Effects of nonequilibrated topological charge distributions on pseudoscalar meson masses and decay constants

C. Bernard[*]

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

D. Toussaint[†]

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

We study the effects of failure to equilibrate the squared topological charge Q^2 on lattice calculations of pseudoscalar masses and decay constants. The analysis is based on chiral perturbation theory calculations of the dependence of these quantities on the QCD vacuum angle θ. For the light-light partially quenched case, we rederive the known chiral perturbation theory results of Aoki and Fukaya, but using the nonperturbatively valid chiral theory worked out by Golterman, Sharpe and Singleton, and by Sharpe and Shoresh. We then extend these calculations to heavy-light mesons. Results when staggered taste violations are important are also presented. The derived Q^2 dependence is compared to that of simulations using the MILC Collaboration's ensembles of lattices with four flavors of dynamical highly improved staggered quarks. We find agreement, albeit with large statistical errors. These results can be used to correct for the leading effects of unequilibrated Q^2, or to make estimates of the systematic error coming from the failure to equilibrate Q^2. In an appendix, we show that the partially quenched chiral theory may be extended beyond a lower bound on valence masses discovered by Sharpe and Shoresh. Subtleties occurring when a sea-quark mass vanishes are discussed in another appendix.

I. INTRODUCTION

In continuum QCD the topological charge Q cannot change in a continuous evolution of the gluon fields. Thus we expect that lattice QCD simulations using approximately continuous evolution algorithms should see very slow evolution of the topological charge, since changing the topological charge involves a tunneling where some of the plaquettes or other loops in the gauge action pass through large values. This expected slow evolution of the topological charge has been observed and studied in Refs. [1–5]. Since the rate at which the topological charge Q changes in a lattice simulation falls off quickly as the lattice spacing decreases, modern QCD simulations are reaching a regime where the distribution of Q cannot be accurately sampled in a simulation with practical length. When this is the case, physical quantities will suffer a systematic error, and we need to either correct for this error or account for it in our error budgets.

Here, we use the MILC Collaboration's ensembles of lattices with a one-loop Symanzik and tadpole-improved gauge action and four flavors of highly improved staggered quarks (HISQ quarks) to study the errors induced by an insufficiently sampled Q distribution. We first demonstrate the expected slow evolution of topological charge as the lattice spacing decreases.

We proceed to calculate, in chiral perturbation theory (χPT), the dependence of the light-light and heavy-light pseudoscalar masses and decay constants on the QCD vacuum angle θ, which is related to their dependence on the average Q^2 in the lattice simulation [6]. The unitary case (valence and sea-quark masses identical) for light-light mesons is treated first, mainly to introduce the methods and set the notation; the results were already reported in Ref. [6], or may be obtained by straightforward generalization of that calculation. We then discuss the light-light partially quenched case, which has also been treated by Aoki and Fukaya [7] using partially quenched chiral perturbation theory (PQχPT) with the replica method [8]. Because the vacuum state changes in the presence of θ, it could in principle be important to use a nonperturbatively valid method for the partially quenched theory. Rather than the replica method [8], which has

[*]cb@wustl.edu
[†]doug@physics.arizona.edu

not been justified nonperturbatively, we therefore employ the approach to PQχPT introduced by Golterman, Sharpe, and Singleton [9] and Sharpe and Shoresh [10]. A potential sticking point, however, is the bound on the values of valence and sea-quark masses found by Sharpe and Shoresh. When this bound is violated, the PQχPT approach of Refs. [9,10] appears to break down. We are able to show (Appendix B) that the bound is actually spurious, and the chiral theory continues to be valid when the bound is violated. Once the partially quenched light-light case is analyzed, it is not difficult to generalize it to the heavy-light case, or to include the leading discretization effects coming from staggered taste violations. The details of the partially quenched light-light and heavy-light calculations constitute the majority of this paper.

Once the χPT predictions are in hand, they are compared to the HISQ simulation data. Although the statistical errors are large, we find good qualitative agreement between predictions and data. We discuss how to use this information to correct for the difference between the average of the squared topological charge in a simulation, $\langle Q^2 \rangle_{\text{sample}}$, and the correct average $\langle Q^2 \rangle$.

