PHYSICAL REVIEW D **87,** 073012 (2013)

Kaon semileptonic vector form factor and determination of $|V_{us}|$ using staggered fermions

A. Bazavov,[1] C. Bernard,[2] C. M. Bouchard,[3] C. DeTar,[4] Daping Du,[5] A. X. El-Khadra,[5] J. Foley,[4] E. D. Freeland,[6] E. Gámiz,[7,*] Steven Gottlieb,[8] U. M. Heller,[9] Jongjeong Kim,[10] A. S. Kronfeld,[11] J. Laiho,[12] L. Levkova,[4] P. B. Mackenzie,[11] E. T. Neil,[11] M. B. Oktay,[4] Si-Wei Qiu,[4] J. N. Simone,[11] R. Sugar,[13] D. Toussaint,[10] R. S. Van de Water,[11] and Ran Zhou[8]

(Fermilab Lattice and MILC Collaborations)

[1]*Physics Department, Brookhaven National Laboratory, Upton, New York, USA*
[2]*Department of Physics, Washington University, St. Louis, Missouri, USA*
[3]*Department of Physics, The Ohio State University, Columbus, Ohio, USA*
[4]*Physics Department, University of Utah, Salt Lake City, Utah, USA*
[5]*Physics Department, University of Illinois, Urbana, Illinois, USA*
[6]*Department of Physics, Benedictine University, Lisle, Illinois, USA*
[7]*CAFPE and Departamento de Física Teórica y del Cosmos, Universidad de Granada, Granada, Spain*
[8]*Department of Physics, Indiana University, Bloomington, Indiana, USA*
[9]*American Physical Society, Ridge, New York, USA*
[10]*Department of Physics, University of Arizona, Tucson, Arizona, USA*
[11]*Fermi National Accelerator Laboratory, Batavia, Illinois, USA*
[12]*SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom*
[13]*Department of Physics, University of California, Santa Barbara, California, USA*
(Received 21 December 2012; published 23 April 2013)

Using staggered fermions and partially twisted boundary conditions, we calculate the K meson semileptonic decay vector form factor at zero momentum transfer. The highly improved staggered quark formulation is used for the valence quarks, while the sea quarks are simulated with the asqtad action (MILC Collaboration $N_f = 2 + 1$ configurations). For the chiral and continuum extrapolation, we use two-loop continuum χPT, supplemented by partially quenched staggered χPT at one loop. Our result is $f_+^{K\pi}(0) = 0.9667 \pm 0.0023 \pm 0.0033$, where the first error is statistical and the second is the sum in quadrature of the systematic uncertainties. This result is the first $N_f = 2 + 1$ calculation with two lattice spacings and a controlled continuum extrapolation. It is also the most precise result to date for the vector form factor, and, although the central value is larger than previous unquenched lattice calculations, it is compatible with them within errors. Combining our value for $f_+^{K\pi}(0)$ with the latest experimental measurements of K semileptonic decays, we obtain $|V_{us}| = 0.2238 \pm 0.0009 \pm 0.0005$, where the first error is from $f_+^{K\pi}(0)$ and the second one is experimental. As a byproduct of our calculation, we obtain the combination of low-energy constants $(C_{12}^r + C_{34}^r - (L_5^r)^2)(M_\rho) = (3.62 \pm 1.00) \times 10^{-6}$.

DOI: 10.1103/PhysRevD.87.073012 PACS numbers: 12.15.Hh, 12.38.Gc, 13.20.Eb

I. INTRODUCTION

The Standard Model (SM) predicts that the matrix that describes the mixing between the different quark flavors, the Cabibbo-Kobayashi-Maskawa (CKM) matrix, has to be unitary. Any deviation from unitarity would indicate the existence of beyond-the-Standard-Model physics, so checking CKM unitarity is very important in the search for new physics (NP). In particular, tests of the unitarity of the first row of the CKM matrix can establish stringent constraints on the scale of allowed NP, even if unitarity is fulfilled. Given current experimental and theoretical inputs, the low-energy constraint coming from these tests ($\Lambda > 11$ TeV, where Λ is the scale of NP) [1] is at the same level as those from Z-pole measurements [2].

The two most precise grounds for the determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{us}|$ are leptonic and semileptonic decays, with competing errors at the 0.6% level [2]. Determinations of $|V_{us}|$ from hadronic τ decays have the potential of being competitive with the above determinations, but they are currently limited by uncertainties in the experimental data [3].

Both leptonic and semileptonic determinations require nonperturbative inputs, which can be precisely calculated on the lattice. In the case of leptonic decays, the ratio of the decay rates $K \to \mu\nu$ and $\pi \to \mu\nu$ can be combined with the ratio of decay constants f_K/f_π, calculated on the lattice with a precision better than 0.5% for the averaged value [4–13], to extract $|V_{us}/V_{ud}|$ [14]. For semileptonic decays, the photon-inclusive decay rate for all $K \to \pi l \nu$ decay modes can be related to the CKM matrix element $|V_{us}|$ via Eq. (4.37) of Ref. [2],

*megamiz@ugr.es

$$\Gamma_{K_{l3(\gamma)}} = \frac{G_F^2 M_K^5 C_K^2}{128\pi^3} S_{\rm EW} |V_{us} f_+^{K^0\pi^-}(0)|^2 I_{Kl}^{(0)} \times (1 + \delta_{\rm EM}^{Kl} + \delta_{\rm SU(2)}^{K\pi}), \quad (1.1)$$

where the Clebsch-Gordan coefficient C_K is equal to 1 or $1/\sqrt{2}$ for neutral and charged kaons, respectively; $S_{\rm EW} = 1.0223(5)$ is the short-distance universal electroweak correction; and $I_{Kl}^{(0)}$ is a phase space integral that depends on the shape of the form factors $f_\pm^{K\pi}$. The parameters $\delta_{\rm EM}^{Kl}$ and $\delta_{\rm SU(2)}^{K\pi}$ contain long-distance electromagnetic and strong isospin-breaking corrections, respectively, and are defined as corrections to the K^0 mode. A discussion of the values and different determinations of those parameters can be found in Ref. [2]. We note that a different normalization from that of Ref. [2] is often used in the literature for the phase space integral $I_{Kl}^{(0)}$, in which case the factor 128 in Eq. (1.1) is replaced by a factor of 192.

The vector form factor $f_+^{K\pi}$, together with the scalar form factor $f_0^{K\pi}$, describes the hadronic matrix element of a vector current

$$\langle \pi(p_\pi)|V^\mu|K(p_K)\rangle = f_+^{K\pi}(q^2)\left[p_K^\mu + p_\pi^\mu - \frac{m_K^2 - m_\pi^2}{q^2} q^\mu\right] + f_0^{K\pi}(q^2)\frac{m_K^2 - m_\pi^2}{q^2} q^\mu, \quad (1.2)$$

where $q = p_K - p_\pi$ is the momentum transfer and $V^\mu = \bar{s}\gamma^\mu u$ is the appropriate flavor-changing vector current. It is convenient to factor out $f_+^{K^0\pi^-}$ in Eq. (1.1) and normalize all experimental measurements to this channel. From now on, we will denote $f_+^{K^0\pi^-}(q^2)$ by $f_+(q^2)$ and $f_0^{K^0\pi^-}(q^2)$ by $f_0(q^2)$. Thus, by definition, $f_+(0)$ is given by the hadronic matrix element (1.2) for the $K^0\pi^-$ channel with meson masses set equal to the physical ones.

The error associated with the lattice-QCD determination of $f_+(q^2 = 0)$ is around 0.5%, but it is still the dominant uncertainty in the extraction of $|V_{us}|$ from experimental data on K semileptonic decays. References [15,16] quote $|V_{us}|f_+(0) = 0.2163(5)$, an experimental error of 0.23%. Improvement in the determination of the form factor is, thus, crucial in order to extract all the information from the available experimental data.

In the search for deviations from SM predictions, one can also compare the values of $|V_{us}|$ as extracted from helicity-allowed semileptonic decays and helicity-suppressed leptonic decays. In particular, it is useful to study the ratio

$$R_{\mu 23} = \left(\frac{f_+^{K\pi}(0)}{f_K/f_\pi}\right)\left(\left|\frac{V_{us}}{V_{ud}}\right|\left|\frac{f_K}{f_\pi}\right)_{\mu 2} |V_{ud}| \frac{1}{[|V_{us}| f_+^{K\pi}(0)]_{l3}}, \quad (1.3)$$

where the subscripts $\mu 2$ and $l3$ indicate that these quantities are obtained from experimental data on leptonic $K_{\mu 2}$ and semileptonic K_{l3} decays, respectively, and the corresponding SM formulas. The ratio in Eq. (1.3) is unity in the SM by construction, but not in some extensions of the SM, such as those with a charged Higgs. Again, the error in the current value $R_{\mu 23} = 0.999(7)$ [15] using the average of $|V_{ud}|$ from superallowed nuclear beta decays in Ref. [17] is limited by the precision of lattice-QCD inputs.

In this paper, we report on a lattice-QCD calculation of the form factor $f_+(0)$ using the highly improved staggered quark (HISQ) action [18] for valence quarks and the asqtad-improved staggered action [19] for sea quarks. The goal of this analysis is to provide, using the staggered formulation, a determination of this parameter that is competitive with other state-of-the-art unquenched determinations [20,21]. In order to avoid the complications associated with the fact that the simplest staggered vector currents are not conserved and need to be renormalized, we use the method developed by the HPQCD Collaboration in Ref. [22] and described in Sec. II. This method uses correlation functions of a scalar current, which do not need a renormalization factor. Some tests of the general methodology and specific aspects such as the use of random wall sources, the stability of the correlator fits, or the chiral and continuum extrapolations were already discussed in Refs. [23–25]. Preliminary results for $f_+(0)$ using different light-quark actions, such as domain wall or overlap, can be found in Refs. [26,27], respectively. An even more precise calculation using HISQ valence quarks on HISQ configurations is already under way [25]. The result presented here, however, is the first $N_f = 2 + 1$ determination of $f_+(0)$ from multiple lattice spacings with a controlled continuum extrapolation, and the errors are already smaller than those of other recently published calculations [20,21].

This paper is organized as follows. In Sec. II, we describe the main characteristics of our approach, and, in Sec. III, we describe the simulation details as well as the fitting strategy for the correlator fits. Section IV is dedicated to the chiral and continuum extrapolation, and Sec. V compiles the discussions on the different sources of systematic errors in our calculation. In Sec. VI, we give the result for the $O(p^6)$ low-energy constants relevant for the calculation of $f_+(0)$. Finally, we present our final results and conclusions in Sec. VII.

II. METHODOLOGY

One of the main components of our analysis, which reduces both systematic and statistical errors, is to use the method developed by the HPQCD Collaboration to study charm semileptonic decays [22]. This method is based on the Ward identity relating the matrix element of a vector current to that of the corresponding scalar current,

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

with Z_V a lattice renormalization factor for the vector current. The scalar current is $S = \bar{s}q$, where field q represents degenerate u and d quarks in our simulations.

In this work, we use the local scalar density of staggered fermions, as described in the following section, so the combination $(m_s - m_q)S^{\mathrm{lat}}$ requires no renormalization ($Z_S = 1$). Using the definition of the form factors in Eq. (1.2) and the identity in Eq. (2.1), one can extract $f_0(q^2)$ at any value of the momentum transfer q^2 by using

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

Kinematic constraints demand that $f_+(0) = f_0(0)$, so this relation can be used to calculate $f_+(0)$ and, thus, to extract $|V_{us}|$, the goal of this work. The main advantage of relation (2.2) is that it does not need a renormalization factor to obtain the form factor $f_0(q^2)$. With staggered fermions, another key point is that one can avoid the calculation of correlation functions with a vector current insertion. The staggered local vector current, unlike the staggered local scalar current, is not a taste singlet. This implies that, in order to have a nonvanishing signal, we would need to use either external Goldstone mesons (with pseudoscalar taste) with a nonlocal vector current or external non-Goldstone mesons with a local vector current [28]. Those correlation functions typically have larger statistical errors than the ones for a correlation function with a local current and external Goldstone mesons [29].

The general structure of the three-point function that gives us access to the matrix element in Eq. (2.2) is depicted in Fig. 1. We generate light quarks at a time slice t_{source} and extended strange propagators at a fixed distance T from the source. At t_{source}, we use random-wall sources in a manner similar to that of Ref. [30]. Specifically, we compute quark propagators from each of four Gaussian-stochastic color vector fields with support on all three color components and all spatial sites at that source time. This method greatly reduces statistical errors, as discussed in Ref. [23]. We place the scalar current at a time position $t_{\mathrm{source}} + t$ between the source and the sink and contract the extended strange propagator with a light propagator. We then study the t dependence to isolate the desired matrix element.

