an uncertainty due to the errors in the inputs or the choice of chiral scale to the statistical fit error.

However, the LECs  $L_7$  and  $L_8$  (which we treat as fixed input parameters, unlike the other LECs), do have an impact on the form-factor error, mainly through their effect on the isospin corrections. We estimate this uncertainty by varying their central values by their respective standard deviations, repeating the fit, and recalculating the form factor (including isospin corrections). We take the shift that this variation produces as the uncertainty associated with these LECs, and add it in quadrature to the fit error, as shown in Table VI. The above procedure does not underestimate the error due to these LECs, since if we treat  $L_7$  and  $L_8$  as constrained fit parameters instead, the same as the other  $O(p^4)$  LECs, we obtain a slightly smaller total error.

### C. Lattice scale

We rewrite all the dimensionful quantities entering in the two-loop ChPT fit function in  $r_1$  units, where the  $r_1$  scale is obtained from the static-quark potential [69,70]. The lattice parameters are converted to  $r_1$  units by multiplying by the values of the relative scales  $r_1/a$  in Table IV, while the physical parameters are converted by using  $r_1 = 0.3117(22)$  fm [63].

The form factor  $f_+^{K\pi}(0)$  is a dimensionless quantity, and thus the effect of the error in the lattice scale is small. When we change the scale  $r_1$  by its error, the central value only shifts by  $\pm 0.0008$ . We include this variation as a systematic error in Table VI. The errors in the relative scales  $r_1/a$ , on the other hand, have no significant impact on our results.

### D. Partial-quenching effects in $m_s$ at NNLO

The valence and sea strange-quark masses differ on some of the ensembles as explained in Sec. II, leading to partial-quenching effects—see Table I. These effects can be exactly treated at NLO within the PQSChPT framework, but at NNLO only the full-QCD ChPT expressions are available. We then have the choice of using either the sea or the valence  $m_s$  at NNLO and beyond. In practice, this ambiguity only affects  $f_4^{\text{cont}}$  in Eq. (4.2) since the factor  $(m_\pi^2 - m_K^2)^2$  in that equation comes from the valence sector.

The result in Eq. (4.6) is obtained using the valence strange-quark masses at NNLO. If we use the sea strange-quark masses at NNLO instead,  $f_+^{K^0\pi^-}(0)$  shifts by 0.013%, which we include on the line labeled “ $m_s^{\text{val}} \neq m_s^{\text{sea}}$ ” in the error budget. This systematic effect is small because the sea strange-quark masses are generally well tuned on the HISQ  $N_f = 2 + 1 + 1$  MILC ensembles, and  $m_s^{\text{val}} = m_s^{\text{sea}}$  on the most relevant ensembles in the chiral-continuum interpolation/extrapolation, the ensembles with physical quark masses and  $a \approx 0.09, 0.06$  fm.

### E. Higher-order finite-volume corrections

In our previous calculation [23,24], the uncertainty due to finite-volume effects was one of the two dominant sources of error. (The other was the fit error.) The finite-volume error was estimated to be of the same order as the statistical error from a comparison of the lattice data from two different volumes, with other parameters held fixed. Then, although very small, 0.2%, this error turned out to be a limiting factor for precision. In this work we have increased the statistics on the ensembles analyzed in Refs. [23,24] to check finite size effects. We have also sharpened this direct comparison by generating data on a third, smaller, volume. The three ensembles are those with  $a \approx 0.12$  fm and  $m_l/m_s^{\text{sea}} = 0.1$  in Table I and Fig. 1. Table III gives the values for  $f_+^{K\pi}(0)$  on these three volumes. The results on the two largest volumes are essentially the same, while that on the smallest volume differs from the others by less than the statistical error. From this comparison alone we could conclude that finite-volume effects are smaller than 0.17%, the smallest statistical error on the three ensembles.

To reduce the error further we use NLO staggered partially twisted partially quenched ChPT [40] to correct the form factor prior to the chiral-continuum fit, as described in Sec. III A. The resulting finite-volume corrections are  $\leq 0.1\%$  on all ensembles. If we did not correct our data for finite-volume effects at one loop, the result for  $f_+^{K^0\pi^-}(0)$  would shift by 0.00051. Although we expect NNLO finite-volume corrections to be suppressed by a typical one-loop suppression factor, we conservatively take this shift as the estimate for the higher-order finite-volume effects. This gives a 0.053% error that we include in the error budget in Table VI.

TABLE VII. Form factor  $f_+^{K^0\pi^-}(0)$  as extracted from the most recent lattice calculations (first half of the table), from phenomenological approaches using two-loop ChPT, and from the 1984 calculation by Leutwyler and Roos, which uses one-loop ChPT and a quark model for higher-order terms. For those calculations based on two-loop ChPT, we also indicate the method used in the estimate of the  $O(p^6)$  LECs.

Group	$f_+^{K^0\pi^-}(0)$	Method
This work	0.9696(15)(12)	Staggered fermions ( $N_f = 2 + 1 + 1$ )
ETM [26]	0.9709(45)(9)	Twisted-mass fermions ( $N_f = 2 + 1 + 1$ )
Fermilab Lattice/MILC [23]	0.9704(24)(22)	Staggered fermions ( $N_f = 2 + 1 + 1$ )
JLQCD [30]	0.9636(36)( <sup>+57</sup> <sub>-35</sub> )	Overlap fermions ( $N_f = 2 + 1$ )
RBC/UKQCD [25]	0.9685(34)(14)	Domain-wall fermions ( $N_f = 2 + 1$ )
Bijnens and Ecker [65,72]	0.970(8)	ChPT + NNLO global fit
Kastner and Neufeld [73]	0.986(8)	ChPT + large $N_c$ + dispersive
Cirigliano <i>et al.</i> [74]	0.984(12)	ChPT + large $N_c$
Jamin, Oller, and Pich [75]	0.974(11)	ChPT + dispersive (scalar form factor)
Bijnens and Talavera [59]	0.976(10)	ChPT + Leutwyler and Roos
Leutwyler and Roos [76]	0.961(8)	One-loop ChPT + quark model

## F. Isospin-breaking corrections

Isospin-breaking corrections accounting for the difference between the up- and down-quark masses can be calculated in the ChPT framework and thus written as a chiral expansion starting at NLO for neutral kaons

$$\begin{aligned} \Delta_{\text{isospin}} f_+^{K^0\pi^-}(0) &\equiv f_+^{K^0\pi^-}(0) - f_{+,\text{isospin limit}}^{K\pi}(0) \\ &= \sqrt{3}(\zeta_{S,K^0\pi^-}^{(4)} + \zeta_{S,K^0\pi^-}^{(6)} \dots), \end{aligned} \quad (5.1)$$

where the parameters  $\zeta_{S,K^0\pi^-}^{(i)}$  are  $O((m_u - m_d)p^i)$  isospin corrections. In our result for  $f_+^{K^0\pi^-}(0)$  in Eq. (4.6), we include both NLO [ $O((m_u - m_d)p^4)$ ] and NNLO [ $O((m_u - m_d)p^6)$ ] corrections calculated in Refs. [68] and [67], respectively. These corrections depend on the lowest-order  $\pi^0 - \eta$  mixing angle  $\epsilon^{(2)}$ , or, alternatively, the quantity  $R \equiv (m_s - \hat{m})/(m_d - m_u)$  with  $\hat{m} \equiv (m_u + m_d)/2$ . In order to arrive at the number in Eq. (4.6), we use the expressions in Ref. [67], the QCD meson masses quoted in Sec. IV A, and the values of the LECs obtained from our fits and shown in Table V (for  $L_7^r$  and  $L_8^r$  we take the input values in Table IV). The only combination of  $O(p^6)$  LECs that enters at this order in the isospin-breaking terms for  $K^0 \rightarrow \pi^-\ell\nu$  decays is  $C_{12} + C_{36}$ . This combination, which we obtain from our fitting procedure, is the same one that appears in the isospin limit.