The remainder of this paper is organized as follows. In Sec. II, we discuss the evolution of topological charge in our simulations. The connection between the dependence of physical quantities on the topological charge in fixed volume, and their dependence on θ in infinite volume, is reviewed in Sec. III. Sections IV and V present the dependence of the mass and decay constants of light-quark pseudoscalar mesons in the unitary and partially quenched cases, respectively. In Sec. VI, the calculation is extended to heavy-light pseudoscalar mesons. We briefly describe the inclusion of staggered taste-violating effects in Sec. VII, and give the results for the light-light and heavy-light cases. Finally, in Sec. VIII, we use the correlation between Q^2 and masses and decay constants in our simulations to estimate the derivatives with respect to θ, compare them to χPT, and discuss how simulation results might be adjusted. A brief conclusion summarizes our main results.

There are three appendices. Appendix A discusses subtleties that occur when one or more sea-quark masses vanish. Appendix B investigates the Sharpe-Shoresh bound [10] on quark masses in the partially quenched chiral theory. A brief discussion of decoupling issues in the current context is presented in Appendix C.

A preliminary report of this work, which did not yet include results for heavy-light mesons or taste violations, was presented at *Lattice 2016* [11].

II. EVOLUTION OF THE TOPOLOGICAL CHARGE IN LATTICE SIMULATIONS

The ensembles we study have lattice spacings ranging from 0.09 to 0.03 fm, and light sea-quark masses (m_l) at either one fifth of the strange-quark mass (m_s) or

approximately the physical light quark mass, which is approximately $m_s/27$. See Refs. [12,13] for the parameters of the ensembles and the details of their generation. We measure the topological charge on these ensembles using the procedure described in Ref. [14]. This procedure consists of three HYP smearings of the lattice [15] followed by an integration of the correlator of an improved topological density operator [16]. In addition to the tests described in Ref. [14], there is a recent study comparing many methods of measuring the topological charge, finding generally good consistency among the methods [17].

Figure 1 shows the time histories of $Q/V^{1/2}$ in our simulations, where V is the lattice volume in fm^4, and there are periodic boundary conditions on the gauge field in all four directions. In this plot the blue lines are for ensembles with a light sea-quark mass that is one fifth of the strange-quark mass, and the red lines are for ensembles with the physical light quark mass. The increasing autocorrelation time of Q as a decreases is obvious, and at $a = 0.03$ fm we see that the simulation has only covered a small range of Q^2. The operator we use to measure Q is noisy enough, and the volume of the lattices large enough, that we do not see plateaus at integer values of Q in Fig. 1, or, for that matter, in histograms of the topological charge.
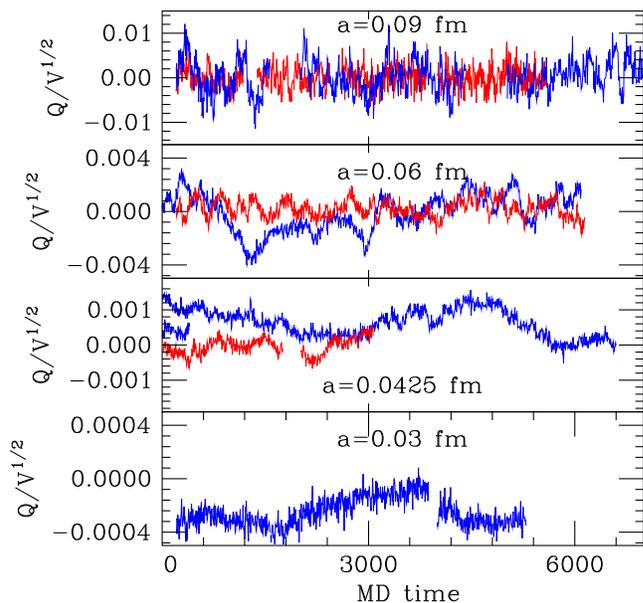