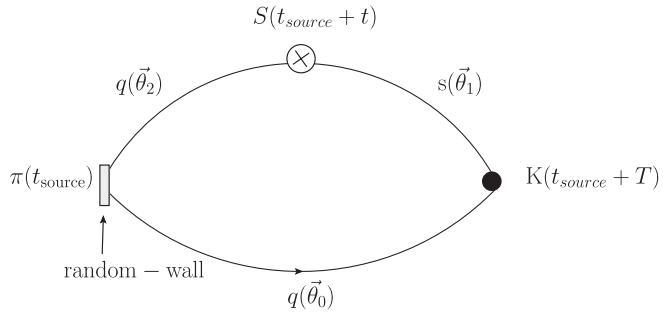

FIG. 1. Structure of the three-point functions needed to calculate $f_0(q^2)$. Light-quark propagators are generated at t_{source} with random-wall sources. An extended strange propagator is generated at $T + t_{\mathrm{source}}$.

Another key ingredient in our calculation is the use of partially twisted boundary conditions [31,32]. The valence quarks are generated with twisted boundary conditions that satisfy the relation

$$\psi(x_k + L) = e^{i\theta_k}\psi(x_k), \tag{2.3}$$

where the subindex k labels the spatial direction, $k = 1, 2, 3$, and L is the spatial extent of the lattice box in which we are simulating. The sea quarks, however, obey periodic boundary conditions. A meson with valence quarks generated with twisted boundary conditions as in Eq. (2.3) acquires a momentum $\vec{p}$ with components $p_k = \pi\Delta\theta_k/L$, where $\Delta\theta_k$ is the difference between the twisting angles of the two valence quarks. We can, thus, inject an arbitrary momentum to external mesons by tuning the twisting angles of their valence quarks. In particular, we can tune the external momentum to simulate directly at $q^2 \approx 0$, avoiding an interpolation in q^2 and the corresponding systematic uncertainty.

In our calculation, in order to get $q^2 \simeq 0$, we inject momentum into either the kaon or the pion. For a nonzero $\vec{p}_K$, we choose $\vec{\theta}_0 = \vec{\theta}_2 = 0$, $\vec{\theta}_1 \neq \vec{0}$, and for a nonzero $\vec{p}_\pi$, we choose $\vec{\theta}_0 = \vec{\theta}_1 = \vec{0}$, $\vec{\theta}_2 \neq \vec{0}$ (see Fig. 1 for definitions of $\vec{\theta}_{0,1,2}$). The twisting angles are, thus, fixed according to

$$\begin{aligned} |\vec{\theta}_1|_{q^2=0} &= \frac{L}{\pi}\sqrt{\left(\frac{m_K^2 + m_\pi^2}{2m_\pi}\right)^2 - m_K^2}, \\ |\vec{\theta}_2|_{q^2=0} &= \frac{L}{\pi}\sqrt{\left(\frac{m_K^2 + m_\pi^2}{2m_K}\right)^2 - m_\pi^2}, \end{aligned} \tag{2.4}$$

in each ensemble. The same choice of twisting angles was first explored for the calculation of $K \to \pi l\nu$ form factors in Ref. [33] (in the quenched approximation) and Ref. [34].

III. NUMERICAL SIMULATIONS

We have performed these calculations on the $N_f = 2 + 1$ MILC ensembles [35–37] with dynamical quarks simulated using the asqtad improved staggered action [19]. The up and down quarks in these configurations are degenerate and heavier than the physical ones. The strange-quark mass is tuned to values near the physical one. The gluon action is a one-loop Symanzik improved and tadpole improved action, with errors of $O(\alpha_s a^2)$ because the one-loop correction from fermion loops was not included in the coefficients—see Ref. [35] and references therein. The configurations on the $N_f = 2 + 1$ MILC ensembles were generated using the fourth-root procedure for eliminating extra degrees of freedom (tastes) originating from fermion doubling. Despite nonlocality and unitarity violations at nonzero lattice spacing [38,39], there are numerical [40–43] and theoretical [44–48] arguments that indicate that the correct theory of QCD is recovered in the continuum limit.

For the valence fermions, we use the HISQ action [18], which has much smaller taste-symmetry violations than the asqtad action. The mixed action setup, with a different light-quark formulation describing the sea and the valence quarks, increases the complexity of the analysis and could also increase some systematic errors, such as that from the chiral extrapolation. However, the effects of having a mixed action can be accounted for by using staggered chiral perturbation theory (SχPT), as described in Appendix A. Indeed, taste-breaking effects are the dominant source of discretization effects in our calculation, so reducing them, even just in the valence sector, is highly desirable.

We have performed simulations at two different values of the lattice spacing for the choice of parameters shown in Table I. The valence strange-quark mass is tuned to its physical value on each ensemble by matching m_{η_s} to its correct value [50,51]. The valence light-quark masses are fixed according to the relation

$$\frac{m_l^{\mathrm{val}}(\mathrm{HISQ})}{m_s^{\mathrm{phys.}}(\mathrm{HISQ})} = \frac{m_l^{\mathrm{sea}}(\mathrm{asqtad})}{m_s^{\mathrm{phys.}}(\mathrm{asqtad})}. \tag{3.1}$$

We average results over four time sources separated by 16 (24) time slices on the $a \approx 0.12$ fm (0.09 fm) ensembles but displaced by a random distance from configuration to configuration to mitigate autocorrelations. The exception is the most chiral ensemble, in which we doubled the statistics and, thus, have time sources separated by 8 time slices. We increased the statistics on this ensemble because we found getting stable results for the correlator fits to be more challenging. In order to disentangle the desired matrix elements from the contamination due to oscillating states in the correlation functions, we consider several source-sink separations (time separation T in Fig. 1), including both odd and even values, for each choice of quark masses and time source. We used the ensemble with $a \approx 0.12$ fm and sea quark masses 0.010/0.050 to tune the optimal parameters for the simulations on the rest of the ensembles, so the number of source-sink separations analyzed is larger than for the other ensembles [24].

For each ensemble, we generate zero-momentum two-point π and K correlation functions and two-point π and K correlation functions with momentum given by the twisting angles in Eq. (2.4). The structure of these correlators is

$$C_{\mathrm{2pt}}^{P}(\vec{p};t) = \frac{1}{L^3}\sum_{\vec{x}}\sum_{\vec{y}}\langle \Phi_P^{\vec{p}}(\vec{y}, t + t_{\mathrm{source}})\Phi_P^{\vec{p}\,\dagger}(\vec{x}, t_{\mathrm{source}})\rangle, \tag{3.2}$$

where $\Phi_P^{\vec{p}}(\vec{x}, t)$ is an interpolating operator creating a meson $P = \pi, K$ at time t carrying momentum $\vec{p}$ generated by using twisted boundary conditions. The $L^{-3}\sum_{\vec{x}}$ is implemented via random wall sources. We also generate three sets of three-point functions with $q^2 = q^2_{\mathrm{max}}$ (both pions and kaons at rest), $q^2 \simeq 0$ with kaons at rest and moving pions, and $q^2 \simeq 0$ with pions at rest and moving kaons. These correlators are obtained using

$$C_{\mathrm{3pt}}^{K\to\pi}(\vec{p}_\pi, \vec{p}_K; t, t_{\mathrm{source}}, T) = \frac{1}{L^3}\sum_{\vec{x}}\sum_{\vec{y}}\sum_{\vec{z}}\langle \Phi_K^{\vec{p}_K}(\vec{x}, t_{\mathrm{source}} + T) \times S(\vec{z}, t)\Phi_\pi^{\vec{p}_\pi\,\dagger}(\vec{y}, t_{\mathrm{source}})\rangle, \tag{3.3}$$

where either $\vec{p}_K$, $\vec{p}_\pi$ or both are zero. The scalar current is a taste-singlet current given by

$$S(\vec{z}, t) = \bar{\psi}_s(\vec{z}, t)\psi_q(\vec{z}, t), \tag{3.4}$$

where the ψ fields are naive fields.

The way in which the correlation functions above are expressed in terms of single-component staggered fields is explained in detail in Ref. [22]. The only difference between the correlation functions in that work and here is that we inject the external momenta using twisted boundary conditions.

TABLE I. Parameters of the ensembles analyzed in this work. The second and third columns show the approximate lattice spacing and the volume. am_l and am_h are the light and strange sea-quark masses, respectively. N_{conf} is the number of configurations analyzed from each ensemble, and am_l^{val} and am_s^{val} are the light and strange valence-quark masses. N_{sources} is the number of values of t_{source} used on each configuration. N_T is the number of source-sink separations T considered. The r_1/a values are obtained by fitting the calculated r_1/a to a smooth function [49], as explained in Ref. [35]. The $a \approx 0.12$ fm ensemble with volume $28^3 \times 64$ is used solely for the study of finite volume effects.

	$\approx a$ (fm)	$(\frac{L}{a})^3 \times \frac{L_t}{a}$	am_l/am_h	N_{conf}	am_s^{val}	am_l^{val}	N_{sources}	N_T	r_1/a	$m_\pi L$
Coarse	0.12	$20^3 \times 64$	0.020/0.050	2052	0.0491	0.02806	4	5	2.65	6.22
		$20^3 \times 64$	0.010/0.050	2243	0.0495	0.01414	4	8	2.62	4.48
		$28^3 \times 64$	0.010/0.050	275	0.0495	0.01414	4	4	2.62	6.27
		$20^3 \times 64$	0.007/0.050	2109	0.0491	0.00980	4	5	2.63	3.78
		$24^3 \times 64$	0.005/0.050	2098	0.0489	0.00670	8	5	2.64	3.84
Fine	0.09	$28^3 \times 96$	0.0124/0.031	1996	0.0337	0.0080	4	5	3.72	5.78
		$28^3 \times 96$	0.0062/0.031	1930	0.0336	0.0160	4	5	3.70	4.14

A. Fitting Method and Statistical Errors

We fit the two-point functions for a pseudoscalar meson P to the expression

$$C^P_{2\mathrm{pt}}(\vec{p}_P;t)=\sum_{m=0}^{N_{\mathrm{exp}}}(-1)^{m(t+1)}(Z^P_m)^2(e^{-E^m_P t}+e^{-E^m_P(L_t-t)}), \tag{3.5}$$

where L_t is the temporal size of the lattice. Oscillating terms with $(-1)^{m(t+1)}$ do not appear for pions with zero momentum. In Table II, we summarize the masses of the valence (Goldstone) pions and kaons that determine the twisting angles and, thus, the external momentum $\vec{p}_P$ injected in the three-point functions to get $q^2=0$, together with these twisting angles and momenta for each ensemble. We also include in Table II the sea Goldstone and root-mean-squared (RMS) pion masses from Ref. [6].

From two-point function fits, we checked whether the continuum dispersion relation is satisfied. This is plotted in Fig. 2, which shows very small deviations from the continuum prediction ($\leq 0.25\%$), indicating small discretization effects. These deviations do not exceed power counting estimates of $O(\alpha_s(a|\vec{p}|)^2)$ errors, except for the point corresponding to a moving pion with $|r_1 p_\pi|^2\approx 0.07$ and $a=0.09$ fm, for which the power counting estimate is about 0.14%. The data corresponding to a moving kaon on the same ensemble (the rightmost red (speckled) triangle in Fig. 2), however, do not show any unusual behavior. In addition, the results for $f_+(0)$ as extracted from data with moving pions and moving kaons agree within one statistical σ ($\leq 0.15\%$ error). This gives us confidence that the errors quoted below for $f_+(0)$ properly take into account the discretization effects seen in the dispersion relation in Fig. 2.

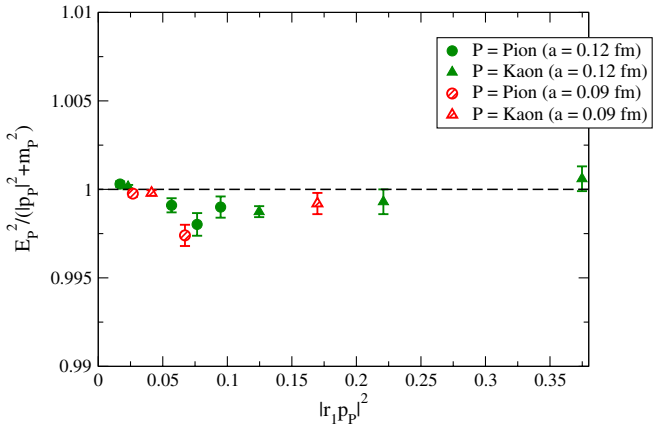

FIG. 2 (color online). Ratio of the measured (lattice) and the continuum dispersion relation for all data analyzed.