We use a power-counting estimate for the error due to isospin corrections not included in our result, N<sup>3</sup>LO and higher, by taking the calculated NNLO correction and multiplying it by a typical chiral-loop suppression factor. For quantities involving a strange quark, we may estimate this factor to be  $m_K^2/(8\pi^2 f_\pi^2) \approx 0.18$ . The size of the ratio of the isospin limit NNLO and NLO contributions to  $f_+^{K^0\pi^-}(0)$  that we obtain in this work is a bit larger,  $\approx 0.26$ . We conservatively multiply the calculated NNLO isospin-breaking correction,  $-0.00057$ , by the larger number, which yields a 0.015% uncertainty.

Another source of error is the parametric uncertainty in the isospin-breaking quantity  $R$  used to obtain the corrections in Eq. (5.1). We use the value

$$R = 35.59(21)_{\text{stat}} \left( \begin{array}{c} +88 \\ -96 \end{array} \right)_{\text{syst}} [35]_{\text{EM-scheme}}. \quad (5.2)$$

The analysis that yields to this result is the same as in Ref. [46], except that we have included more configurations at the ensembles with  $a \approx 0.06$  fm and  $a \approx 0.042$  fm, and included the  $a \approx 0.15$  fm data in the central fit. The electromagnetic errors are estimated as in Ref. [71].

We estimate the error on the form factor coming from the uncertainty on  $R$  by varying this quantity within its error and redoing the fit. As expected, the impact on the form factor for the neutral mode is nearly negligible, 0.002%. Nevertheless, we include it in our error budget.

## G. Nonequilibrated topological charge

As described in Sec. III B, a correction due to improper sampling of the topological charge is needed only on the  $a \approx 0.042$  fm ensemble with  $m_l = 0.2m_s$ , where we obtain  $\Delta_Q f_+^{K\pi}(0) = 0.00018$ . Not surprisingly, given that (i) this ensemble has little influence on the chiral-continuum interpolation/extrapolation (see Fig. 8 for the effect of removing the ensemble completely), and (ii) the correction is much smaller than the statistical error on the ensemble (see Table III), the effect of the correction on the physical value of  $f_+^{K\pi}(0)$  is negligible. We therefore do not add an uncertainty due to this effect to our error budget.

## VI. RESULTS

Our final result for the vector form factor is

$$f_+^{K^0\pi^-}(0) = 0.9696(15)_{\text{stat}}(12)_{\text{syst}} = 0.9696(19), \quad (6.1)$$

where the first error in the middle expression is the combined statistical, discretization and chiral interpolation

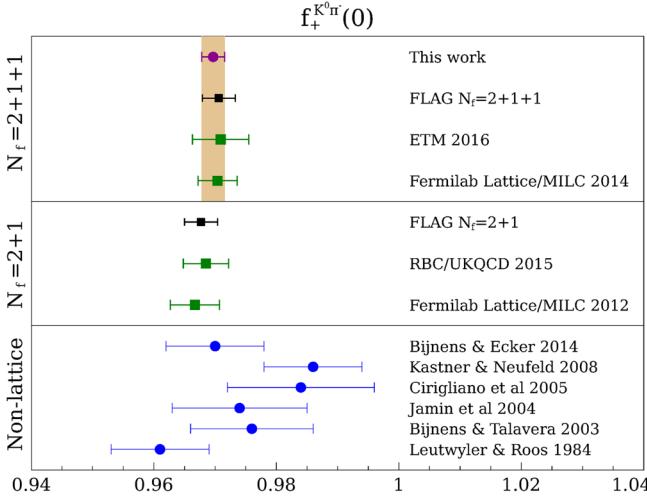


FIG. 9. Comparison of  $f_+^{K^0\pi^-}(0)$  from this analysis with previous lattice results entering in the FLAG averages [6] together with those averages for  $N_f = 2 + 1 + 1$  and  $N_f = 2 + 1$ , as well as nonlattice determinations based on ChPT. The beige band corresponds to our result. The references and numerical results for all determinations are given in Table VII.

error discussed in Sec. IV A, and the second the sum in quadrature of all the systematic errors discussed in Sec. V. Table VI summarizes all sources of error in our calculation. The total uncertainty is the smallest achieved to date.

We compare our result for  $f_+^{K^0\pi^-}(0)$  with the results from the most recent lattice calculations and phenomenological approaches in Table VII, and with the results entering the FLAG average and those from phenomenological approaches in Fig. 9. Our value for  $f_+^{K^0\pi^-}(0)$  agrees within errors with previous  $N_f = 2 + 1$  and  $N_f = 2 + 1 + 1$  lattice calculations. In particular, the value is close to the other  $N_f = 2 + 1 + 1$  results, but with significantly smaller errors. It also agrees with the most recent phenomenological determinations [65,72], which are based on two-loop ChPT with LECs determined by NNLO global fits. The lattice results in Table VII and in Fig. 9 do not include isospin corrections, with the exception of the Fermilab Lattice/MILC result in Ref. [23] (only NLO corrections) and our result here (up to NNLO corrections).

### A. $O(p^6)$ LEC combination $C_{12}^r + C_{34}^r$

The parameter  $\tilde{C}_4$  in the two-loop ChPT fit function that we use to interpolate  $f_+^{K\pi}(0)$  to the physical point—see Eq. (4.2)—is related to the combination of  $O(p^4)$  and  $O(p^6)$  LECs

$$\tilde{C}_4 = -\frac{8}{f_\pi^4} [C_{12} + C_{34} - L_5^2]. \quad (6.2)$$

We can thus use the values of  $\tilde{C}_4$  and  $L_5^r$  from our fit output in Table V to extract the combination of  $O(p^6)$  LECs involved. Taking correlations into account, we find

TABLE VIII. Error budget for the LEC combinations of order  $p^6$ :  $[C_{12}^r + C_{34}^r](M_\rho)$  and  $[C_{12}^r + C_{34}^r - (L_5^r)^2](M_\rho)$ .

Source of uncertainty	$[C_{12}^r + C_{34}^r](M_\rho) \times 10^6$	$[C_{12}^r + C_{34}^r - (L_5^r)^2](M_\rho) \times 10^6$
Stat. + disc. + chiral inter.	0.36	0.23
$L_{7,8}^r$	0.12	0.13
Scale $r_1$	0.13	0.14
$m_s^{\text{val}} \neq m_s^{\text{sea}}$	0.02	0.02
Finite volume	0.09	0.08
Total error	0.41	0.031

$$[C_{12}^r + C_{34}^r](M_\rho) = 3.93(36)_{\text{stat}}(20)_{\text{syst}} \times 10^{-6}. \quad (6.3)$$

The first error in Eq. (6.3) includes statistics, chiral extrapolation and discretization errors, as well as the uncertainty from the LECs (except  $L_7$  and  $L_8$ ) and the taste-violating hairpin parameters, as discussed in Sec. V. The second error is the sum in quadrature of the rest of the systematic uncertainties. The detailed error budget is in Table VIII. We obtain all the errors in the same way as for  $f_+^{K\pi}(0)$ . Isospin corrections do not apply to this quantity since it is defined in the isospin limit. In practice, the values of LECs coming from a fit may be significantly affected by the presence or absence of higher-order chiral terms in the fit function. Therefore, applications of our result in Eq. (6.3) should allow the same type of corrections as in (the continuum limit of) Eq. (4.2). The complete error budget for this quantity can be found in Table VIII.