FIG. 1. Topological charge time histories for various lattice spacings; note that the vertical scale decreases as the lattice spacing decreases (top to bottom). Blue lines are for ensembles with a light sea-quark mass that is one fifth of the strange-quark mass, and the red lines are for ensembles with a light sea-quark mass at its physical value, $\approx m_s/27$. Notice the narrower distributions and shorter autocorrelation times for physical quark mass ensembles. Breaks in the traces separate multiple runs at the same couplings. The second short blue trace at $a = 0.0425$ fm is from a run with 3 times longer molecular dynamics trajectories than the main run.
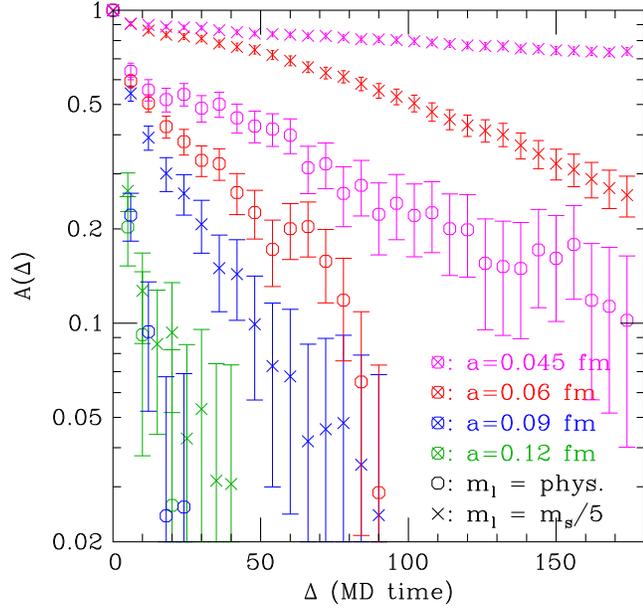
FIG. 2. Autocorrelations of the squared topological charge, where $A(\Delta)$ is defined in Eq. (1). For each lattice spacing, the crosses are the ensemble with $m_l = m_s/5$ and the octagons the ensemble with m_l at its physical value, $\approx m_s/27$.

For each lattice spacing, the local structure of the time evolution is similar for the $m_l = m_s/5$ ensemble and the physical m_l ensemble: Q typically changes by about the same amount in each time unit. However, in the $m_l = m_s/5$ ensembles Q ranges over larger values, and so it takes longer to random walk over this range, leading to a longer autocorrelation time. This is as expected, since the gauge action controls the tunneling rate for Q, so the average squared change in Q per unit volume per unit simulation time is approximately independent of the light quark mass. However, the fermion determinant suppresses the average Q^2, and the topological susceptibility, $\langle Q^2/V \rangle$, is approximately proportional to m_l. Figure 2 shows the autocorrelation of the squared topological charge for four different lattice spacings,

$$A(\Delta) = \frac{\langle Q^2(t)Q^2(t+\Delta) \rangle - \langle Q^2(t) \rangle^2}{\langle Q^4(t) \rangle}. \quad (1)$$

We use Q^2 rather than Q because it is Q^2 that controls the effects on masses and decay constants (and all other CP-conserving correlators). In this graph we see the expected increase of the autocorrelation times as the lattice spacing decreases, and also that the autocorrelations are smaller for the physical quark mass ensembles (octagons) than for the $m_l = m_s/5$ ensembles (crosses).

We define ΔQ as the change in Q over molecular dynamics time Δt. Figure 3 shows the tunneling rate per volume, $\langle (\Delta Q)^2/(V\Delta t) \rangle$ with octagons, where the blue symbols are for the $m_s/5$ ensembles and the red for the
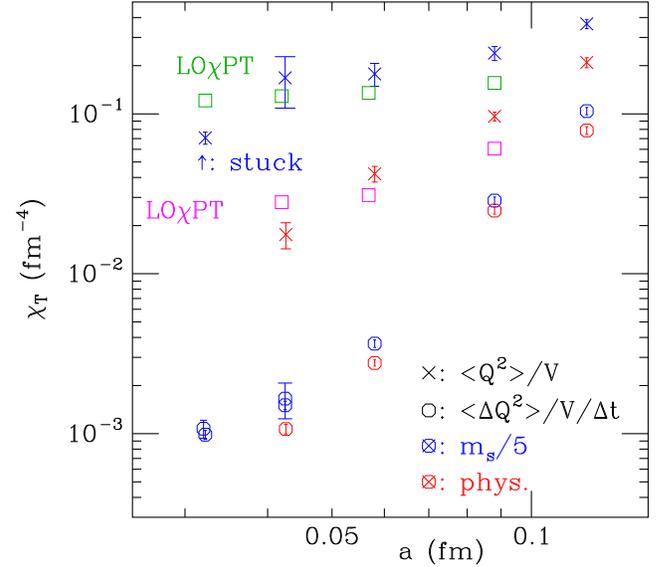