The functional form for the three-point functions is

$$\begin{aligned}C^{K\to\pi}_{3\mathrm{pt}}(\vec{p}_\pi,\vec{p}_K;t,T)=&\sum_{m,n=0}^{N^{3\mathrm{pt}}_{\mathrm{exp}}}(-1)^{m(t+1)}(-1)^{n(T-t+1)}\\&\times A^{mn}(q^2)Z^\pi_m Z^K_n(e^{-E^m_\pi t}+e^{-E^m_\pi(L_t-t)})\\&\times(e^{-E^n_K(T-t)}+e^{-E^n_K(T-L_t+t)}),\end{aligned} \tag{3.6}$$

where the factors Z^P_i are the amplitudes of the two-point functions in Eq. (3.5). The three-point parameter $A^{00}(q^2)$ in Eq. (3.6) is related to the desired form factor $f_0(q^2)$ via

$$f_0(q^2)=\frac{1}{2}A^{00}(q^2)\sqrt{2E_\pi E_K}\,\frac{m_s-m_q}{m_K^2-m_\pi^2}, \tag{3.7}$$

where we have used Eq. (2.2) and taken into account overall factors involved in the parametrization of the correlation function. We extract the form factors $f_0(q^2)$, using the expression above, directly from simultaneous Bayesian

TABLE II. Sea Goldstone and RMS pion masses, valence Goldstone pion and kaon masses, twisting angles, and external momenta injected in the three-point functions for each ensemble. The quark masses am_l and am_h are the same as in Table I, and the twisting angles $\vec{\theta}_1$ and $\vec{\theta}_2$ are defined in Fig. 1 and Eq. (2.4). The subscript P in the meson masses refers to the pseudoscalar taste, as defined in Eq. (A1). The first line for each ensemble corresponds to moving pions and the second line to moving kaons.

$\approx a$ (fm)	am_l/am_h	$aM^{\mathrm{sea}}_{\pi,\mathrm{P}}$	$aM^{\mathrm{sea}}_{\pi,\mathrm{RMS}}$	$aM^{\mathrm{val}}_{\pi,\mathrm{P}}$	$aM^{\mathrm{val}}_{K,\mathrm{P}}$	$\lvert\vec{\theta}_1\rvert$	$\lvert\vec{\theta}_2\rvert$	$\lvert a\vec{p}_P\rvert$
0.12	0.020/0.050	0.31125	0.3728	0.31315	0.36617	0	0.31310	0.04918
						0.36610	0	0.05751
	0.010/0.050	0.22447	0.3041	0.22587	0.33456	0	0.57960	0.09104
						0.85851	0	0.13485
	0.007/0.050	0.18891	0.2789	0.18907	0.32119	0	0.66807	0.10494
						1.13486	0	0.17826
	0.005/0.050	0.15971	0.2600	0.15657	0.31225	0	0.88804	0.11624
						1.76565	0	0.23112
0.09	0.0124/0.031	0.20635	0.2247	0.20341	0.25241	0	0.39257	0.04405
						0.48667	0	0.05460
	0.0062/0.031	0.14789	0.1726	0.14572	0.23171	0	0.62420	0.07003
						0.99253	0	0.11136

fits of the relevant three- and two-point functions. In these fits, we include several three-point functions with different values of the source-sink separation T, with at least one odd T and one even T to control the contributions from the oscillating states. We perform three sets of fits: including only the data corresponding to moving pions, including only the data corresponding to moving kaons, or including all data. In the last case, we assume that the parameters $A^{mn}(q^2)$ depend only on q^2, so they are the same for moving pions or moving kaons. This is true up to discretization errors of the order of the deviations from the continuum dispersion relation of those data ($\leq 0.25\%$). The quality of all the fits included in our determination of the form factor is rather high, with p values ranging between 0.6 and 1. This may indicate that there is some mild overestimation of statistical errors.

In this analysis, it is especially relevant to check for the stability of our fits under the choice of fitting parameters and techniques, since we have very small statistical errors and need to be sure that these results are not dependent on our methodology. One check is to vary the time fitting ranges and number of states included in the fits. The fitting range is $t \in [t_{\mathrm{min}}, (L_t - t_{\mathrm{min}})]$ for two-point functions and $t \in [t_{\mathrm{min}}, (T - t_{\mathrm{min}})]$ for three-point functions—see Fig. 1 and Table I for notation. The number of states included is the same in the regular and oscillating sectors, so $N_{\mathrm{exp}}/2 = N_{\mathrm{regular\ states}} = N_{\mathrm{oscillating\ states}}$. We also keep the number of states the same in all correlation functions included in a given simultaneous fit. Fixing N_{exp} and changing t_{min} from 3 (5) for $a \approx 0.12$ fm (0.09 fm) ensembles up to the maximum allowed by the source-sink separation gives us a plateau for the central values with only small variations in the errors. Analogously, fixing t_{min} to our preferred value, we do not find any significant variation of the results for $N_{\mathrm{exp}} \geq 6$–8. Examples of these variation tests for two of the ensembles analyzed are shown in Fig. 3.

We also study which combination of T's is optimal. We find that the central values are very insensitive to the number of three-point functions included and the values of T in the range we are analyzing. Errors and stability are better when $15 \leq T \leq 24$ with three values of T in the three-point functions for the $a \approx 0.12$ fm ensembles and $18 \leq T \leq 33$ with four values of T in the three-point functions for the $a \approx 0.09$ fm ensembles.

Finally, we check an alternative fit procedure, using the superaverage method described in Ref. [52]. This takes an explicit combination of three-point functions with consecutive values of T and the time slice t that suppresses the contribution from both the first regular excited state and the first oscillating state:

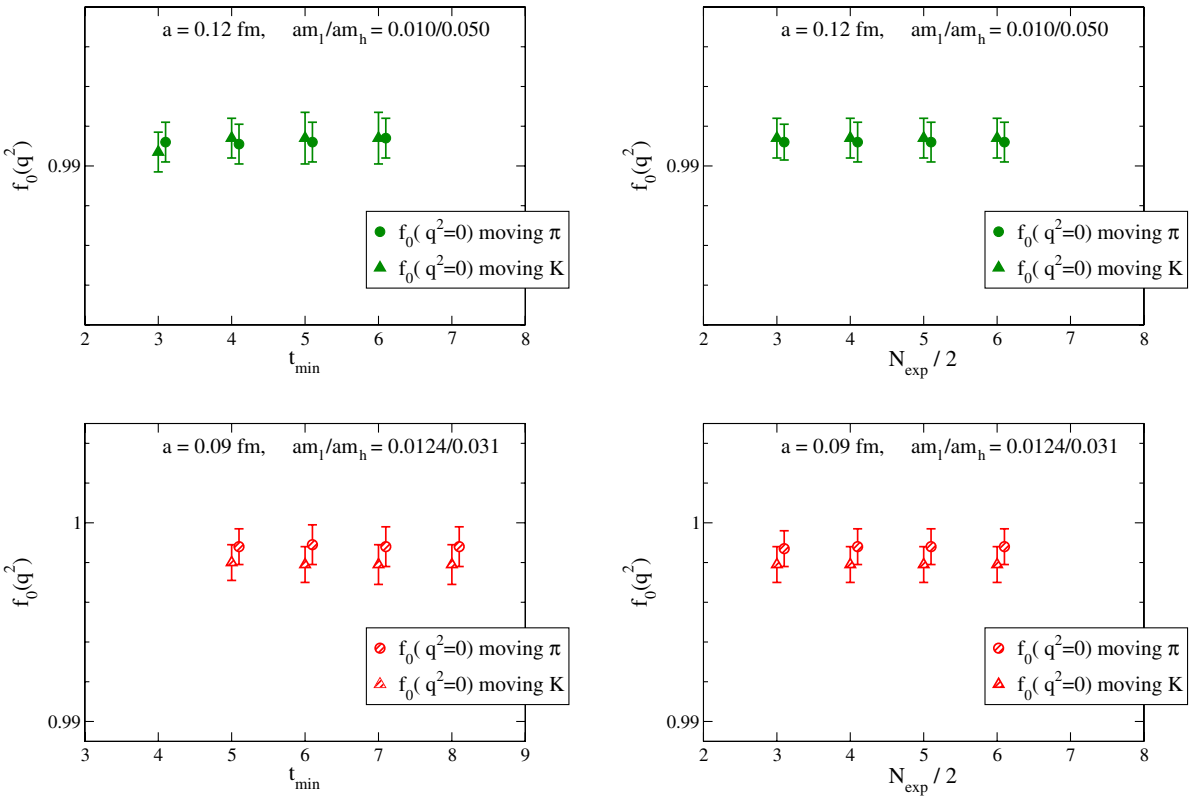

FIG. 3 (color online). Variation of $f_0(q^2)$ with t_{min} (left panel) and N_{exp} (right panel) for $q^2 = 0$ generated by injecting momentum in the π or the K and for $q^2 = q^2_{\mathrm{max}}$ (zero external momentum) for two of the ensembles analyzed.

$$\bar{C}^{K\to\pi}_{3\mathrm{pt}}(t,T) = \frac{e^{-E^{(0)}_\pi t}e^{-E^{(0)}_K(T-t)}}{8}\Bigg[\frac{C^{K\to\pi}_{3\mathrm{pt}}(t,T)}{e^{-E^{(0)}_\pi t}e^{-E^{(0)}_K(T-t)}} + \frac{C^{K\to\pi}_{3\mathrm{pt}}(t,T+1)}{e^{-E^{(0)}_\pi (t)}e^{-E^{(0)}_K(T+1-t)}} + \frac{2C^{K\to\pi}_{3\mathrm{pt}}(t+1,T)}{e^{-E^{(0)}_\pi (t+1)}e^{-E^{(0)}_K(T-t-1)}}$$
$$+ \frac{2C^{K\to\pi}_{3\mathrm{pt}}(t+1,T+1)}{e^{-E^{(0)}_\pi (t+1)}e^{-E^{(0)}_K(T-t)}} + \frac{C^{K\to\pi}_{3\mathrm{pt}}(t+2,T)}{e^{-E^{(0)}_\pi (t+2)}e^{-E^{(0)}_K(T-t-2)}} + \frac{C^{K\to\pi}_{3\mathrm{pt}}(t+2,T+1)}{e^{-E^{(0)}_\pi (t+2)}e^{-E^{(0)}_K(T-t-1)}}\Bigg]$$
$$\approx A^{00}\sqrt{Z^K_0 Z^\pi_0}\,e^{-E^{(0)}_\pi t}e^{-E^{(0)}_K(T-t)} + (-1)^T\bar{A}^{11}e^{-E^{(1)}_\pi t}e^{-E^{(1)}_K(T-t)} + \bar{A}^{20}e^{-E^{(2)}_K(T-t)}e^{-E^{(0)}_\pi t}$$
$$+ \bar{A}^{02}e^{-E^{(2)}_\pi t}e^{-E^{(0)}_K(T-t)} + O(\Delta E^2_\pi, \Delta E_\pi \Delta E_K, \Delta E^2_K). \quad (3.8)$$

Again, results are compatible with our preferred fitting method within errors.

We also checked for autocorrelations in our data and did not observe any effect on the relevant correlation functions. Central values and errors of the ground-state parameters remain unchanged when blocking by 2, 3, and 4. We, thus, do not perform any blocking of our data.

In Fig. 4, we collect the results for $f_+(0)$ from our preferred correlator fit methodology as a function of $am_l/am_s^{\mathrm{phys}}$. The errors shown are statistical only and generated with a 500 bootstrap distribution. They are very small, around 0.1–0.15%, except for the cases noted below. We plot results coming from fits to three-point functions with moving pions (circles) and moving kaons (upward-pointing triangles), as well as combined fits (downward-pointing triangles) of both sets of three-point functions. The correlator fits for the 0.005/0.050 and 0.007/0.050 ensembles with $a \approx 0.12$ fm and moving kaons are not as stable as for the other ensembles, and errors are considerably larger. These results correspond to the two first filled green upward-pointing triangles from the left in Fig. 4. This is due to the fact that we need a larger momentum in lattice units to get $q^2 = 0$ when reducing the light-quark masses, and correlators become noisier. For this reason, we drop these two sets of data from our analysis.