Our result in Eq. (6.3) agrees with nonlattice determinations in Refs. [74–76]. In those papers, the contribution to  $f_+(0)$  from  $C_{12} + C_{34}$  was calculated using the large  $N_c$  approximation, a coupled-channel dispersion relation analysis, and a quark model, respectively. However, the value for  $C_{12} + C_{34} - L_5^2$  found in Ref. [73], which is based on ChPT, large  $N_c$  estimates of the LECs, and dispersive methods, is  $\sim 3\sigma$  smaller than our value,  $[C_{12}^r + C_{34}^r - (L_5^r)^2](M_\rho) = (2.92 \pm 0.31) \times 10^{-6}$ .

The result in Eq. (6.3) also agrees very well with our previous calculation of this combination of LECs in Ref. [27], on the MILC  $N_f = 2 + 1$  asqtad configurations, although with greatly reduced errors. In fact, all sources of error are reduced due to several factors: the use of the MILC  $N_f = 2 + 1 + 1$  HISQ configurations with smaller discretization errors than the asqtad action, data at smaller lattice spacings, data with physical light-quark masses, better tuning of the strange sea quark masses, and including NLO finite-volume corrections explicitly. The agreement with the JLQCD result in Ref. [30] is borderline, but the JLQCD calculation relies on simulations at a single lattice spacing, although a systematic error is quoted for it, and it does not include data at the physical light-quark masses. Those systematics could affect more strongly the value of the combination of LECs than the form factor itself.

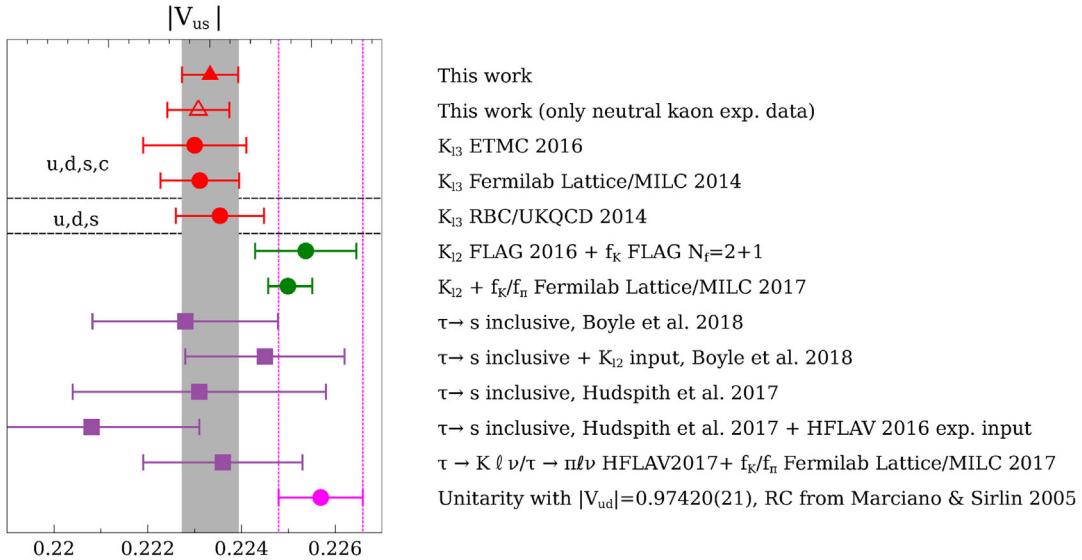


FIG. 10. Summary of recent  $|V_{us}|$  determinations. The semileptonic determinations, labeled  $K_{13}$ , use inputs for  $f_+^{K\pi}(0)$  from the most recent lattice calculations in Refs. [23,25,26], respectively. The leptonic determinations, labeled  $K_{12}$ , use as inputs the 2 + 1-flavor lattice-QCD average  $f_K$  from FLAG [6], which only includes calculations where the lattice scale is set from physical inputs other than  $f_\pi$ , and the most recent and precise determination of  $f_{K^\pm}/f_{\pi^\pm}$  from Ref. [46]. The inclusive hadronic  $\tau$ -decay determinations are the most recent ones, from Boyle *et al.* 2018 [36] and Hudspith *et al.* 2017 [35]. The second value from Ref. [36] comes from relating the  $\tau \rightarrow K \ell \nu$  branching fraction to the  $K_{\mu 2}$  branching fraction to get the experimental contribution from the  $K$  pole. The two values in Ref. [35] correspond to using the normalization for  $\tau$  decays into  $K\pi$  modes as obtained in Ref. [37] or as given by HFLAV [39]. For the exclusive  $\tau$  determination we follow the calculation by the HFLAV group [39], but we update the value of the ratio  $f_{K^\pm}/f_{\pi^\pm}$  to that in Ref. [46]. The unitarity value is taken to be  $|V_{us}| = \sqrt{1 - |V_{ud}|^2}$  with  $|V_{ud}|$  from Ref. [2]. RC stands for radiative corrections. The dotted magenta vertical lines correspond to this unitarity value. The gray vertical band corresponds to our result in Eq. (7.1).

## VII. PHENOMENOLOGICAL IMPLICATIONS

### A. Determination of $|V_{us}|$

Combining the form factor in Eq. (6.1) with the latest experimental average  $|V_{us}| f_+^{K^0\pi^-} = 0.21654(41)$  from Ref. [7], we obtain

$$|V_{us}| = 0.22333(44)_{f_+(0)}(42)_{\text{exp}} = 0.22333(61), \quad (7.1)$$

where the first error is from the uncertainty on the form factor, and the second is the experimental uncertainty. Both errors are now of the same size. The experimental error in Eq. (7.1) includes the uncertainty on the long-distance electromagnetic and strong isospin-breaking corrections,  $\delta_{\text{EM}}^{K\ell}$  and  $\delta_{\text{SU}(2)}^{K\pi}$ , which are taken into account when doing the experimental average of the neutral and charged modes [7]. This uncertainty is however dominated by the errors in the lifetime and branching-ratio measurements of the neutral-kaon modes [7]. Other uncertainties such as those from the phase-space integrals are insignificant [7].

In Fig. 10 we compare our extraction of  $|V_{us}|$  from  $K$  semileptonic decays with other determinations using  $K$  semileptonic and leptonic decays, and hadronic  $\tau$  decays. Our semileptonic determination of  $|V_{us}|$  is the most precise to date not relying on an external input for  $|V_{ud}|$ . The central value agrees very well with the most recent lattice

and nonlattice semileptonic calculations, as well as with those based on hadronic tau decay; the latter have much larger errors. Our result, however, is in tension with the leptonic determination using  $f_K/f_\pi$  and with the unitarity prediction given by  $|V_{us}| = \sqrt{1 - |V_{ud}|^2}$  with  $|V_{ud}|$  from Ref. [2]. The agreement with the leptonic determination using  $f_K$  is borderline. The sizes of the disagreements — $2.6\sigma$  with unitarity and  $2.2\sigma$  with the leptonic determination using  $f_K/f_\pi$ —are similar to those using other recent lattice calculations for the semileptonic vector form factor.