FIG. 3. Average topological susceptibility $\langle Q^2/V \rangle$ (crosses) and tunneling rate per unit volume $\langle (\Delta Q)^2/(V\Delta t) \rangle$ (octagons) versus lattice spacing. Red and blue points are results for ensembles with $m_l = m_s/5$ and m_l approximately physical, respectively. Similarly, the magenta and green squares are the lowest-order chiral perturbation theory predictions for the susceptibility at $m_l = m_s/5$ and m_l at its physical value, $\approx m_s/27$. The leading-order χPT results shown include staggered corrections and are taken from Eq. (95) below.

physical m_l ensembles. The tunneling rate does not depend strongly on the quark mass, but decreases as expected as the lattice spacing gets small. (In the cases where there are two blue octagons, there were two subensembles with different molecular dynamics trajectory lengths.) The crosses in Fig. 3 show the topological susceptibility, $\langle Q^2/V \rangle$. Here we see the expected strong dependence on the light quark mass. The small error bar on the 0.03 fm point is unrealistic; it simply reflects the fact that Q is basically stuck near this value in this simulation.

III. QUANTITIES AT FIXED Q

In this section, we outline the relation between the behavior of physical quantities at fixed topological charge and their dependence on the vacuum angle θ. The discussion relies heavily on that in Ref. [6]. For nonzero θ, the partition function is

$$Z(\theta) = \int \mathcal{D}A \mathcal{D}\bar{\Psi} \mathcal{D}\Psi \exp(-S[A, \bar{\Psi}, \Psi])$$
$$\times \exp(-i\theta Q[A]). \quad (2)$$

The topological susceptibility χ_T is defined by [18–20]

$$\chi_T \equiv -\frac{1}{V}\left(\frac{1}{Z}\frac{\partial^2 Z}{\partial \theta^2}\right)\bigg|_{\theta=0} = \frac{1}{V}\langle Q^2 \rangle. \quad (3)$$

We assume that both the time extent T of our system and the three-dimensional volume V_3 are large, so the four-dimensional volume $V = TV_3$ is also large. The partition function is then dominated by the vacuum energy density $\epsilon_0(\theta)$,

$$Z(\theta) \approx C \exp(-V\epsilon_0(\theta)), \tag{4}$$

$$\epsilon_0 = \frac{1}{2}\chi_T\theta^2 + \gamma\theta^4 + \cdots, \tag{5}$$

where C is a constant. The fact that χ_T is the coefficient of the quadratic term follows from Eq. (3). Parity symmetry (or, more precisely, *extended parity*; see below) implies that only even powers of θ appear in Eq. (5).

Quantities evaluated at fixed Q can be found by Fourier transforming

$$Z_Q = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta \exp(i\theta Q) Z(\theta), \tag{6}$$

$$\begin{aligned} G_Q &= \langle \mathcal{O}_1 \mathcal{O}_2 ... \mathcal{O}_n \rangle_Q \\ &= \frac{1}{Z_Q} \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta \exp(i\theta Q) Z(\theta) G(\theta), \end{aligned} \tag{7}$$

with $G(\theta) = \langle \mathcal{O}_1 \mathcal{O}_2 ... \mathcal{O}_n \rangle_\theta$. Since V is large, we can do the θ integrals by the saddle-point method. Using Eqs. (4) and (5), the saddle occurs at

$$\theta_s = i\frac{Q}{\chi_T V} + \mathcal{O}\left(\frac{Q^3}{V^3}\right). \tag{8}$$