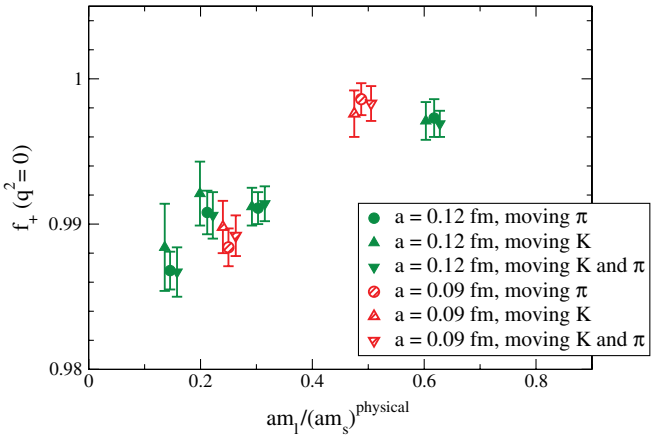

FIG. 4 (color online). Form factor $f_+(0)$ obtained from the different ensembles in Table I as a function of the valence light-quark mass normalized to the physical strange-quark mass. The squares, upward-pointing triangles, and downward-pointing triangles correspond to fits to data with only moving pions, only moving kaons, and both moving pions and moving kaons, respectively. The filled green (speckled red) symbols label ensembles with $a \approx 0.12$ fm (0.09 fm). We give a horizontal offset to two of the three points at each light-quark mass for clarity.

The results from fits of three-point functions with moving K, moving π, or with both sets of correlators all agree within our very small statistical errors, as can be seen in Fig. 4. This constitutes a strong test of our methodology and quoted errors. It also suggests that the discretization errors that break the continuum dispersion relation (which contribute differently for $f_+(0)$ as extracted from moving pions and moving kaons) are smaller than our statistical errors.

For the rest of the analysis, we take separately the data with moving pions and moving kaons, omitting the two ensembles with moving kaons mentioned above. This choice allows us to identify more clearly the different sources of discretization errors. The numerical values of the data included in this analysis are listed in Table III.

IV. EXTRAPOLATION TO THE PHYSICAL POINT

In this section, we describe the use of chiral perturbation theory (χPT) to extrapolate our data at several values of the light quark masses down to the physical point. The form factor $f_+(0)$ can be written as a χPT expansion in the following way:

$$f_+(0) = 1 + f_2 + f_4 + f_6 + \cdots = 1 + f_2 + \Delta f, \quad (4.1)$$

where the f_{2i}'s contain corrections of $O(p^{2i})$ in the chiral power counting. The Ademollo-Gatto (AG) theorem [53], which follows from vector current conservation, ensures

TABLE III. Values of $f_+(0)$ included in the chiral and continuum extrapolation. Errors are statistical only, from 500 bootstrap ensembles.

$\approx a$ (fm)	am_l/am_h	$aM^{\mathrm{val}}_{\pi,\mathrm{P}}$	$f_+^{\mathrm{moving}\ \pi}(0)$	$f_+^{\mathrm{moving\ K}}(0)$
0.12	0.020/0.050	0.31315	0.9973(13)	0.9971(13)
	0.010/0.050	0.22587	0.9911(11)	0.9912(13)
	0.007/0.050	0.18907	0.9908(15)	-
	0.005/0.050	0.15657	0.9868(13)	-
0.09	0.0124/0.031	0.20341	0.9984(13)	0.9898(18)
	0.0062/0.031	0.14572	0.9986(11)	0.9976(16)

that $f_+(0) \to 1$ in the SU(3) limit and, further, that the SU(3) breaking effects are second order in $(m_K^2 - m_\pi^2)$. This fixes f_2 completely in terms of experimentally measurable quantities. At finite lattice spacing, however, we have violations of the AG theorem due to symmetry-breaking discretization effects in the form-factor decomposition, Eq. (1.2), and in the continuum dispersion relation needed to derive the relation between $f_0(0)$ and the correlation functions, Eq. (2.2). These and other discretization effects are very small in our data, though, as can be deduced from Fig. 4.

We study the light-quark mass dependence and the discretization effects in our calculation using two-loop continuum χPT [54], supplemented by partially quenched staggered χPT (PQSχPT) at one loop. The small variation with a in our data suggests that addressing those effects at one loop should be enough for our target precision. We also incorporate partial quenching effects in the continuum expression at one loop [55,56]. The staggered one-loop fitting function can be found in Appendix A, while the details of its calculation will be given in Ref. [57].

In order to test the dependence of our results on the fit function, as well as the systematics associated with disregarding higher-order terms in both the chiral and the a^2 expansions, we also use another approach for the chiral and continuum extrapolation. We replace the continuum two-loop χPT with a general analytic next-to-next-to-leading-order (NNLO) parametrization. The implementation and results obtained with each of these methods are discussed in the following sections.

A. Partially quenched SχPT at NLO plus continuum NNLO χPT

We take the next-to-leading-order (NLO) PQSχPT expression in Appendix A and add the two-loop full QCD continuum calculation from Ref. [54]. The one-loop corrections that relate the pion decay constant f_π to the SU(3) chiral limit, f_0, must be accounted for consistently when f_π appears in the one-loop expression, and we are working through two-loop order. Our two-loop expression is chosen so that f_π evaluated at the physical pion mass is the appropriate value to use in the coefficient of the one-loop chiral logarithms. We also add a term that parametrizes discretization effects arising from the violation of the continuum dispersion relation, which breaks the AG theorem, and another term of $O(a^2)$ that respects the AG theorem. They are proportional to $K_1^{(a)}$ and $K_2^{(a)}$, respectively, in Eq. (4.2). This gives us the general fit function

$$f_+(0) = 1 + f_2^{\mathrm{PQS}\chi\mathrm{PT}}(a) + K_1^{(a)}\left(\frac{a}{r_1}\right)^2 + f_4^{\mathrm{cont.}}(\mathrm{logs}) + f_4^{\mathrm{cont.}}(L_i) + r_1^4(m_\pi^2 - m_K^2)^2\left[C_6^{\prime(1)} + K_2^{(a)}\left(\frac{a}{r_1}\right)^2\right], \tag{4.2}$$

where the constants $K_1^{(a)}$, $K_2^{(a)}$, and $C_6^{\prime(1)} \propto C_{12} + C_{34} - L_5^2$ are constrained parameters to be determined by the chiral fits using Bayesian techniques. All dimensionful quantities entering in the chiral fit functions are converted to r_1 units using the values of r_1/a in Table I. We split the two-loop contribution, f_4, into a piece containing the logarithmic terms alone and another one depending on the $O(p^4)$ low-energy constants (LECs) L_i. The first piece does not contain any free parameters. For the NLO LECs, we can either choose to fix them or leave them as constrained parameters in the fits. This choice does not make much difference in our final results, though, as explained in Sec. IV C. The C_{ij} are $O(p^6)$ LECs defined in Ref. [58]. Notice that only one combination of $O(p^6)$ LECs contributes to the chiral expansion for $f_+(0)$ at this order, so only one extra free parameter enters at NNLO.

The NLO PQSχPT function, $f_2^{\mathrm{PQS}\chi\mathrm{PT}}(a)$, incorporates the dominant lattice artifacts from taste-symmetry breaking. The only unknown quantities entering in this function are the "mixed" taste-violating hairpin parameters, δ_V^{mix} and δ_A^{mix}, defined in Appendix A. These coefficients appear because we use a mixed action, and we expect their values to be small. Thus, we take them as constrained parameters in the fit, with prior central values equal to zero and prior widths equal to or larger than the hairpin parameters in the

TABLE IV. Inputs for the fixed parameters needed in the NLO PQSχPT part of the fit functions in Eqs. (4.2) and (4.3). We do not consider errors on the slope μr_1, the pion decay constant $f_\pi r_1$, or the taste splittings $r_1^2 a^2 \Delta_\Xi$ because they have negligible effect on the final results.

	$a \approx 0.12$ fm	$a \approx 0.09$ fm
$f_\pi r_1$	0.20725	
$\Lambda_\chi r_1 = M_\rho r_1$	1.2225	
μr_1 (asqtad)	6.234	6.382
$r_1^2 a^2 \Delta_P^{\mathrm{asqtad}}$	0	0
$r_1^2 a^2 \Delta_V^{\mathrm{asqtad}}$	0.439	0.152
$r_1^2 a^2 \Delta_T^{\mathrm{asqtad}}$	0.327	0.115
$r_1^2 a^2 \Delta_A^{\mathrm{asqtad}}$	0.205	0.0706
$r_1^2 a^2 \Delta_I^{\mathrm{asqtad}}$	0.537	0.206
$r_1^2 a^2 \Delta_P^{\mathrm{HISQ}}$	0	0
$r_1^2 a^2 \Delta_V^{\mathrm{HISQ}}$	0.171	0.0573
$r_1^2 a^2 \Delta_T^{\mathrm{HISQ}}$	0.112	0.0390
$r_1^2 a^2 \Delta_A^{\mathrm{HISQ}}$	0.0575	0.0204
$r_1^2 a^2 \Delta_I^{\mathrm{HISQ}}$	0.222	0.0709
$r_1^2 a^2 \Delta_P^{\mathrm{mixed}}$	0.0828	0.0828×0.35
$r_1^2 a^2 \Delta_V^{\mathrm{mixed}}$	0.325	0.325×0.35
$r_1^2 a^2 \Delta_T^{\mathrm{mixed}}$	0.263	0.263×0.35
$r_1^2 a^2 \Delta_A^{\mathrm{mixed}}$	0.174	0.174×0.35
$r_1^2 a^2 \Delta_I^{\mathrm{mixed}}$	0.403	0.403×0.35

valence sector, δ_V^{HISQ} and δ_A^{HISQ}. The remaining inputs needed to obtain $f_2^{\mathrm{PQS}\chi\mathrm{PT}}(a)$ have been calculated by our collaboration or can be taken from experiment. Their values are collected in Table IV, and their definitions appear in Appendix A. The asqtad taste splittings are taken from Ref. [6], while the HISQ splittings are preliminary results from the MILC Collaboration that agree well with their final results in Ref. [59]. We extract the mixed taste splittings from simulations on the $a \approx 0.12$ fm ensemble with sea quark masses 0.010/0.050 and valence light-quark mass $am_l = 0.2am_s$. With the exception of the pseudoscalar taste splitting, which is nonzero for a mixed action, the other mixed taste splittings agree fairly well with the values given by the approximation used in Ref. [25]: $\Delta_\Xi^{\mathrm{mix}} = (\Delta_\Xi^{\mathrm{asqtad}} + \Delta_\Xi^{\mathrm{HISQ}})/2$. However, we observe a systematic effect that the real taste splittings are between 10% and 30% larger than those from this approximation. Since the fits for the tensor taste are not stable enough to extract the corresponding taste splitting, we use this effect as a guide and estimate Δ_T^{mix} to be 20% larger than the sea-valence average. The final result is insensitive to this choice, however. Finally, the taste-violating hairpin parameters for the HISQ action, collected also in Table IV, are taken from chiral fits to the HISQ light pseudoscalar data [60].

B. Partially quenched SχPT at NLO plus NNLO analytical parametrization

To estimate systematic errors, we use a second approach. We take the same NLO staggered partially quenched χPT expressions as before, but, instead of the two-loop continuum calculation, we add a general analytic parametrization of higher-order terms and a^2 corrections of the form

$$f_+(0) = 1 + f_2^{\mathrm{PQS}\chi\mathrm{PT}}(a) + K_1^{(a)}\left(\frac{a}{r_1}\right)^2 + r_1^4(m_\pi^2 - m_K^2)^2 \times\Bigg[C_6^{(1)} + C_8^{(1)}(r_1 m_\pi)^2 + C_8^{(2)}(r_1 m_\pi)^2 \times\ln(m_\pi^2/\Lambda_\chi^2) + C_{10}^{(1)}(r_1 m_\pi)^4 + K_2^{(a)}\left(\frac{a}{r_1}\right)^2\Bigg], \quad (4.3)$$

where Λ_χ is the chiral scale, which we take equal to the mass of the ρ throughout this analysis. The discretization effects in Eq. (4.3) are parametrized in the same way as in the fit function in Eq. (4.2). The terms proportional to the coefficients $C_{2j}^{(i)}$ are of $O(p^{2j})$ in the chiral expansion, so $C_6^{(1)}$ is NNLO as before, $C_8^{(i)}$ are next-to-next-to-next-to-leading order (NNNLO), and $C_{10}^{(1)}$ is next-to-next-to-next-to-next-to-leading order (NNNNLO). Notice that we do not include a term of order $(r_1 m_K)^2$ in the expression above since the valence strange-quark masses are tuned to the physical value and are, thus, the same in units of r_1. The effect produced by the difference between the strange-quark mass in the valence and the sea sectors is discussed in Sec. VA.