As a consistency check of the semileptonic extraction of  $|V_{us}|$ , we can consider the neutral- and charged-kaon modes separately. Using our result in Eq. (6.1) together with the experimental average for neutral modes only [7],  $|V_{us}| f_+^{K^0\pi^-}(0) = 0.2163(5)$ ,<sup>5</sup> we can compare  $|V_{us}|$  as extracted exclusively from neutral-kaon decays:  $|V_{us}|_{K^0\pi^-} = 0.22309(44)_{f_+(0)}(44)_{\text{exp}}(25)_{\delta_{\text{EM}}^{K\ell}} = 0.22309(67)$ . In this case, we can disentangle the purely experimental

<sup>5</sup>Notice that in order to perform the separate averages, Moulson [7] uses the phase-space integrals as extracted from the overall average of form-factor parameters. Although the phase-space factors are affected by isospin-breaking corrections, those corrections are expected to have a negligible impact at this level of precision since the uncertainty on the phase-space integrals currently has an insignificant impact on the experimental averages [7].

error from the uncertainty in the long-distance electromagnetic corrections,  $\delta_{\text{EM}}^{K1}$ , which is the same for all neutral modes,  $\sim 0.22\%$  [77,78]. This result is in very good agreement with the value in Eq. (7.1) within errors, which constitutes a good test of the ChPT calculation of isospin (larger for the charged modes) and EM (larger for the neutral modes) corrections included in the total experimental average, as was already made clear by the results in Ref. [7].

## B. Tests of CKM unitarity

Using our main result for  $|V_{us}|$  in Eq. (7.1), the value  $|V_{ud}| = 0.97420(21)$  from superallowed nuclear  $\beta$  decays [2], and noting that  $|V_{ub}|^2$  is negligible, we find that the measure of deviation from first-row CKM unitarity in Eq. (1.1) is

$$\begin{aligned}\Delta_u &\equiv |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 - 1 \\ &= -0.00104(27)_{V_{us}}(41)_{V_{ud}},\end{aligned}\quad (7.2)$$

which is  $\sim 2.1\sigma$  away from the unitarity prediction, with an error dominated by the uncertainty on  $|V_{ud}|$ . This makes revisiting the determination of  $|V_{ud}|$  a priority for CKM tests. In this vein, one should examine not only superallowed  $\beta$  decays but also other approaches.

At present, the precision in the extraction of  $|V_{ud}|$  from the measurement of the neutron lifetime [3] or pion  $\beta$  decays [4] is still far from that obtained from superallowed  $\beta$  decays. In the case of superallowed  $\beta$  decays, additional measurements will have a small effect on  $|V_{ud}|$ . At the moment, the greatest improvement would come from a calculation of the short-distance radiative correction, which is the main source of uncertainty [2]. A very recent calculation of the nucleus-independent contribution to those corrections, following a new methodology based on dispersion relations [17], obtains a value around  $2\sigma$  larger than the current best determination by Marciano and Sirlin [79] and with a significant reduction of the error. The increased electroweak radiative correction, when combined with the superallowed  $\beta$  decay results [2], results in a lower value of  $|V_{ud}|$ . The authors of Ref. [17] quote  $|V_{ud}| = 0.97366(15)$ . Together with our result for  $|V_{us}|$ , this value of  $|V_{ud}|$  considerably increases the tension with unitarity:

$$\begin{aligned}\Delta_u &\equiv |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 - 1 \\ &= -0.00209(27)_{V_{us}}(29)_{V_{ud}},\end{aligned}\quad (7.3)$$

a more than  $5\sigma$  discrepancy. We discuss further phenomenological implications of this new calculation in Sec. VII D. For the remainder of this section, we use the result by Marciano and Sirlin [79], which leads to  $|V_{ud}| = 0.97420(21)$  and Eq. (7.2).

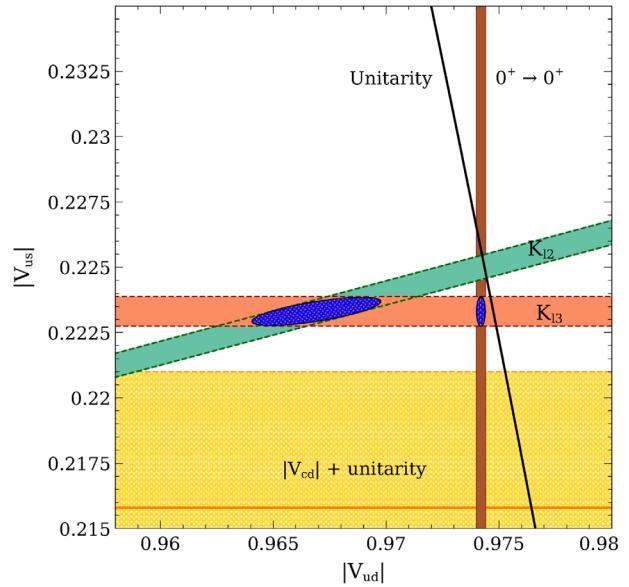


FIG. 11. Constraints on  $|V_{ud}|$  and  $|V_{us}|$  from our results ( $K_{13}$ ), kaon leptonic decays ( $K_{12}$ ), superallowed nuclear  $\beta$  decays, unitarity, and  $|V_{cd}|$ , as discussed in the text. Blue ellipses correspond to the allowed region from  $K_{13}$  and one of the other two constraints with a 68% probability. Both regions have no overlap with unitarity (black line). Correlations between  $K_{12}$  and  $K_{13}$  are not taken into account. The orange horizontal line in the yellow region corresponds to the central value for  $|V_{us}|$  as extracted from  $|V_{cd}|$ .

To avoid using  $|V_{ud}|$  as an input, we can instead perform a unitarity test relying only on experimental kaon-decay measurements [7], on the lattice input from the most recent determination of  $f_{K^+}/f_{\pi^+}$  [46], and on our result in Eq. (6.1) for  $f_+^{K^0\pi^-}(0)$ . The result of the unitarity test using those inputs, noting again that  $|V_{ub}|$  is negligible, is<sup>6</sup>

$$\begin{aligned}\Delta_u &\equiv |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 - 1 \\ &= -0.0151(39)_{f_+(0)}(36)_{f_{K^\pm}/f_{\pi^\pm}}(36)_{\text{exp}}(27)_{\text{EM}},\end{aligned}\quad (7.4)$$

where the  $2.2\sigma$  deviation from unitarity is a reflection of the tension between the leptonic and semileptonic determinations of CKM matrix elements. These results are shown in Fig. 11, together with the test that takes  $|V_{ud}|$  from superallowed  $\beta$  decays as input. No correlation between  $K_{12}$  and  $K_{13}$  inputs, either on the theory or experimental sides, has been taken into account in this test.

<sup>6</sup>The disentanglement of the EM and experimental errors in Eq. (7.4) is approximate, and intended only to indicate the relative size of these errors. The separation of the sources of error is precise for leptonic decays, but for semileptonic decays we assume an overall 0.11% EM error in the uncertainty of the experimental average. This should be a fairly good approximation, however, since the average is dominated by the neutral modes for which the error is indeed 0.11%.