This gives

$$G_Q = G(\theta_s) + \frac{1}{2\chi_T V} \frac{\partial^2 G}{\partial\theta^2}\bigg|_{\theta=\theta_s} + \cdots, \tag{9}$$

which in turn implies [6,21]

$$B|_{Q,V} = B + \frac{1}{2\chi_T V} B'' \left(1 - \frac{Q^2}{\chi_T V}\right) + \mathcal{O}\left(\frac{1}{(\chi_T V)^2}\right), \tag{10}$$

where B is the mass M or the decay constant f, and the primes here and below indicate derivatives with respect to θ evaluated at $\theta = 0$. The terms B and $-(B''/2)(Q/\chi_T V)^2$ on the right-hand side of Eq. (10) come from expanding $G(\theta_s)$ in Eq. (9), while the term $B''/(2\chi_T V)$ comes from the term proportional to $\partial^2 G/\partial\theta^2$. By Eq. (3), the correction to B vanishes when averaged over Q. Equations (7), (9) and (10) are valid under the assumption $Q \sim \sqrt{\chi_T V}$, i.e., a "typical" value in a random walk of Q around $Q = 0$ with $\langle Q^2 \rangle = \chi_T V$.

Here and below we use the fact that $B' = 0$. This is true for any parity-conserving quantity, since, although parity symmetry is broken at the QCD level by the θ term, what we can call *extended parity* is preserved. In extended parity,

we take $\theta \to -\theta$ along with a normal parity transformation. This means that a matrix element that does not violate parity at $\theta = 0$ must be even in θ, and therefore its first derivative at $\theta = 0$ vanishes. For masses and decay constants, the χPT results will confirm this.

IV. CHIRAL PERTURBATION THEORY: UNITARY CASE

The quantity B'' ($B = M$ or f) appearing in Eq. (10) is physical and therefore has dependence on the three-dimensional volume $V_3 = L^3$ that is exponentially suppressed, $\propto \exp(-M_\pi L)$. This is in contrast with the quantity at fixed Q, $B|_{Q,V}$, which has a power-law dependence on $1/V$. We may thus get a handle on topological effects by calculating B'' in infinite-volume χPT. For now, we ignore possible discretization errors and consider χPT in the continuum only. Corrections due to discretization effects for staggered quarks are calculated in Sec. VII. For two dynamical flavors, a first calculation of M'' in χPT for full (unitary) QCD in infinite volume and the continuum appears in Ref. [6].

In the presence of a vacuum angle θ, the leading-order (LO) Euclidean chiral Lagrangian is

$$\mathcal{L}_\chi = \frac{f^2}{8}\text{tr}(\partial_\mu\Sigma\partial_\mu\Sigma^\dagger) - \frac{B_0 f^2}{4}\text{tr}(\mathcal{M}_A^*\Sigma + \mathcal{M}_A\Sigma^\dagger), \tag{11}$$

where the normalization is such that $f \approx 130$ MeV, and where $\mathcal{M}_A \equiv e^{i\theta/N}\mathcal{M}$, where N is the number of flavors and $\mathcal{M}$ is the mass matrix in the absence of θ. Complex conjugation is denoted by $*$. We always take $\mathcal{M}$ to be diagonal in this paper. The change from $\mathcal{M}$ to $\mathcal{M}_A$ is effected by an anomalous (flavor-singlet) chiral transformation, which simultaneously removes the $i\theta Q$ from the Euclidean gauge action. See Appendix A for more discussion of the phases in the mass matrix.

In order to set out notation and make various points that will be useful later, we examine both the two-flavor and the three-flavor unitary cases in detail. We emphasize that none of the results in this section are new to the literature; while the authors of Ref. [6] did not discuss decay constants or the $N = 3$ case, those results can easily be obtained as limits of the general partially quenched results given in Ref. [7].

A. Two flavors

We consider the $N = 2$ case with nondegenerate quark masses $m_u \neq m_d$. With $\theta \neq 0$, Σ can have a nontrivial vacuum expectation value, $\langle\Sigma\rangle$. When the quark masses are nondegenerate, an argument for arbitrary N by Gasser and Leutwyler [22] shows that $\langle\Sigma\rangle$ must be diagonal.[1]

---

[1]If there are degeneracies, then $\langle\Sigma\rangle$ can be put in diagonal form by making a vector (flavor) rotation that leaves the chiral Lagrangian unchanged.

The intuitive reason for this is that only the diagonal elements of Σ enter into the potential energy, which has an overall negative sign in Eq. (11). Since Σ is unitary, off-diagonal elements would reduce the absolute size of the diagonal elements and result in a higher potential energy.