C. Results

We obtain the central value for $f_+(0)$ by extrapolating our data to the physical meson masses and the continuum limit using the fitting functional form in Eq. (4.2) with $K_2^{(a)} = 0$ and allowing for a nonzero value of $K_1^{(a)}$. We use Bayesian techniques and estimate the statistical errors by fitting to a set of 500 bootstrap samples for each ensemble. In these fits, we include results coming from only injecting external momentum in the π and only injecting external momentum in the K, circles and upward-pointing triangles in Fig. 4, respectively. We omit the data with moving kaons in the two most chiral $a \approx 0.12$ fm ensembles, due to the lower quality of those data, as explained in Sec. . For the meson masses, we use the Particle Data Group values, $m_{\pi^-} = 139.570$ MeV and $m_{K^0} = 497.614$ MeV.

The prior central values and widths we use in our fits for the constrained parameters are summarized in Table V. For the priors for $C_6^{\prime(1)}$, we assume that the NNLO analytical

TABLE V. Priors for the fit parameters entering the expressions in Eqs. (4.2) and (4.3). The dimensionless χPT parameter s is given by the quantity $1/(8\pi^2(r_1 f_\pi)^2)$. See the text for the explanation of the choice of priors. The priors listed for the hairpin parameters are for the $a \approx 0.12$ fm ensembles, and those for the $a \approx 0.09$ fm ensembles are obtained by multiplying by 0.35. This factor comes from assuming that the hairpin parameters scale like the Δ_Ξ. The central values for the NLO LECs are from fit 10 in Ref. [61].

Fit parameters	(Central value ± width)
$r_1^2 a^2 \delta_V^{\mathrm{HISQ}}$	0.057 ± 0.033
$r_1^2 a^2 \delta_A^{\mathrm{HISQ}}$	−0.0782 ± 0.0040
$r_1^2 a^2 \delta_V^{\mathrm{mix}}$	0.0 ± 0.1
$r_1^2 a^2 \delta_A^{\mathrm{mix}}$	0.0 ± 0.1
$K_{1,2}^{(a)}$	0 ± 1
$C_6^{\prime(1)}$	0 ± s^2
$C_6^{(1)}$	0 ± s^2
$C_8^{(i)}$	0 ± s^3
$C_{10}^{(1)}$	0 ± s^4
$L_1^r(M_\rho) \times 10^3$	0.43 ± 0.24
$L_2^r(M_\rho) \times 10^3$	0.73 ± 0.24
$L_3^r(M_\rho) \times 10^3$	−2.30 ± 0.74
$L_4^r(M_\rho) \times 10^3$	0.0 ± 0.6
$L_5^r(M_\rho) \times 10^3$	0.97 ± 0.22
$(2L_6^r - L_4^r)(M_\rho) \times 10^3$	0.0 ± 0.4
$L_7^r(M_\rho) \times 10^3$	−0.31 ± 0.28
$L_8^r(M_\rho) \times 10^3$	0.60 ± 0.36

terms are of the same order as the NNLO logarithmic terms, which are $\sim s^2(r_1 m_\pi)^4$ with s a typical χPT scale, $s \equiv 1/(8\pi^2(r_1 f_\pi)^2)$. Following the same reasoning, we choose the priors' widths to be s^3 and s^4 for the coefficients of the NNNLO and NNNNLO terms, respectively. We set the priors for the coefficients of the a^2 terms to 0 ± 1 and check that the output of the fits for those coefficients is not constrained by that choice. For example, in our main fit, the output for $K_1^{(a)}$ is $-0.011(9)$, well under the limits set by the corresponding prior.

We treat the HISQ hairpin parameters as constrained parameters in our fits, with prior central values and widths equal to the values and errors obtained in Ref. [60] from fits to other light quantities. Those priors are also listed in Table V. In this way, the uncertainty associated with the error of those parameters is included directly in the statistical error for the extrapolated form factor. The same strategy is adopted for the NLO LECs needed in the fit function, L_i with $i = 1$–5. We constrain their values using the results from fit 10 in Refs. [61,62] for the prior central values and 2 times the error in the same reference for the prior widths; see Table V. Since Ref. [61] sets L_4 and L_6 exactly to zero, in accord with large N_c expectations [64], for the width of L_4 and $2L_6 - L_4$, we use the error in the determination by the MILC Collaboration in Ref. [7]. The exact choice of NLO LECs priors has a minimal impact on our final result. We have checked several equally reasonable choices for treating these parameters, including using the errors in Ref. [61] as prior widths instead of widths 2 times larger; using widths 10 times larger; fixing the LEC's (instead of leaving them as parameters of the fit) to the values in Ref. [61]; fixing $L_{4,5,6}$ instead to the values determined in Ref. [7], where different lattice light quantities were computed on the same ensembles as here; shifting the values of each individual L_i within its corresponding error as determined in Ref. [61]; and keeping $L_4 = 0$ according to its large N_c limit. The conclusion from these checks is that our data are unable to disentangle the individual values of the L_i's, but the value of the form factor in the continuum limit and at the physical quark masses changes by less than 0.03% with any of the choices listed above.

The results from our preferred fitting strategy are shown in Fig. 5. The form factor in the continuum limit and at the physical quark masses is

$$f_+(0) = 0.9667 \pm 0.0023 \pm \text{systematics}, \qquad (4.4)$$

where the error, for now, is statistical plus the uncertainty from the errors of the hairpin parameters $\delta_{V,A}^{\text{HISQ}}$ and the NLO LECs. The p value of the fit leading to the form factor in Eq. (4.4) and for which results are displayed in Fig. 5 is 0.62. We note that approximately 15–25% of the apparent discretization effects in the 0.12 fm and 0.09 fm extrapolations to the physical point (green and red lines in Fig. 5) are actually due to mistunings of the sea strange-quark mass.

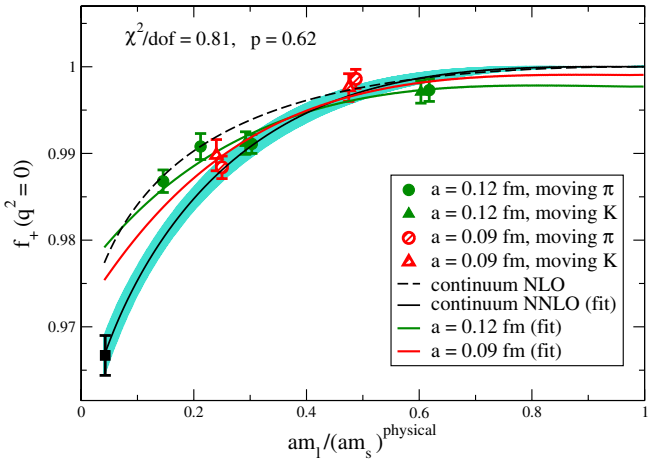

FIG. 5 (color online). Form factor $f_+(0)$ vs light-quark mass. Filled green (speckled red) symbols are the $a \approx 0.12$ fm ($a \approx 0.09$ fm) data points included in the chiral fit. The dashed black line is the continuum NLO χPT prediction, $1 + f_2$. The solid black line is the extrapolation in the light-quark mass, keeping m_s equal to its physical value and turning off all discretization effects. The turquoise band shows the error for the continuum line. The green and red lines are also extrapolations in the light-quark masses with $a \approx 0.12$ fm and 0.09 fm, respectively. Errors in the data points are statistical only and were obtained from 500 bootstrap samples. The differences between the green, red, and solid black lines reflect not only the discretization errors but also the effect of the mistuned sea quark masses, which are particularly large in the $a \approx 0.12$ fm case. The error in the continuum value at the physical quark masses includes also the uncertainty in the $O(p^4)$ LECs and the HISQ hairpin parameters.

V. SYSTEMATIC ERROR ANALYSIS

In this section, we discuss the systematic errors in the calculation of $f_+(0)$ and the different strategies followed to estimate them. The uncertainties discussed in this section are summarized in Table. VI.

A. Chiral extrapolation and fit function

The fit function is completely determined at NLO in χPT, given the values of the taste-breaking parameters. In particular, the value of the decay constant that enters

TABLE VI. Complete error budget and total error for $f_+(0)$. All errors are given in percent.

Source of uncertainty	Error $f_+(0)$ (%)
Statistics	0.24
Chiral extrapolation & fitting	0.3
Discretization	0.1
Scale	0.06
Finite volume	0.1
Total Error	0.42

in $f_2^{\mathrm{PQ}\chi\mathrm{PT}}(a)$, the experimental value for f_π, is fixed by the two-loop formulas employed in our analysis. Any other choice for this parameter requires the use of a different NNLO expression. Hence, we can estimate the uncertainty associated with the choice of fit function and the truncation of the χPT series by exploring different choices we make for the NNLO description of the form factor.

First, we perform a set of chiral fits using the fit function in Eq. (4.3), in which we describe NNLO and higher-order chiral effects with an analytical parametrization. In order to compare with our result in Eq. (4.4) and disentangle discretization effects from the chiral extrapolation error, we parametrize the discretization effects in Eq. (4.3) in the same way as for our preferred fit function using Eq. (4.2): we take $K_2^{(a)} = 0$ and $K_1^{(a)} \neq 0$. We perform several fits in which we include the NNLO term $C_6^{(1)}$ along with one or two of the other terms in Eq. (4.3).

The fit function equivalent to the one we use in our preferred fit, namely, the pure NNLO expression, which includes terms proportional to $K_1^{(a)}$ and $C_6^{(1)}$ in Eq. (4.3), and choosing the physical value for the decay constant f_π, agrees with our central value in Eq. (4.4) within statistical errors. The difference between the determinations is $\sim$0.15%, although using the two-loop χPT gives a better fit. If one includes the NNNLO term proportional to $(r_1 m_\pi)^2$ in Eq. (4.3), the quality of the fit improves, and the result for $f_+(0)$ is nearly identical to Eq. (4.4). Including the logarithmic term proportional to $C_8^{(2)}$, instead of the term proportional to $C_8^{(1)}$, gives a value of $f_+(0)$ $\sim$0.2% lower than the one in Eq. (4.4) but with a worse confidence level and much larger errors. Going beyond the inclusion of two chiral correction terms makes the fits unstable and considerably increases the statistical errors of the extrapolated point. With our data, we are not able to determine more than three coefficients, including the one proportional to the lattice spacing squared, $K_1^{(a)}$.

Another issue related to the use of the analytical parametrization in Eq. (4.3) is the freedom to choose the value of the decay constant used at NLO, i.e., in $f_2^{\mathrm{PQ}\chi\mathrm{PT}}$ (with the two-loop calculation, the freedom appears at NNLO). In our preferred NNLO χPT analysis, using f_K instead of f_π changes the central value by a negligible 0.02%, while using $f_0 = (113.6 \pm 3.6 \pm 7.7)$ MeV [7] instead of f_π increases the result only by 0.2%.

The last ambiguity in our fit function is related to the fact that we have strange quarks simulated with different masses in the valence and sea sectors. We are using the physical values for the valence strange-quark masses on each ensemble, but the sea strange-quark masses have values that differ from their physical values by a factor of $\sim$1.5 on the $a \approx 0.12$ fm ensembles and $\sim$1.2 on the ≈ 0.09 fm ensembles. We correct for this mistuning using PQSχPT at one loop, but we cannot correct at the two-loop level since the partially quenched χPT expressions are not available. We obtain the value in Eq. (4.4) using the valence meson masses in the two-loop χPT function f_4. If we use the sea meson masses instead, keeping the factor $(m_K^2 - m_\pi^2)^2$ from the valence sector, the extrapolated result for $f_+(0)$ decreases by 0.3%. At NLO, the difference between the correct $1 + f_2^{\mathrm{PQS}\chi\mathrm{PT}}$ and the one taking all the masses equal to the valence masses is 0.2% on the coarse ensembles and 0.07% on the fine ensembles, and the NLO contribution is around 3 times larger than the NNLO. Thus, the 0.3% shift most likely overestimates the effect of the mistuning of m_s in the sea.

Finally, the addition of a NNNLO term of the form $(m_\pi^2 - m_K^2)^2 \times m_\pi^2$ to our fit function in Eq. (4.2) results in a negligible shift in the extrapolated result for $f_+(0)$. In summary, all the effects described above provide estimates of the error associated with our choice of fit function. We take the largest shift in the central value, 0.3%, as the systematic error.