One can perform another test of the unitarity of the CKM matrix by comparing |V<sub>us</sub>| with |V<sub>cd</sub>|, which in the SM should be equal up to corrections of O(λ<sup>5</sup>), with |V<sub>us</sub>| = λ + O(λ<sup>7</sup>). Including the O(λ<sup>5</sup>) corrections, which only affect the last significant digit, the most precise determination of |V<sub>cd</sub>| = 0.2151(6)<sub>f<sub>D</sub></sub>(49)<sub>expt</sub>(6)<sub>EM</sub> from leptonic decays [46] implies the value |V<sub>us</sub>|<sub>|V<sub>cd</sub>|</sub> = 0.2158(52). This value of |V<sub>us</sub>| agrees at the 1.4σ level with our result in Eq. (7.1), although with an uncertainty that is an order of magnitude larger. The uncertainty is dominated by the experimental error on the leptonic decay rate D<sup>+</sup> → ℓ<sup>+</sup>ν, which is expected to be reduced by BESIII and Belle II. This result is also depicted in Fig. 11. As is the case for our main result, |V<sub>us</sub>|<sub>|V<sub>cd</sub>|</sub> is in tension with first-row CKM unitarity by about 2σ when it is used together with |V<sub>ud</sub>| from superallowed nuclear β decays in Eq. (1.1).

Note that, in order to perform this test, we change the normalization of the decay constant f<sub>D<sup>+</sup></sub> obtained in Ref. [46] to account for a change in the scale-setting quantity in that work, f<sub>π<sup>+</sup></sub>, from the PDG value f<sub>π<sup>+</sup></sub> = 130.50 ± 0.13 MeV [3] to the FLAG average f<sub>π<sup>+</sup></sub> = 130.2 ± 0.8 MeV [6]. That gives us f<sub>D<sup>+</sup></sub> = 212.2(0.3)<sub>stat</sub>(0.4)<sub>syst</sub>(1.2)<sub>f<sub>π</sub>,FLAG</sub>[0.2]<sub>EMscheme</sub>.<sup>7</sup> The reason for that change is that the PDG value relies on an external input for |V<sub>ud</sub>|, which is taken from superallowed nuclear β decays, which obscures the comparison. The FLAG number, however, is an average of direct lattice determinations of f<sub>π<sup>+</sup></sub>. With this choice of f<sub>π<sup>+</sup></sub>, the errors are fairly large, and the value of |V<sub>ud</sub>| extracted from experimental data on pion leptonic decays [16] agrees within ∼1.5σ with both |V<sub>ud</sub>| from superallowed nuclear β decays and the value from kaon decays only that we discuss below.

### C. Ratio of leptonic and semileptonic decays

Another way of analyzing the tension between SM kaon leptonic and semileptonic decays is by looking at ratios of decay widths of leptonic and semileptonic decays, where the dependence on |V<sub>us</sub>| cancels. We can construct two ratios

$$\frac{\Gamma(K \rightarrow \ell\nu)}{\Gamma(K \rightarrow \pi\ell\nu)} \propto \left( \frac{f_{K^\pm}}{f_+^{K\pi}(0)} \right)^2, \\ \frac{\Gamma(K \rightarrow \ell\nu)/\Gamma(\pi \rightarrow \ell\nu)}{\Gamma(K \rightarrow \pi\ell\nu)} \propto \frac{1}{|V_{ud}|^2} \left( \frac{f_{K^\pm}/f_{\pi^\pm}}{f_+^{K\pi}(0)} \right)^2. \quad (7.5)$$

The first ratio does not depend on any CKM matrix elements, while the second one is proportional to

<sup>7</sup>Although the dependence of f<sub>D<sup>+</sup></sub> on the scale-setting quantity is much more complicated than a simple linear relation, this estimate should capture most of the effect and, thus, be good enough for this comparison, since its uncertainty is dominated by that of the D<sup>+</sup> → ℓ<sup>+</sup>ν decay rate.

1/|V<sub>ud</sub>|<sup>2</sup>. In addition, the short-distance radiative corrections cancel between numerator and denominator in the first ratio, but not in the second.

Taking experimental averages for the kaon decays and assuming the SM, we obtain<sup>8</sup> [7,16]

$$\left. \frac{f_{K^\pm}}{f_+^{K^0\pi^-}(0)} \right|_{\text{exp.}} = 162.05(40) \text{ MeV}, \\ \left. \frac{1}{|V_{ud}|} \frac{f_{K^\pm}/f_{\pi^\pm}}{f_+^{K^0\pi^-}(0)} \right|_{\text{exp.}} = 1.2745(30). \quad (7.6)$$

With our result in Eq. (6.1) for f<sub>+</sub><sup>K<sup>0</sup>π<sup>-</sup>(0), the average of lattice calculations for f<sub>K<sup>±</sup></sub> = 155.6(0.4) MeV from Ref. [16], f<sub>K<sup>±</sup></sub>/f<sub>π<sup>±</sup></sub> from Ref. [46], and |V<sub>ud</sub>| = 0.97420(21) from [2], those ratios are</sup>

$$\left. \frac{f_{K^\pm}}{f_+^{K^0\pi^-}(0)} \right|_{\text{latt}} = 160.58(79) \text{ MeV}, \\ \left. \frac{1}{|V_{ud}|} \frac{f_{K^\pm}/f_{\pi^\pm}}{f_+^{K^0\pi^-}(0)} \right|_{\text{latt}} = 1.2651^{(+31)}_{(-35)}, \quad (7.7)$$

where we have not taken into account any correlation between the decay constants and the form factor. Comparing Eqs. (7.6) and (7.7), we see some tension, ∼1.7σ and 2.2σ, respectively, between the SM predictions and the experimental measurements. The error from lattice QCD is the main limiting factor in this comparison, but that can be reduced by taking into account the correlation between the numerator and denominator in Eq. (7.7), which we plan to do in the future.

Alternatively, one can compare the ratio [f<sub>K<sup>±</sup></sub>/f<sub>π<sup>±</sup></sub>]/[|V<sub>ud</sub>|f<sub>+</sub><sup>K<sup>0</sup>π<sup>-</sup>(0)] as extracted from experiment and theory to get a value of the CKM matrix element |V<sub>ud</sub>|, and compare it with the value from superallowed nuclear β decays. The result of such an exercise is |V<sub>ud</sub>| = 0.9669(19)<sub>f<sub>+</sub></sub>(+13)<sub>-19</sub><sub>f<sub>K</sub>/f<sub>π</sub></sub>(23)<sub>exp</sub> = 0.9669(+32)<sub>-35</sub>, approximately 2.1σ lower than the value from superallowed β decays. This result is seen in Fig. 11 at the intersection of the two bands for K<sub>ℓ3</sub> and K<sub>ℓ2</sub>. It also deviates from the unitarity condition.</sup>

The unitarity test comparing |V<sub>us</sub>|/|V<sub>ud</sub>| with |V<sub>us</sub>|<sub>|V<sub>cd</sub>|</sub>/|V<sub>ud</sub>|, again including corrections up to O(λ<sup>5</sup>), and taking the decay constants f<sub>K<sup>+</sup></sub>/f<sub>π<sup>+</sup></sub> and f<sub>D<sup>+</sup></sub>/f<sub>π<sup>+</sup></sub> from Ref. [46] and the experimental data on leptonic experimental data from Ref. [16], fails at the 2σ level. This test is

<sup>8</sup>We take Γ(K → ℓν) from Ref. [16], which does not use the same value of the universal short-distance electroweak correction S<sub>EW</sub> as [7] (from which we take the other experimental averages). The imperfect cancellation is too small to affect the conclusion drawn here.