We therefore let

$$\langle \Sigma \rangle = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix}, \tag{12}$$

where the diagonal elements are constrained by $\det \Sigma = 1$. The potential energy term we need to minimize is then

$$V = -\frac{B_0 f^2}{2}(m_u \cos(\alpha - \theta/2) + m_d \cos(\alpha + \theta/2)). \tag{13}$$

Differentiating with respect to α gives the condition

$$m_u \sin(\alpha - \theta/2) + m_d \sin(\alpha + \theta/2) = 0, \tag{14}$$

with the solution

$$\tan(\alpha) = r \tan(\theta/2), \qquad r \equiv \frac{m_u - m_d}{m_u + m_d}. \tag{15}$$

We can now expand the potential energy to quadratic order to find the pion mass. The meson field Φ, which characterizes fluctuations around the vacuum expectation value, is defined most conveniently by

$$\Sigma = s e^{2i\Phi/f} s, \qquad s \equiv \sqrt{\langle \Sigma \rangle}, \tag{16}$$

$$\Sigma^\dagger = s^\dagger e^{-2i\Phi/f} s^\dagger. \tag{17}$$

The field Φ may be written as usual in terms of individual meson fields as

$$\Phi = \begin{pmatrix} \pi^0/\sqrt{2} & \pi^+ \\ \pi^- & -\pi^0/\sqrt{2} \end{pmatrix}. \tag{18}$$

The definition of Φ in Eq. (16) is convenient for two reasons. First of all, it transforms normally under extended parity (usual parity plus $\theta \to -\theta$):

$$\Phi \to -\Phi, \tag{19}$$

$$s \to s^\dagger, \tag{20}$$

$$\Sigma \to \Sigma^\dagger. \tag{21}$$

This means that the fields $\pi^\pm$, π^0 have the usual interpretation as pion fields. [Note that Eq. (20) follows from the fact that $\alpha \to -\alpha$ when $\theta \to -\theta$.] The chiral Lagrangian, Eq. (11), is easily seen to be invariant under extended

parity, as expected. Second, with the definition (16), the kinetic energy term in Eq. (11) takes the same form in terms of Φ as it does in the standard case when $\langle \Sigma \rangle = I$. This means that there is no wave-function renormalization at leading order, which simplifies calculations. An alternative definition,

$$\Sigma = e^{i\Phi/f}\langle \Sigma \rangle e^{i\Phi/f}, \tag{22}$$

also leads to Eqs. (19) and (21) under extended parity, but generates a nontrivial leading-order wave-function renormalization factor. Of course, physical results must be the same with any appropriate field definition. It is straightforward to check that the meson mass and decay constant results below, found in the first instance using Eq. (16), can also be obtained with Eq. (22).

Expanding to quadratic order in Φ, we find, for the charged pion mass,

$$M_\pi^2(\theta) = M_\pi^2(0)\cos\left(\frac{\theta}{2}\right)\sqrt{1 + r^2\tan^2\left(\frac{\theta}{2}\right)}, \tag{23}$$

where $M_\pi^2(0) = B_0(m_u + m_d)$. Equation (23) agrees with Eq. (4.15) of Ref. [6].

For the decay constant we need the axial current $\mathcal{A}_\mu^{ij}$ in χPT that corresponds to the QCD current $\bar{q}^j\gamma_\mu\gamma_5 q^i$, where i, j are flavor indices. With the definitions of Ref. [22],

$$\mathcal{A}_\mu^{ij} = i\frac{f^2}{4}(\partial_\mu\Sigma\Sigma^\dagger + \Sigma^\dagger\partial_\mu\Sigma)^{ij}. \tag{24}$$

Note that the axial current comes from the kinetic energy term in the chiral Lagrangian, and its form in terms of Σ is unaffected by a nonzero θ. For f_π, we need $\mathcal{A}_\mu^{12}$. Plugging in Eq. (16), gives, to leading order,

$$\mathcal{A}_\mu^{12} = -f\cos(\alpha)\partial_\mu\pi^+. \tag{25}$$

With Eq. (15), this implies

$$f_\pi(\theta) = f\cos(\alpha) = f_\pi(0)\cos(\alpha)$$
$$= \frac{f_\pi(0)}{\sqrt{1 + r^2\tan^2(\theta/2)}}. \tag{26}$$