B. Discretization errors

The dominant sources of discretization errors in our calculation are taste-violating and finite-momentum effects. The leading-order contribution from the first source is removed by using PQSχPT at one loop. The remaining taste-breaking errors are of $O(a^2\alpha_s^3\Lambda_{\mathrm{QCD}}^2, a^4\Lambda_{\mathrm{QCD}}^4)$. The latter source yields errors of $O(\alpha_s(ap)^2)$.

Since we have results only at two different values of the lattice spacing, we can reliably determine the value of one fit term that depends on the lattice spacing. For our central value, we choose the simplest parametrization of discretization effects, a term given by $(a/r_1)^2$ times a constant $K_1^{(a)}$—see Eq. (4.3). We choose to allow this term to break the AG theorem since we expect errors of $O((ap)^2)$ to be larger than $O(a^2\alpha_s^2\Lambda_{\mathrm{QCD}}^2, a^4\Lambda_{\mathrm{QCD}}^4)$ errors. The value obtained from the fit for the a^2 coefficient, $K_1^{(a)} = -0.011(9)$, suggests that discretization errors after removing the dominant taste-breaking effects are very small. Indeed, our data are well described by a fit that includes only the one-loop PQSχPT terms and omits the $(a/r_1)^2$ term. With $K_1^{(a)} = 0$, we obtain a result that agrees with Eq. (4.4) at the 0.05% level for both the central value and the error. However, we do need to include taste-breaking effects at one loop, since the statistical errors in our data are very small, 0.1–0.15%.

We conclude that the remaining $O(a^2)$ errors after the continuum extrapolation are negligible. Nevertheless, we try alternate parametrizations of the generic lattice-spacing dependence to further test the size of the remaining discretization effects. We perform fits in which we replace the $K_1^{(a)}$ term by 1) the $K_2^{(a)}$ term in Eq. (4.2), 2) a term of order $(a/r_1)^2(m_K^2 - m_\pi^2)$, 3) a term proportional to $(a/r_1)^2 \times (r_1|\vec{p}_P|)^2$, where $P = \pi(K)$ for a moving pion (moving kaon), and 4) a term proportional to $(a/r_1)^2(r_1 q)^2$.

The last of these tests is motivated by the observed deviations of the simulation q^2 from zero. These small nonzero values of the momentum transfer, $(r_1 q)^2 \sim 10^{-4}$, are due to momentum-dependent discretization errors. None of the four fits increases the central value of Eq. (4.4) by more than 0.1%. We, therefore, take 0.1% as our estimate of the remaining discretization effects.

C. Finite volume effects

We carried out an additional simulation on an ensemble with the same parameters as the $a \approx 0.12$ fm, $am_l = 0.2am_s$ ensemble (second line in Table I) but with a larger volume (third line in Table I). This simulation in a larger volume gives a value for $f_+(0)$ that is ~0.1% lower. Note that the physical volumes of the other ensembles in Table I are comparable to that of the $a \approx 0.12$ fm, $am_l = 0.2am_s$ ensemble, and, in particular, the most chiral $a \approx 0.09$ fm ensemble has approximately equal size as the 0.12 fm ensemble where the finite size effects were checked explicitly and has roughly the same sea quark masses. We, therefore, take the 0.1% difference as the estimate of the finite volume error.

D. Lattice scale

Since we are calculating a dimensionless quantity, $f_+(0)$, uncertainties from the conversion from lattice units to physical units are small. We perform this conversion through an intermediate scale, the r_1 scale [65,66], by rewriting all dimensionful quantities entering in the chiral fit function in r_1 units. The lattice parameters are converted to r_1 units by using the values of r_1/a listed in Table I for each MILC ensemble. The physical parameters are expressed in r_1 units using the determination of the physical r_1 in Ref. [50], $r_1 = (0.3133 \pm 0.0023)$ fm. We use this value for consistency with the determination of the valence strange-quark masses in our simulations, which were tuned to their physical values in Ref. [50] using it. This value agrees within errors with the one obtained from the 2009 MILC determination of $r_1 f_\pi$ [67] combined with the Particle Data Group value of f_π [68], $r_1 = 0.3117 \pm 0.0022$ [69].

If we change the value of r_1 to the extremes allowed by its total error, the form factor $f_+(0)$ shifts 0.06% from our central value. We add this shift as a systematic error associated with the choice of scale.

E. Chiral scale

We checked that our result is independent of the chiral scale when varying Λ_χ between $M_\rho \pm 0.5$ GeV. Indeed, we find it is independent up to small fluctuations due mainly to the fact that the sunset two-loop contribution is handled numerically by means of a lookup table. We, thus, do not need to add any extra uncertainty from the choice of scale.

VI. CALCULATION OF THE $O(p^6)$ LEC COMBINATION $C^r_{12} + C^r_{34}$

As a byproduct of our calculation, from the parameter $C_6^{\prime(1)}$ in Eq. (4.2), we can extract the combination of $O(p^4)$ and $O(p^6)$ LECs:

$$(C^r_{12} + C^r_{34} - (L^r_5)^2)(M_\rho) = (3.62 \pm 0.33) \times 10^{-6}, \tag{6.1}$$

where the error is statistical only. Using the result in Eq. (6.1) and the LEC $L^r_5(M_\rho) = (0.97 \pm 0.22) \times 10^{-3}$, which we also get as an output of our chiral fit (although it is the same as the corresponding prior within the given precision), we obtain

$$(C^r_{12} + C^r_{34})(M_\rho) = (4.57 \pm 0.44 \pm 0.90) \times 10^{-6}, \tag{6.2}$$

where the first error is statistical and also includes the uncertainty from the errors of the hairpin parameters $\delta^{\mathrm{HISQ}}_{V,A}$ and the NLO LECs and the second error is from the systematic errors added in quadrature. The detailed error budget for this quantity is given in Table VII. In this case, we do not vary the value of the decay constant used in the chiral expansion at NNLO to estimate the systematic error associated with the chiral extrapolation and choice of fit function, since the LECs are not physical quantities and depend on convention. The value in Eq. (6.2) should be used taking into account the type of fit performed for its extraction. We obtain the other systematic errors in the same way as for $f_+(0)$, with the exception of finite volume effects. We estimate that error to be of the same size as that for $1 - f_+(0)$, which highly enhances the effect as compared to $f_+(0)$.

The result in Eq. (6.2) agrees with nonlattice determinations in Refs. [70–72]. The authors of those papers used the large N_c approximation, a coupled-channel dispersion relation analysis that provides the slope and the curvature of the scalar $K\pi$ form factor, and a quark model, respectively, to calculate the contribution from the combination $C_{12} + C_{34}$ to the form factor $f_+(0)$. The calculation of $C_{12} + C_{34} - L_5^2$ in Ref. [73], based on χPT, large N_c estimates of the LECs, and dispersive methods is $\sim 3\sigma$ smaller than our result in Eq. (6.1).

TABLE VII. Complete error budget and total error for the combination of $O(p^6)$ LEC's $C^r_{12} + C^r_{34}$. All errors are given in percent.

Source of uncertainty	$(C^r_{12} + C^r_{34})$ (%)
Statistics	9.6
Chiral extrapolation & fitting	12
Discretization	5.7
Scale	0.2
Finite volume	~11
Total Error	19.7

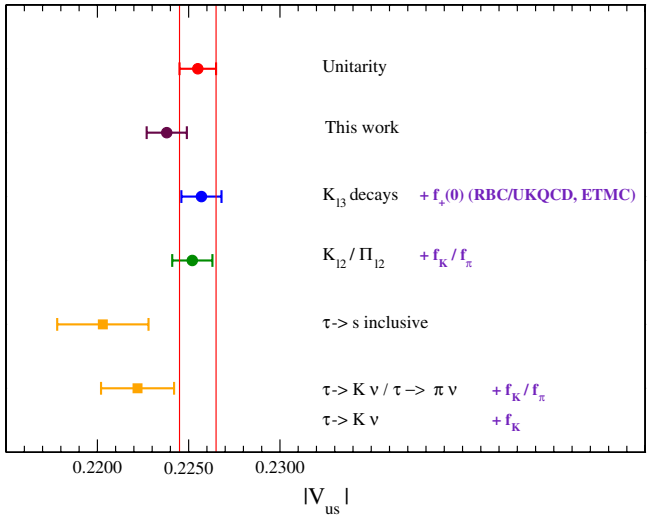

FIG. 6 (color online). Comparison of $|V_{us}|$ as extracted from K leptonic decays, K semileptonic decays, τ decays, and the value obtained in this work. Theoretical inputs for $f_+(0)$, f_K/f_π, and f_K are taken from Refs. [20,21], Refs. [4,5,7–9,13], and Refs. [4,7,13,77,78], respectively. The unitarity value of $|V_{us}|$ is obtained from $|V_{ud}| = 0.97425(22)$ [17] and neglects the contribution from $|V_{ub}|$. The vertical lines correspond to the unitarity prediction. See the text for further explanation.

VII. DISCUSSION OF RESULTS AND FUTURE IMPROVEMENTS

We summarize the error budget and total error for our calculation of $f_+(0)$ discussed in Sec. V in Table VI. Our final result is

$$f_+(0) = 0.9667 \pm 0.0023 \pm 0.0033 = 0.9667 \pm 0.0040, \tag{7.1}$$

where the first error in the middle expression is statistical and the second is the sum in quadrature of the different systematic errors. In Table VIII, we list our result together with other determinations from both unquenched lattice methods and different analytical approaches. Our result is about 1σ larger than previous unquenched lattice determinations by the RBC/UKQCD and ETMC Collaborations and has somewhat smaller errors mainly due to the use of two-loop χPT in the chiral extrapolation and the use of data at two lattice spacings. On the other hand, the form factor in Eq. (7.1) is smaller than those coming from analytical approaches based on the use of two-loop χPT and model estimates of the $O(p^6)$ LECs, although it is compatible within errors (except for the calculation in Ref. [73]).

With the form factor in Eq. (7.1) and the latest average of experimental measurements of K semileptonic decays [16], $|V_{us}|f_+(0) = 0.2163(5)$, we obtain the following value for the CKM matrix element $|V_{us}|$:

$$|V_{us}| = 0.2238 \pm 0.0009 \pm 0.0005 = 0.2238 \pm 0.0011, \tag{7.2}$$

where the first error is from $f_+(0)$ and the second one is experimental. Using the average over superallowed nuclear beta decay determinations, $|V_{ud}| = 0.97425(22)$ [17], and neglecting the value of $|V_{ub}|$, the result in Eq. (7.2) allows us to check unitarity in the first row of the CKM matrix,

$$\Delta_{\mathrm{CKM}} \equiv |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 - 1 = -0.0008(6). \tag{7.3}$$

In Fig. 6, we summarize the status of the determination of $|V_{us}|$ from different sources: this work, semileptonic decays using as theory input previous unquenched lattice calculations of $f_+(0)$ listed in Table VIII, leptonic decays using as input the average over $N_f = 2 + 1$ lattice calculations of f_K/f_π [4,5,7–9,13], inclusive hadronic τ decays [3,74,75], and exclusive τ decay modes [3] that require either f_K/f_π [4,5,7–9,13] or f_K [4,7,13,77,78] from the lattice as theory inputs. The value predicted by unitarity of the CKM matrix is also included in the figure.

Using staggered fermions, we have obtained the most precise lattice-QCD value of $f_+(0)$ to date. Three elements of our work are the key to this precision. First, the MILC asqtad ensembles yield very small statistical errors, ~0.1–0.15%, providing a solid foundation for the whole analysis. Second, using the HISQ action for the valence quarks leads to very small discretization errors. Third, the χPT description of the chiral behavior

TABLE VIII. Form factor $f_+(0)$ as extracted from unquenched lattice calculations (first half of the table), phenomenological approaches based on the use of two-loop χPT, and the pioneering calculation by Leutwyler and Roos using a quark model. For those calculations based on two-loop χPT, we also indicate the method used in the estimate of the $O(p^6)$ LECs.

Group	$f_+(0)$	Method	Sea content
This work	0.9667(23)(33)	staggered fermions	$N_f = 2 + 1$
RBC/UKQCD [20]	$0.9599(34)(^{+31}_{-43})$	domain wall fermions	$N_f = 2 + 1$
ETMC [21]	0.9560(57)(62)	twisted mass fermions	$N_f = 2$
Kastner & Neufeld [73]	0.986(8)	χPT + large N_c limit + dispersion relations	-
Cirigliano *et al.* [70]	0.984(12)	χPT + large N_c limit	-
Jamin, Oller, & Pich [71]	0.974(11)	χPT + dispersion relations (scalar form factor)	-
Bijnens & Talavera [54]	0.976(10)	χPT + Leutwyler & Roos	-
Leutwyler & Roos [72]	0.961(8)	Quark model	-

of our data—including taste-breaking, partially quenched, and mixed-action effects at one loop—permit a controlled extrapolation to the continuum limit and physical light-quark mass. In particular, our analysis shows that one-loop PQSχPT may be enough to account for the discretization effects observed in our data, so our results do not change even if we disregard any additional sources of discretization error.