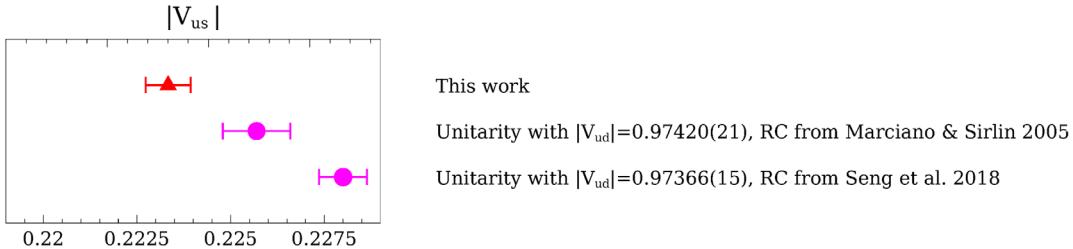


FIG. 12. Comparison of the unitarity point using  $|V_{ud}| = 0.97366(15)$  with the results in this work, and with the unitarity point corresponding to  $|V_{ud}| = 0.97420(21)$ . RC stands for radiative corrections.

limited by the experimental error on the  $D^+$  leptonic decay rate.

#### D. Implications of the new extraction of $|V_{ud}|$

If the decrease of the central value and uncertainty of the nucleus-independent electroweak radiative corrections involved in the extraction of  $|V_{ud}|$  from superallowed  $\beta$  decays in Ref. [17] is confirmed, the new value  $|V_{ud}| = 0.97366(15)$  would exacerbate some of the tensions we have just discussed.

First, as shown above, this value of  $|V_{ud}|$  and our semileptonic result for  $|V_{us}|$  would imply a greater than  $5\sigma$  violation of first-row CKM unitarity. The tension between our semileptonic value of  $|V_{us}|$  and the one extracted from kaon leptonic decays and  $f_{K^\pm}/f_{\pi^\pm}$ , however, would be slightly reduced to  $2\sigma$ , since a smaller value of  $|V_{ud}|$  would give a smaller value of the leptonic  $|V_{us}|$ , closer to our semileptonic extraction. For the same reason, the tension between the ratios involving  $|V_{ud}|$  in Eqs. (7.6) and (7.7) would be slightly lessened.

In Fig. 12, as an example, we show the comparison of the unitarity prediction  $\sqrt{1 - |V_{ud}|^2}$  for  $|V_{us}|$  using both  $|V_{ud}| = 0.97366(15)$  and  $|V_{ud}| = 0.97420(21)$ , together with the results in this work. Given the important implications of a value of  $|V_{ud}|$  with a smaller error and a smaller central value, it is very important to confirm the new calculation of radiative corrections in Ref. [17], and to understand the discrepancy with the previous best determination in Ref. [79].

### VIII. CONCLUSIONS AND OUTLOOK

Using the HISQ  $N_f = 2 + 1 + 1$  MILC ensembles, we have performed the most precise computation to date of the vector form factor at zero momentum transfer,  $f_+^{K^0\pi^-}(0)$ , and the first one to include the dominant FV effects, as calculated in ChPT at NLO. Our result for the form factor enables a direct determination of the CKM matrix element  $|V_{us}|$  from semileptonic kaon decays with a theory error that is, for the first time, at the same level as the experimental error. Further, the uncertainty in this direct determination is now similar to those from indirect determinations based on leptonic decays with  $|V_{ud}|$  as input.

A key to achieving this level of precision is simulating at near-physical values of the quark masses, which drastically reduces the systematic errors associated with the chiral extrapolation (replacing it with an interpolation), as well as the error coming from the chiral LECs that are inputs to the analysis. The finite-volume effects, one of the main sources of uncertainty in our previous analyses, have also been significantly reduced by explicitly including them at NLO (the leading nontrivial order) in ChPT. The dominant remaining source of error is now statistics, which could be reduced by extending the key ensembles with physical quark masses, and including the existing MILC physical-mass ensemble with a finer lattice spacing of  $a \approx 0.042$  fm.

Another important error arises from the uncertainty in the ChPT LECs of order  $p^6$ . That uncertainty could be reduced by performing a combined analysis of form-factor data together with light meson masses and decay constants, which would put more constraints on the ChPT LECs. In particular, the error from  $L_8$  is comparable to, but greater than, that from  $L_7$ , and the combined analysis could significantly reduce the  $L_8$  error. Errors from  $L_4$ ,  $L_5$ , and  $L_6$  would also be reduced, but they have a much smaller effect on the total error here.

We find that the extraction of  $|V_{us}|$  from semileptonic kaon decays is in tension both with the extraction from leptonic kaon decays and with unitarity at the  $\sim 2$ – $2.6\sigma$  level. In particular, the unitarity test based only on kaon decay data, without any external input for  $|V_{ud}|$ , and having as nonperturbative inputs  $f_+^{K^0\pi^-}(0)$  from this work and  $f_{K^\pm}/f_{\pi^\pm}$  from Ref. [46], shows a  $\sim 2.2\sigma$  tension. While unitarity tests based on  $|V_{ud}|$  are currently limited by the uncertainty in that matrix element, the tension with unitarity would raise to the  $5\sigma$  level if the new calculation of radiative corrections involved in the extraction of  $|V_{ud}|$  from superallowed  $\beta$  decays [17] is confirmed.

The test based on kaon-decay data has similarly sized uncertainties arising from both theory and experiment. In order to shed light on these tensions, improvements from both the theoretical and experimental sides are urgently needed, as are improvements in other approaches. A new round of experiments is expected to reduce the experimental error to  $\sim 0.12\%$  in the next few years [7].

More importantly, the new high-statistics data will help to check the consistency of current fits, and to perform a more thorough study of systematic errors on the experimental averages.

For the experimental determination of |V<sub>us</sub>|f<sub>+</sub><sup>K<sup>0</sup>π<sup>-</sup>, electromagnetic and isospin effects are currently being estimated using phenomenology and ChPT techniques. Although they are not yet a dominant source of error (EM effects make a 0.11% correction to the individual neutral channels), with the reduction of other sources of error and the forthcoming improvement in the experimentally measured branching ratios and lifetimes, they will eventually need to be included directly in the lattice-QCD simulations. Recent efforts in that direction can be found in Refs. [80–83].</sup>

Isospin corrections are numerically important for the charged kaon channels, where those effects enter already at LO through  $\pi^0$ - $\eta$  mixing. The NNLO ChPT estimate of the corrections for the charged modes has large errors [67] due to the unknown value of the O( $p^6$ ) LECs. Fortunately, the experimental average is dominated by the neutral-kaon channels, so the charged-mode uncertainty does not have a large effect on the final experimental average. The strong isospin-breaking correction  $\delta_{SU(2)}$  used in the experimental average is a NLO ChPT estimate that partially includes NNLO effects; it does not include the uncertainty associated with higher-order terms in the chiral expansion. However, the fact that the value used in the average and the one extracted from experiment are so close [2.45(19)% vs 2.82(38)% [7]], that the result for |V<sub>us</sub>| using only the neutral modes agrees with the one using all decay modes (see Sec. VII A), and that neutral modes are the dominant ones in the average, indicates that the experimental average using this estimate is robust.

The uncertainties from the phase-space integrals are insignificant at present in the final error for the experimental average. It is therefore not crucial at present to have a better representation of those, i.e., to have the  $q^2$  dependence of the form factors. In the future, however, lattice calculations of f<sub>+</sub><sup>Kπ</sup>( $q^2$ ) could provide better determinations of the form-factor slope than those relying on experimental data [7,26].

An important future step in the investigation of the tensions observed in the first-row unitarity relation, and in the value of |V<sub>us</sub>| extracted from different sources, will be to perform a correlated analysis of semileptonic and leptonic kaon decays. That analysis would provide a more precise value of the ratio [f<sub>K<sup>±</sup></sub>/f<sub>π<sup>±</sup></sub>]/[|V<sub>ud</sub>|f<sub>+</sub><sup>Kπ</sup>(0)] and potentially give an insight into the tensions. Another key point in the study of those tensions is clarifying the role of the electroweak radiative corrections in the extraction of |V<sub>ud</sub>| from superallowed  $\beta$  decays, as well as reducing the error of that CKM matrix element as extracted, not only from superallowed  $\beta$  decays, but from other sources.