Note that f_π is independent of θ in the degenerate case, $r = 0$. To apply Eq. (10), we need the second derivative of M_π or f_π at $\theta = 0$. From Eqs. (23) and (26), we obtain

$$M_\pi'' = M_\pi(0)\frac{r^2 - 1}{8} = -M_\pi(0)\frac{m_u m_d}{2(m_u + m_d)^2}, \tag{27}$$

$$f_\pi'' = -f_\pi(0)\frac{r^2}{4} = -f_\pi(0)\frac{(m_u - m_d)^2}{4(m_u + m_d)^2}. \tag{28}$$

As expected, the first derivatives, M'_π and f'_π, vanish. Note that f''_π does not vanish when one mass (m_u, say) goes to zero. This seems to contradict the expectation that the theory is θ independent when one quark mass vanishes. However, the θ independence does not apply to quantities, such as f_π, that depend on external currents. We explain this in more detail in Appendix A.

B. Three flavors

For $N = 3$, we will work in the limit $m_u = m_d \equiv m$, but $m \neq m_s$ in general. Since isospin is preserved, we can assume

$$\langle \Sigma \rangle = \begin{pmatrix} e^{i\alpha} & 0 & 0 \\ 0 & e^{i\alpha} & 0 \\ 0 & 0 & e^{-2i\alpha} \end{pmatrix}. \qquad (29)$$

The potential energy term is then

$$V = -\frac{B_0 f^2}{2}(2m\cos(\alpha - \theta/3) + m_s \cos(2\alpha + \theta/3)). \qquad (30)$$

Differentiating with respect to α gives the condition

$$2m\sin(\alpha - \theta/3) + 2m_s \sin(2\alpha + \theta/3) = 0. \qquad (31)$$

Although Eq. (31) does not have a simple analytic solution, we really only need derivatives of quantities at $\theta = 0$, which can all be calculated by implicit differentiation. Note first that the solution for $\langle \Sigma \rangle$ is invariant under $\theta \to -\theta$, $\alpha \to -\alpha$, so the solution $\alpha(\theta)$ has only odd powers of θ. In particular, the second derivative of $\alpha(\theta)$ at $\theta = 0$ vanishes: $\alpha'' = 0$. We then write a physical quantity $W(\theta)$ as $W(\alpha(\theta), \theta)$, where the second argument is the explicit θ dependence, and the first is the dependence through α. Using $\alpha'' = 0$, we have

$$W'' \equiv \frac{d^2 W}{d\theta^2}\bigg|_{\theta=0} = \frac{\partial^2 W}{\partial \alpha^2}(\alpha')^2 + 2\frac{\partial^2 W}{\partial \alpha \partial \theta}\alpha' + \frac{\partial^2 W}{\partial \theta^2}, \qquad (32)$$

where all derivatives on the right-hand side are to be evaluated at $\theta = 0$ (which implies $\alpha = 0$). Thus, all we need from the solution to Eq. (31) is α', which is easily calculated to be

$$\alpha' = \frac{m - m_s}{3(m + 2m_s)}. \qquad (33)$$

We now merely need to find the masses and decay constants as functions of α and θ. Expanding the potential energy term in Eq. (11) to quadratic order in the now 3×3 meson matrix Φ, we find

$$M_\pi^2 = B_0(2m\cos(\alpha - \theta/3)),$$
$$M_K^2 = B_0(m\cos(\alpha - \theta/3) + m_s \cos(2\alpha + \theta/3)). \qquad (34)$$

Similarly, using Eq. (16) and expanding Eq. (24) to linear order in Φ gives

$$f_\pi = f, \qquad f_K = f\cos(3\alpha/2). \qquad (35)$$

Using Eq. (32) now gives

$$M''_\pi = -M_\pi(0)\frac{m_s^2}{2(m + 2m_s)^2}, \qquad (36)$$

$$M''_K = -M_K(0)\frac{mm_s}{2(m + 2m_s)^2}, \qquad (37)$$

$$f''_\pi = 0, \qquad (38)$$

$$f''_K = -f_K(0)\frac{(m_s - m)^2}{4(m + 2m_s)^2}. \qquad (39)$$

Like f''_π in the two-flavor case, f''_K does not vanish when one of the masses goes to zero; see Appendix A for an explanation.