The uncertainty in the form factor still dominates the determination of $|V_{us}|$ from K semileptonic decays, so further improvements are necessary. We are in the process of reducing the dominant sources of systematic error in our analysis by performing simulations with physical light- and strange-quark masses on the HISQ $N_f = 2 + 1 + 1$ MILC ensembles [59]. With lattice data at the physical quark masses, we expect to greatly reduce the chiral extrapolation errors. Discretization errors are also considerably smaller for the HISQ action than for the asqtad action, as can be seen explicitly in Ref. [25]. Another advantage of the HISQ ensembles is that the strange sea-quark masses are much better tuned than for the asqtad ensembles, which will reduce the dominant contribution to the chiral-fit systematic error. Finally, the MILC HISQ ensembles include the effects of charm quarks in the sea.

ACKNOWLEDGMENTS

We thank Jon Bailey, Johan Bijnens, Christine Davies, Eduardo Follana, Pere Masjuan, and Heechang Na for useful discussions. We thank Johan Bijnens for making his NLO partially quenched χPT and NNLO full QCD χPT codes available to us. We thank Jon Bailey for the careful reading of this manuscript. Computations for this work were carried out with resources provided by the USQCD Collaboration, the Argonne Leadership Computing Facility, the National Energy Research Scientific Computing Center, and the Los Alamos National Laboratory, which are funded by the Office of Science of the United States Department of Energy, and with resources provided by the National Institute for Computational Science, the Pittsburgh Supercomputer Center, the San Diego Supercomputer Center, and the Texas Advanced Computing Center, which are funded through the National Science Foundation's Teragrid/XSEDE Program. This work was supported in part by the U.S. Department of Energy under Grants No. DE-FG02-91ER40628 (C. B.), No. DOE FG02-91ER40664 (Y. M.), No. DE-FC02-06ER41446 (C. D., J. F., L. L., M. B. O.), No. DE-FG02-91ER40661 (S. G., R. Z.), No. DE-FG02-91ER40677 (D. D., A. X. K.), and No. DE-FG02-04ER-41298 (J. K., D. T.); by the National Science Foundation under Grants No. PHY-1067881, No. PHY-0757333, No. PHY-0703296 (C. D., J. F., L. L., M. B. O.), and No. PHY-0757035 (R. S.); by the Science and Technology Facilities Council and the Scottish Universities Physics Alliance (J. L.); by the MICINN (Spain) under Grant No. FPA2010-16696 and the Ramón y Cajal program (E. G.); by the Junta de Andalucía (Spain) under Grants No. FQM-101, No. FQM-330, and No. FQM-6552 (E. G.); and by European Commission (EC) under Grant No. PCIG10-GA-2011-303781 (E. G.). This manuscript has been coauthored by employees of Brookhaven Science Associates, LLC, under Contract No. DE-AC02-98CH10886 with the U.S. Department of Energy. Fermilab is operated by Fermi Research Alliance, LLC, under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy.

APPENDIX: STAGGERED CHIRAL PERTURBATION THEORY AT NLO

As described in Sec. IV, we use SχPT to incorporate the dominant discretization effects at one loop. For staggered quarks, the taste-nonsinglet pseudoscalar mesons masses are split according to five taste representations $\Xi = P, A, T, V, I$:

$$M^2_{ij,\Xi} = \mu(m_i + m_j) + a^2\Delta_\Xi, \tag{A1}$$

with m_i and m_j the quark masses. All the taste-nonsinglet pseudoscalar mesons contribute via the loops in the SχPT calculation, even when we choose our external states to be the ones with $\Xi = P$.

Since we are using a different staggered action in the sea (asqtad) and in the valence (HISQ) sectors, there are mixed-action effects that we need to incorporate in our expressions. At one loop, those effects arise as changes to the taste splittings Δ_Ξ, extra terms in the disconnected propagators, and the appearance of three different hairpin coefficients in each channel. These modifications affect the structure of the disconnected propagator only in the axial and vector channels, changing them from the usual form [79]

$$\mathcal{D}^{\Xi}_{XY}(p) = -\frac{a^2\delta_\Xi}{(p^2 + m^2_{X_\Xi})(p^2 + m^2_{Y_\Xi})} \times \frac{(p^2 + m^2_{U_\Xi})(p^2 + m^2_{D_\Xi})(p^2 + m^2_{S_\Xi})}{(p^2 + m^2_{\pi_\Xi})(p^2 + m^2_{\eta_\Xi})(p^2 + m^2_{\eta'_\Xi})}, \tag{A2}$$

with $\Xi = V, A$ and U, D, S the $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ sea-quark mesons, respectively, to the following form [57]:

$$\mathcal{D}^{\Xi}_{XY}(p) = -\frac{a^2(\delta^{v\sigma}_\Xi)^2/\delta^{\sigma\sigma}_\Xi}{(p^2 + m^2_{X_\Xi})(p^2 + m^2_{Y_\Xi})} \times \frac{(p^2 + m^2_{U_\Xi})(p^2 + m^2_{D_\Xi})(p^2 + m^2_{S_\Xi})}{(p^2 + m^2_{\pi_\Xi})(p^2 + m^2_{\eta_\Xi})(p^2 + m^2_{\eta'_\Xi})} - \frac{a^2[\delta^{vv} - (\delta^{v\sigma}_\Xi)^2/\delta^{\sigma\sigma}_\Xi]}{(p^2 + m^2_{X_\Xi})(p^2 + m^2_{Y_\Xi})}, \tag{A3}$$

where v and σ label the valence and sea quark, respectively. It seems reasonable to expect $\delta_\Xi^{v\sigma} \sim \sqrt{\delta_\Xi^{vv}\delta_\Xi^{\sigma\sigma}}$, so the mixed action effects should be small. To take this into account in our fits, we define the parameter $\delta_\Xi^{\rm mix} = \delta^{vv} - (\delta_\Xi^{v\sigma})^2/\delta_\Xi^{\sigma\sigma}$ and write the SχPT expressions in terms of $\delta_\Xi^{\rm mix}$ (which we expect to be suppressed with respect to the parameters in the sea and the valence sectors) and δ_Ξ^{vv}:

$$\mathcal{D}_{XY}^{\Xi}(p) = -\frac{a^2(\delta_\Xi^{vv} - \delta_\Xi^{\rm mix})}{(p^2 + m_{X_\Xi}^2)(p^2 + m_{Y_\Xi}^2)} \times \frac{(p^2 + m_{U_\Xi}^2)(p^2 + m_{D_\Xi}^2)(p^2 + m_{S_\Xi}^2)}{(p^2 + m_{\pi_\Xi}^2)(p^2 + m_{\eta_\Xi}^2)(p^2 + m_{\eta'_\Xi}^2)} - \frac{a^2\delta_\Xi^{\rm mix}}{(p^2 + m_{X_I}^2)(p^2 + m_{Y_I}^2)}. \tag{A4}$$

Although we can not determine very precisely the value of $\delta_\Xi^{\rm mix}$ from our data, the chiral fits prefer nonzero values that are the same sign as but an order of magnitude smaller than δ_Ξ^{vv}. After these mixed-action expressions were derived, we discovered that they had been previously worked out in Ref. [80], including the comments about the expected size of the $\delta_\Xi^{v\sigma}$.

The partially quenched one-loop calculation involves mesons made of two valence, two sea, and one valence and one sea quarks. The corresponding masses are given by the definition in Eq. (A1) with the taste splittings Δ_Ξ being $\Delta_\Xi^{\rm HISQ}$, $\Delta_\Xi^{\rm asqtad}$, and $\Delta_\Xi^{\rm mix}$, respectively.

1. Result for f_2

A detailed description of the one-loop PQSχPT calculation with and without a mixed action setup will be given elsewhere [57]. Here, we just give the final result needed for the extrapolation of the form factor $f_+(q^2 = 0)$. The NLO PQSχPT expression for the form factor at zero momentum transfer in the isospin limit and for $N_f = 2 + 1$ flavors of sea quarks, $f_2(0) \equiv f_2^{\rm PQ\chi PT}(q^2 = 0, a)$, is (for x and y, the light and strange valence quarks, respectively)

$$\begin{aligned}
f_2(0) = \frac{-1}{2(4\pi f)^2}\Bigg\{&\frac{1}{16}\sum_{f,\Xi}[\ell(m_{xf,\Xi}) + \ell(m_{yf,\Xi})] + \frac{1}{3}\Bigg[\sum_j \frac{\partial}{\partial m_{X,I}^2}(R_j^{[2,2]}(\mathcal{M}_I^{(2,x)};\mu_I^{(2)})\ell(m_{j,I})) \\
&+ \sum_j \frac{\partial}{\partial m_{Y,I}^2}(R_j^{[2,2]}(\mathcal{M}_I^{(2,y)};\mu_I^{(2)})\ell(m_{j,I})) + 2\sum_j R_j^{[3,2]}(\mathcal{M}_I^{(3,xy)};\mu_I^{(2)})\ell(m_{j,I})\Bigg] + \frac{1}{4}\sum_{f,\Xi} B_{22}(m_{xf,\Xi}, m_{yf,\Xi}, 0) \\
&+ \frac{4}{3}\frac{\partial}{\partial m_{X,I}^2}\Bigg\{\sum_j \tilde{B}_{22}(m_{xy,I}^2, m_{j,I}^2, 0)R_j^{[2,2]}(\mathcal{M}_I^{(2,x)};\mu_I^{(2)})\Bigg\} + \frac{4}{3}\frac{\partial}{\partial m_{Y,I}^2}\Bigg\{\sum_j \tilde{B}_{22}(m_{xy,I}^2, m_{j,I}^2, 0)R_j^{[2,2]}(\mathcal{M}_I^{(2,y)};\mu_I^{(2)})\Bigg\} \\
&+ \frac{8}{3}\sum_j \tilde{B}_{22}(m_{xy,I}^2, m_{j,I}^2, 0)R_j^{[3,2]}(\mathcal{M}_I^{(3,xy)};\mu_I^{(2)}) + a^2(\delta_V^{vv} - \delta_V^{\rm mix})\Bigg[\sum_j \frac{\partial}{\partial m_{X,V}^2}(R_j^{[3,2]}(\mathcal{M}_V^{(3,x)};\mu_V^{(2)})\ell(m_{j,V})) \\
&+ \sum_j \frac{\partial}{\partial m_{Y,V}^2}(R_j^{[3,2]}(\mathcal{M}_V^{(3,y)};\mu_V^{(2)})\ell(m_{j,V})) + 2\sum_j R_j^{[4,2]}(\mathcal{M}_V^{(4,xy)};\mu_V^{(2)})\ell(m_{j,V}) \\
&+ 4\frac{\partial}{\partial m_{X,V}^2}\Bigg\{\sum_j \tilde{B}_{22}(m_{xy,V}^2, m_{j,V}^2, 0)R_j^{[3,2]}(\mathcal{M}_V^{(3,x)};\mu_V^{(2)})\Bigg\} + 4\frac{\partial}{\partial m_{Y,V}^2}\Bigg\{\sum_j \tilde{B}_{22}(m_{xy,V}^2, m_{j,V}^2, 0)R_j^{[3,2]}(\mathcal{M}_V^{(3,y)};\mu_V^{(2)})\Bigg\} \\
&+ 8\sum_j \tilde{B}_{22}(m_{xy,V}^2, m_{j,V}^2, 0)R_j^{[4,2]}(\mathcal{M}_V^{(4,xy)};\mu_V^{(2)})\Bigg] + a^2\delta_V^{\rm mix}\Bigg[\frac{\partial\ell(m_{X,V})}{\partial m_{X,V}^2} + \frac{\partial\ell(m_{Y,V})}{\partial m_{Y,V}^2} + 2\frac{(\ell(m_{X,V}) - \ell(m_{Y,V}))}{m_{Y,V}^2 - m_{X,V}^2} \\
&+ 4\frac{\partial}{\partial m_{X,V}^2}\tilde{B}_{22}(m_{xy,V}^2, m_{X,V}^2, 0) + 4\frac{\partial}{\partial m_{Y,V}^2}\tilde{B}_{22}(m_{xy,V}^2, m_{Y,V}^2, 0) + \frac{8}{m_{Y,V}^2 - m_{X,V}^2}(\tilde{B}_{22}(m_{xy,V}^2, m_{X,V}^2, 0) \\
&- \tilde{B}_{22}(m_{xy,V}^2, m_{Y,V}^2, 0))\Bigg] + [V \to A]\Bigg\},
\end{aligned} \tag{A5}$$

where Ξ runs over the sixteen independent meson tastes, f runs over sea-quark flavors, $\ell(m)$ and $\tilde{B}_{22}$ are chiral logarithm functions defined below, and $R_j^{[n,k]}(\mathcal{M};\mu)$ are residue functions introduced in Ref. [79], with $\mathcal{M}$ and μ various sets of meson masses:

$$\begin{aligned}
&\{\mathcal{M}_\Xi^{(2,z)}\} \equiv \{m_{\eta,\Xi}, m_{Z,\Xi}\}, \qquad \{\mathcal{M}_\Xi^{(3,z,z')}\} \equiv \{m_{\eta,\Xi}, m_{Z,\Xi}, m_{Z',\Xi}\}, \qquad \{\mathcal{M}_\Xi^{(3,z)}\} \equiv \{m_{\eta,\Xi}, m_{\eta',\Xi}, m_{Z,\Xi}\}, \\
&\{\mathcal{M}_\Xi^{(4,z,z')}\} \equiv \{m_{\eta,\Xi}, m_{\eta',\Xi}, m_{Z,\Xi}, m_{Z',\Xi}\}, \qquad \{\mu_\Xi^{(2)}\} \equiv \{m_{U,\Xi}, m_{S,\Xi}\}.
\end{aligned} \tag{A6}$$

Here, z can be either x or y, and Z is the corresponding $z\bar{z}$ meson (X or Y).