## ACKNOWLEDGMENTS

We thank Matthew Moulson for useful discussions and Bipasha Chakraborty for participating in an early stage of this analysis. We thank Johan Bijnens for making his isospin-breaking NLO partially quenched ChPT and isospin-breaking NNLO full QCD codes available to us, and Bijnens and Johan Relefors for making their FV ChPT code available to us. We thank Zechariah Gelzer for discussions on autocorrelations in the MILC ensembles. Computations for this work were carried out with resources provided by the USQCD Collaboration, the National Energy Research Scientific Computing Center, the Argonne Leadership Computing Facility, the Blue Waters sustained-petascale computing project, the National Institute for Computational Science, the National Center for Atmospheric Research, and the Texas Advanced Computing Center. USQCD resources are acquired and operated thanks to funding from the Office of Science of the U.S. Department of Energy. The National Energy Research Scientific Computing Center is a Department of Energy Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research used resources of the Argonne Leadership Computing Facility, which is a Department of Energy Office of Science User Facility supported under Contract No. DE-AC02-06CH11357. The Blue Waters sustained-petascale computing project is supported by the National Science Foundation (Grants No. OCI-0725070 and No. ACI-1238993) and the State of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. This work is also part of the “Lattice QCD on Blue Waters” and “High Energy Physics on Blue Waters” PRAC allocations supported by the National Science Foundation (Grants No. 0832315 and No. 1615006) and used an allocation received under the “Blue Waters for Illinois faculty” program. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant No. ACI-1548562 [161]. Allocations under the Teragrid and XSEDE programs included resources at the National Institute for Computational Sciences (NICS) at the Oak Ridge National Laboratory Computer Center, the Texas Advanced Computing Center and the National Center for Atmospheric Research, all under NSF teragrid allocation TG-MCA93S002. Computer time at the National Center for Atmospheric Research was provided by NSF MRI Grant No. CNS-0421498, NSF MRI Grant No. CNS-0420873, NSF MRI Grant No. CNS-0420985, NSF sponsorship of the National Center for Atmospheric Research, the University of Colorado, and a grant from the IBM Shared University Research (SUR) program.

This work was supported in part by the U.S. Department of Energy under Grants No. DE-FG02-91ER40628 (C. B.), No. DE-FC02-12ER41879 (C. D.), No. DE-SC0010120 (S. G.), No. DE-FG02-91ER40661 (S. G.), No. DE-FG02-13ER42001 (A. X. K.), No. DE-SC0015655 (A. X. K.), No. DE-SC0010005 (E. T. N.), No. DE-FG02-13ER41976 (D. T.); by the U.S. National Science Foundation under Grants No. PHY14-14614 and No. PHY17-19626 (C. D.), No. PHY14-17805 (J. L.), and No. PHY13-16748 and No. PHY16-20625 (R. S.); by the MINECO (Spain) under Grants No. FPA2013-47836-C-1-P and No. FPA2016-78220-C3-3-P (E. G.); by the Junta de Andalucía (Spain) under Grant No. FQM-101

(E. G.); by the European Commission (EC) under Grant No. PCIG10-GA-2011-303781 (E. G.); by the U.K. Science and Technology Facilities Council (J. K.); by the German Excellence Initiative and the European Union Seventh Framework Program under Grant No. 291763 as well as the European Union Marie Curie COFUND program (J. K., A. S. K.). Brookhaven National Laboratory is supported by the United States Department of Energy, Office of Science, Office of High Energy Physics, under Contract No. DE-SC0012704. Fermilab is operated by Fermi Research Alliance, LLC, under Contract No. DE-AC02-07CH11359 with the United States Department of Energy, Office of Science, Office of High Energy Physics.

---

[1] M. González-Alonso and J. Martín Camalich, *J. High Energy Phys.* **12** (2016) 052.

[2] J. C. Hardy and I. S. Towner, [arXiv:1807.01146](https://arxiv.org/abs/1807.01146).

[3] C. Patrignani *et al.* (Particle Data Group), *Chin. Phys. C* **40**, 100001 (2016).

[4] D. Počanić *et al.*, *Phys. Rev. Lett.* **93**, 181803 (2004).

[5] D. Počanić, Talk at the Ninth International Workshop on the CKM Unitarity Triangle, 2016, Mumbai, <https://indico.tifr.res.in/indico/contributionDisplay.py?contribId=46&confId=5095>.

[6] S. Aoki *et al.* (Flavour Lattice Averaging Group), *Eur. Phys. J. C* **77**, 112 (2017).

[7] M. Moulson, *Proc. Sci., CKM2016* (2017) 033 [[arXiv:1704.04104](https://arxiv.org/abs/1704.04104)].

[8] S. Dürr *et al.*, *Phys. Rev. D* **95**, 054513 (2017).

[9] N. Carrasco *et al.* (ETM Collaboration), *Phys. Rev. D* **91**, 054507 (2015).

[10] T. Blum *et al.* (RBC and UKQCD Collaborations), *Phys. Rev. D* **93**, 074505 (2016).

[11] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), *Phys. Rev. D* **90**, 074509 (2014).

[12] R. J. Dowdall, C. T. H. Davies, G. P. Lepage, and C. McNeile (HPQCD Collaboration), *Phys. Rev. D* **88**, 074504 (2013).

[13] A. Bazavov *et al.* (MILC Collaboration), *Proc. Sci., LATTICE2010* (2010) 074 [[arXiv:1012.0868](https://arxiv.org/abs/1012.0868)].

[14] S. Dürr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, A. Ramos, and K. K. Szabó (Budapest-Marseille-Wuppertal Collaboration), *Phys. Rev. D* **81**, 054507 (2010).

[15] E. Follana, C. T. H. Davies, G. P. Lepage, and J. Shigemitsu (HPQCD Collaboration), *Phys. Rev. Lett.* **100**, 062002 (2008).

[16] J. L. Rosner, S. Stone, and R. S. Van de Water, in Review of Particle Physics, [arXiv:1509.02220](https://arxiv.org/abs/1509.02220).

[17] C. Y. Seng, M. Gorchtein, H. H. Patel, and M. J. Ramsey-Musolf, *Phys. Rev. Lett.* **121**, 241804 (2018).

[18] V. Cirigliano, G. Ecker, H. Neufeld, A. Pich, and J. Portolés, *Rev. Mod. Phys.* **84**, 399 (2012).

[19] A. Sirlin, *Rev. Mod. Phys.* **50**, 573 (1978); **50**, 905(E) (1978).

[20] A. Sirlin, *Nucl. Phys.* **B196**, 83 (1982).

[21] W. J. Marciano and A. Sirlin, *Phys. Rev. Lett.* **71**, 3629 (1993).

[22] V. Cirigliano, M. Giannotti, and H. Neufeld, *J. High Energy Phys.* **11** (2008) 006.

[23] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), *Phys. Rev. Lett.* **112**, 112001 (2014).

[24] E. Gámiz *et al.* (Fermilab Lattice and MILC Collaborations), *Proc. Sci., LATTICE2013* (2014) 395 [[arXiv:1311.7264](https://arxiv.org/abs/1311.7264)].