The chiral logarithm function ℓ is given by

$$\ell(m) \equiv m^2 \ln(m^2/\Lambda_\chi^2), \tag{A7}$$

with Λ_χ the chiral scale. The function $\tilde{B}_{22}(m_1^2, m_2^2, q^2)$ is related to the function $\bar{B}_{22}$ defined in Ref. [81] by

$$\tilde{B}_{22}(m_1^2, m_2^2, q^2) = (4\pi)^2 \bar{B}_{22}(m_1^2, m_2^2, q^2). \tag{A8}$$

In the special case of $q^2 = 0$, $\tilde{B}_{22}$ takes the simple form

$$\tilde{B}_{22}(m_1^2, m_2^2, 0) = -\frac{1}{4}\left(\frac{m_2^2\ell(m_2^2) - m_1^2\ell(m_1^2)}{m_2^2 - m_1^2}\right) + \frac{1}{8}(m_1^2 + m_2^2). \tag{A9}$$

It is not difficult to check that Eq. (A5) obeys the Ademollo-Gato theorem [53] in the valence masses and vanishes as $(m_y - m_x)^2$ as $m_y \to m_x$. To show this, one needs the identities obeyed by the $R_j^{[n,k]}$ [79] (or, more simply, the structure of the integrals from which they arise) as well as Eqs. (A7) and (A9) for the terms involving the sum over f. Eq. (A5) analytically agrees with the isospin symmetric partially quenched continuum calculation [56].

[1] V. Cirigliano, J. Jenkins, and M. González-Alonso, Nucl. Phys. **B830**, 95 (2010).
[2] V. Cirigliano, G. Ecker, H. Neufeld, A. Pich, and J. Portolés, Rev. Mod. Phys. **84**, 399 (2012).
[3] Y. Amhis *et al.* (Heavy Flavor Averaging Group), arXiv:1207.1158.
[4] J. Laiho and R. S. Van de Water, Proc. Sci., LATTICE2011 (2011) 293 [arXiv:1112.4861].
[5] S. Dürr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, A. Ramos, and K. K. Szabó, Phys. Rev. D **81**, 054507 (2010).
[6] C. Aubin, C. Bernard, C. DeTar, J. Osborn, Steven Gottlieb, E. Gregory, D. Toussaint, U. Heller, J. Hetrick, and R. Sugar (MILC Collaboration), Phys. Rev. D **70**, 114501 (2004).
[7] A. Bazavov *et al.* (MILC Collaboration), Proc. Sci., LATTICE2010 (2010) 074 [arXiv:1012.0868].
[8] Y. Aoki *et al.* (RBC and UKQCD Collaborations), Phys. Rev. D **83**, 074508 (2011).
[9] E. Follana, C. Davies, G. Lepage, and J. Shigemitsu (HPQCD and UKQCD Collaborations), Phys. Rev. Lett. **100**, 062002 (2008).
[10] R. Arthur *et al.* (RBC and UKQCD Collaborations), arXiv:1208.4412.
[11] A. Bazavov *et al.* (MILC Collaboration), arXiv:1301.5855.
[12] G. Colangelo *et al.*, Eur. Phys. J. C **71**, 1695 (2011).
[13] J. Laiho, E. Lunghi, and R. S. Van de Water, Phys. Rev. D **81**, 034503 (2010); updated information can be found at http://www.latticeaverages.org.
[14] W. J. Marciano, Phys. Rev. Lett. **93**, 231803 (2004).
[15] M. Antonelli *et al.*, Eur. Phys. J. C **69**, 399 (2010).
[16] M. Moulson, arXiv:1209.3426.
[17] J. C. Hardy and I. S. Towner, Phys. Rev. C **79**, 055502 (2009).
[18] E. Follana, Q. Mason, C. Davies, K. Hornbostel, P. Lepage, and H. Trottier (HPQCD Collaboration), Nucl. Phys. B, Proc. Suppl. **129-130**, 447 (2004); **129-130**, 384 (2004); E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD and UKQCD Collaborations), Phys. Rev. D **75**, 054502 (2007).
[19] G. P. Lepage, Phys. Rev. D **59**, 074502 (1999).
[20] P. A. Boyle, J. M. Flynn, A. Jüttner, C. Kelly, C. Maynard, H. Pedroso Lima, C. T. Sachrajda, and J. M. Zanotti (RBC-UKQCD Collaboration), Eur. Phys. J. C **69**, 159 (2010).
[21] V. Lubicz, F. Mescia, S. Simula, and C. Tarantino (ETM Collaboration), Phys. Rev. D **80**, 111502 (2009).
[22] H. Na, C. T. H. Davies, E. Follana, G. Peter Lepage, and J. Shigemitsu (HPQCD Collaboration), Phys. Rev. D **82**, 114506 (2010).
[23] J. A. Bailey *et al.* (Fermilab Lattice and MILC Collaboration), Proc. Sci., LATTICE2010 (2010) 306 [arXiv:1011.2423].
[24] E. Gámiz *et al.* (Fermilab Lattice and MILC Collaboration), Proc. Sci., LATTICE2011 (2011) 281 [arXiv:1111.2021].
[25] E. Gámiz *et al.* (Fermilab Lattice and MILC Collaboration), Proc. Sci., LATTICE2012 (2012) 113 [arXiv:1211.0751].
[26] T. Kaneko *et al.* (JLQCD Collaboration), Proc. Sci., LATTICE2012 (2012) 111 [arXiv:1211.6180].
[27] P. A. Boyle, J. M. Flynn, A. Jüttner, C. Sachrajda, K. Sivalingam, and J. M. Zanotti, Proc. Sci., LATTICE2012 (2012)112 [arXiv:1212.3188].
[28] H. Na, C. T. H. Davies, E. Follana, P. Lepage, and J. Shigemitsu, Proc. Sci., LATTICE2009 (2009) 247 [arXiv:0910.3919].
[29] J. Koponen *et al.* (HPQCD Collaboration), arXiv:1208.6242.
[30] C. McNeile and C. Michael (UKQCD Collaboration), Phys. Rev. D **73**, 074506 (2006).
[31] P. F. Bedaque, J.-W. Chen, Phys. Lett. B **616**, 208 (2005).
[32] C. T. Sachrajda and G. Villadoro, Phys. Lett. B **609**, 73 (2005).
[33] D. Guadagnoli, F. Mescia, and S. Simula, Phys. Rev. D **73**, 114504 (2006).
[34] P. A. Boyle, J. M. Flynn, A. Jüttner, C. T. Sachrajda, and J. M. Zanotti, J. High Energy Phys. 05 (2007) 016.
[35] A. Bazavov *et al.*, Rev. Mod. Phys. **82**, 1349 (2010).

[36] C. Bernard, T. Burch, C. DeTar, S. Gottlieb, L. Levkova, U. Heller, J. Hetrick, R. Sugar, and D. Toussaint, Phys. Rev. D **75**, 094505 (2007).

[37] C. W. Bernard, T. Burch, K. Orginos, D. Toussaint, T. DeGrand, C. DeTar, S. Datta, S. Gottlieb, U. Heller, and R. Sugar, Phys. Rev. D **64**, 054506 (2001).

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

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

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

[41] S. Dürr and C. Hoelbling, Phys. Rev. D **69**, 034503 (2004).

[42] S. Dürr, C. Hoelbling, and U. Wagner, Phys. Rev. D **70**, 094501 (2004).

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

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

[45] C. Bernard, M. Golterman, Y. Shamir, and S. R. Sharpe, Phys. Lett. B **649**, 235 (2007).

[46] Y. Shamir, Phys. Rev. D **71**, 034509 (2005); **75**, 054503 (2007).

[47] C. Bernard, M. Golterman, and Y. Shamir, Phys. Rev. D **77**, 074505 (2008).

[48] D. H. Adams, Phys. Rev. D **77**, 105024 (2008).

[49] C. R. Allton, arXiv:hep-lat/9610016.

[50] C. T. H. Davies, E. Follana, I. D. Kendall, G. P. Lepage, and C. McNeile (HPQCD Collaboration), Phys. Rev. D **81**, 034506 (2010).

[51] The η_s is a ficticious particle made of a strange quark-antiquark pair, considered without the disconnected contribution. Its "physical value" was determined in Ref. [50].

[52] J. A. Bailey *et al.* (Fermilab Lattice and MILC Collaboration), Phys. Rev. D **79**, 054507 (2009).

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

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

[55] D. Becirevic, G. Martinelli, and G. Villadoro, Phys. Lett. B **633**, 84 (2006).

[56] Johan Bijnens, private communication. We agree with J. Bijnens that there is a misprint in the $N_f = 3$ formula for f_2 in the appendix of Ref. [55].

[57] C. Bernard and E. Gámiz (unpublished).

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

[59] A. Bazavov *et al.* (MILC Collaboration), arXiv:1212.4768.

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

[61] G. Amorós, J. Bijnens, and P. Talavera, Nucl. Phys. **B602**, 87 (2001).

[62] The authors of Ref. [61] fit experimental data on K_{l4} decays to the corresponding two-loop χPT expressions to extract the value of the NLO LECs. In fit 10, L_4 and L_6 are set to zero. More details on the type of fit performed can be found in Ref. [63].

[63] G. Amorós, J. Bijnens, and P. Talavera, Phys. Lett. B **480**, 71 (2000).

[64] J. Gasser and H. Leutwyler, Nucl. Phys. **B250**, 465 (1985).

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

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

[67] A. Bazavov *et al.* (MILC Collaboration), Proc. Sci., CD09, (2009) 007 [arXiv:0910.2966].

[68] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).

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

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

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

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

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

[74] E. Gámiz, M. Jamin, A. Pich, J. Prades, and F. Schwab, Phys. Rev. Lett. **94**, 011803 (2005); arXiv:hep-ph/0610246.

[75] The value shown in Fig. 6 uses the theoretical calculation in Ref. [74] together with the latest experimental averages for hadronic τ decay modes [3], except for a few modes for which the experimental measurements are substituted by theoretical predictions as described in Ref. [76].

[76] V. Bernard, D. R. Boito, and E. Passemar, Nucl. Phys. B, Proc. Suppl. **218**, 140 (2011).

[77] C. T. H. Davies, C. McNeile, E. Follana, G. P. Lepage, H. Na, and J. Shigemitsu (HPQCD Collaboration), Phys. Rev. D **82**, 114504 (2010).

[78] C. Kelly (RBC and UKQCD Collaborations), Proc. Sci., LATTICE2011 (2011) 285 [arXiv:1201.0706].

[79] C. Aubin and C. Bernard, Phys. Rev. D **68**, 034014 (2003); **68**, 074011 (2003).

[80] T. Bae, Y.-C. Jang, C. Jung, H.-J. Kim, J. Kim, K. Kim, W. Lee, and S. R. Sharpe, Phys. Rev. D **82**, 114509 (2010).

[81] J. Bijnens and P. Talavera, J. High Energy Phys. 03 (2002) 046.