[25] P. A. Boyle *et al.* (RBC and UKQCD Collaborations), *J. High Energy Phys.* **06** (2015) 164.

[26] N. Carrasco, P. Lami, V. Lubicz, L. Riggio, S. Simula, and C. Tarantino (ETM Collaboration), *Phys. Rev. D* **93**, 114512 (2016).

[27] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), *Phys. Rev. D* **87**, 073012 (2013).

[28] P. A. Boyle, J. M. Flynn, N. Garron, A. Jüttner, C. T. Sachrajda, K. Sivalingam, and J. M. Zanotti (RBC and UKQCD Collaborations), *J. High Energy Phys.* **08** (2013) 132.

[29] P. A. Boyle, J. M. Flynn, A. Jüttner, C. Kelly, C. Maynard, H. P. de Lima, C. T. Sachrajda, and J. M. Zanotti (RBC and UKQCD Collaborations), *Eur. Phys. J. C* **69**, 159 (2010).

[30] S. Aoki, G. Cossu, X. Feng, H. Fukaya, S. Hashimoto, T. Kaneko, J. Noaki, and T. Onogi (JLQCD Collaboration), *Phys. Rev. D* **96**, 034501 (2017).

[31] T. Primer *et al.* (Fermilab Lattice and MILC Collaborations), *Proc. Sci., LATTICE2014* (2014) 374 [[arXiv:1411.1651](https://arxiv.org/abs/1411.1651)].

[32] E. Gámiz *et al.* (Fermilab Lattice and MILC Collaborations), *Proc. Sci., LATTICE2016* (2016) 286 [[arXiv:1611.04118](https://arxiv.org/abs/1611.04118)].

[33] V. Mateu and A. Pich, *J. High Energy Phys.* **10** (2005) 041.

[34] E. Gámiz, M. Jamin, A. Pich, J. Prades, and F. Schwab, *Phys. Rev. Lett.* **94**, 011803 (2005).

[35] R. J. Hudspith, R. Lewis, K. Maltman, and J. Zanotti, *Phys. Lett. B* **781**, 206 (2018).

[36] P. Boyle *et al.* (RBC and UKQCD Collaborations), arXiv:1803.07228.

[37] M. Antonelli, V. Cirigliano, A. Lusiani, and E. Passemar, *J. High Energy Phys.* **10** (2013) 070.

[38] E. Kou *et al.*, arXiv:1808.10567.

[39] Y. Amhis *et al.* (HFLAV Collaboration), *Eur. Phys. J. C* **77**, 895 (2017).

[40] C. Bernard, J. Bijnens, E. Gámiz, and J. Relefors, *J. High Energy Phys.* **03** (2017) 120.

[41] C. Bernard and D. Toussaint (MILC Collaboration), *Phys. Rev. D* **97**, 074502 (2018).

[42] H. Na, C. T. H. Davies, E. Follana, P. Lepage, and J. Shigemitsu, *Proc. Sci., LATT2009* (2009) 247 [arXiv: 0910.3919].

[43] J. Koponen *et al.* (HPQCD Collaboration), arXiv: 1208.6242.

[44] A. Bazavov *et al.* (MILC Collaboration), *Phys. Rev. D* **82**, 074501 (2010).

[45] A. Bazavov *et al.* (MILC Collaboration), *Phys. Rev. D* **87**, 054505 (2013).

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

[47] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD and UKQCD Collaborations), *Phys. Rev. D* **75**, 054502 (2007).

[48] C. McNeile and C. Michael, *Phys. Rev. D* **73**, 074506 (2006).

[49] E. Gámiz *et al.* (Fermilab Lattice and MILC Collaborations), *Proc. Sci., LATTICE2012* (2012) 113 [arXiv: 1211.0751].

[50] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), *Phys. Rev. D* **93**, 113016 (2016).

[51] A. Bazavov *et al.* (MILC Collaboration), *Phys. Rev. D* **81**, 114501 (2010).

[52] R. Brower, S. Chandrasekharan, J. W. Negele, and U. J. Wiese, *Phys. Lett. B* **560**, 64 (2003).

[53] S. Aoki, H. Fukaya, S. Hashimoto, and T. Onogi, *Phys. Rev. D* **76**, 054508 (2007).

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

[55] H. Leutwyler and A. V. Smilga, *Phys. Rev. D* **46** (1992) 5607.

[56] B. Billeter, C. E. Detar, and J. Osborn, *Phys. Rev. D* **70**, 077502 (2004).

[57] M. Ademollo and R. Gatto, *Phys. Rev. Lett.* **13**, 264 (1964).

[58] C. Bernard, J. Bijnens, and E. Gámiz, *Phys. Rev. D* **89**, 054510 (2014).

[59] J. Bijnens and P. Talavera, *Nucl. Phys.* **B669**, 341 (2003).

[60] H. W. Fearing and S. Scherer, *Phys. Rev. D* **53**, 315 (1996).

[61] J. Bijnens, G. Colangelo, and G. Ecker, *J. High Energy Phys.* **02** (1999) 020.

[62] A. Bazavov *et al.*, *Rev. Mod. Phys.* **82**, 1349 (2010).

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

[64] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), *Phys. Rev. D* **90**, 074509 (2014).

[65] J. Bijnens and G. Ecker, *Annu. Rev. Nucl. Part. Sci.* **64**, 149 (2014).

[66] A. Bazavov *et al.* (MILC Collaboration), *Proc. Sci., LATTICE2011* (2011) 107 [arXiv:1111.4314].

[67] J. Bijnens and K. Ghorbani, arXiv:0711.0148.

[68] J. Gasser and H. Leutwyler, *Nucl. Phys.* **B250**, 517 (1985).

[69] C. 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).

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

[71] S. Basak *et al.* (MILC Collaboration), *Phys. Rev. D* **99**, 034503 (2019).

[72] G. Ecker, *Proc. Sci., CD15* (2015) 011 [arXiv:1510.01634].

[73] A. Kastner and H. Neufeld, *Eur. Phys. J. C* **57**, 541 (2008).

[74] V. Cirigliano, G. Ecker, M. Eidemüller, R. Kaiser, A. Pich, and J. Portolés, *J. High Energy Phys.* **04** (2005) 006.

[75] M. Jamin, J. A. Oller, and A. Pich, *J. High Energy Phys.* **02** (2004) 047.

[76] H. Leutwyler and M. Roos, *Z. Phys. C* **25**, 91 (1984).

[77] V. Cirigliano, H. Neufeld, and H. Pichl, *Eur. Phys. J. C* **35**, 53 (2004).

[78] V. Cirigliano, M. Knecht, H. Neufeld, H. Rupertsberger, and P. Talavera, *Eur. Phys. J. C* **23**, 121 (2002).

[79] W. J. Marciano and A. Sirlin, *Phys. Rev. Lett.* **96**, 032002 (2006).

[80] N. Carrasco, V. Lubicz, G. Martinelli, C. T. Sachrajda, N. Tantalo, C. Tarantino, and M. Testa, *Phys. Rev. D* **91**, 074506 (2015).

[81] D. Giusti, V. Lubicz, G. Martinelli, C. T. Sachrajda, F. Sanfilippo, S. Simula, N. Tantalo, and C. Tarantino, *Phys. Rev. Lett.* **120**, 072001 (2018).

[82] A. Patella, *Proc. Sci., LATTICE2016* (2017) 020. [arXiv: 1702.03857].

[83] C. T. Sachrajda, Proceedings of the 36th International Symposium on Lattice Field Theory, East Lansing, Michigan, 2018. <https://indico.fnal.gov/event/15949/contribution/163>.