V. PARTIALLY QUENCHED χPT

Since most of our lattice data is partially quenched, we need to extend the calculation of M'' and f'' to PQχPT. This was done by Aoki and Fukaya [7] using the replica method to remove the determinant of the valence quarks. However, the required calculation is nonperturbative, at least on its face, since the vacuum state changes in the presence of θ. The replica method is only justified perturbatively, so a nonperturbatively safe method is preferable. The Lagrangian approach of Ref. [23], which introduces ghost (bosonic) quarks to cancel the valence-quark determinant, is also only valid perturbatively, since it ignores the requirement that the bosonic path integral be convergent.

The nonperturbatively correct version of PQχPT has been worked out by Golterman, Sharpe, and Singleton [9] and Sharpe and Shoresh [10]. The nonperturbative problems of the Lagrangian approach are fixed by taking into account the convergence requirement.

In terms of Σ, the chiral Lagrangian is

$$\mathcal{L}_{\chi,PQ} = \frac{f^2}{8}\text{str}(\partial_\mu \Sigma \partial_\mu \Sigma^{-1})$$
$$- \frac{B_0 f^2}{4}\text{str}(\mathcal{M}\Sigma + \mathcal{M}\Sigma^{-1}), \qquad (40)$$

where "str" is the supertrace. The main difference from the standard perturbative version of PQχPT [23] is that Σ is not unitary, which is why Σ^{-1} appears instead of $\Sigma^\dagger$ in Eq. (40).

For definiteness, we work with three sea quarks ($N = 3$) and two valence quarks ($N_v = 2$), and take the isospin limit in the sea: $m_u = m_d \equiv m$, but $m \neq m_s$ in general. The valence quarks are x and y with masses m_x and m_y, respectively. Corresponding ghost quarks $\tilde{x}$ and $\tilde{y}$ with masses m_x and m_y are included to cancel the valence-quark determinant. The chiral field Σ and the quark mass matrix $\mathcal{M}$ are then 7×7 matrices. The mass matrix is given by

$$\mathcal{M} = \mathrm{diag}(m, m, m_s, m_x, m_y, m_x, m_y), \quad (41)$$

where the quarks are ordered (sea, valence, ghost).

In expanding Σ in terms of pseudoscalar meson fields, it is useful to separate out a special diagonal meson field, ϵ, which is a linear combination of flavor-neutral quark-antiquark and ghost-antighost mesons. We write

$$\Sigma = \exp(2i\Phi/f), \quad (42)$$

$$\Phi = \Phi' + i\epsilon T_6, \quad (43)$$

$$\Phi' = \begin{pmatrix} \phi & \eta \\ \eta^\dagger & i\tilde{\phi} \end{pmatrix}, \quad (44)$$

where ϕ is the quark-antiquark block (both sea and valence), $\tilde{\phi}$ is the ghost-antighost block, and η and $\eta^\dagger$ are the quark-antighost and ghost-antiquark blocks, respectively. The diagonal generator[2] T_6 is nonanomalous (straceless), and is given by

$$T_6 = \sqrt{\frac{2}{15}}\mathrm{diag}(1, 1, 1, 1, 1, 5/2, 5/2). \quad (45)$$

Both ϕ and $\tilde{\phi}$ are Hermitian and traceless. The factor of i in the ghost-antighost ($\tilde{\phi}$) block in Eq. (44) comes ultimately from the careful consideration of the true symmetries of the theory with ghosts. These are more complicated than those assumed in Ref. [23] because of the necessity of keeping the ghost (bosonic) path integrals convergent in a non-perturbative treatment. At the chiral level, the integrals over the independent real fields in $\tilde{\phi}$ run from $-\infty$ to ∞, and the factor of i multiplying $\tilde{\phi}$ in Eq. (44), in combination with the supertrace, guarantees that the action for these fields is positive definite, so the kinetic and mass terms have the proper sign for convergence of the path integral. There is no problem with the convergence of the ϕ and η, $\eta^\dagger$ integrals because the former are over a compact region (they are angles), and the latter are Grassmann variables.

In Eq. (43), we have followed the prescription of Sharpe and Shoresh [10] and have included a factor of i with the field ϵ multiplying T_6. Because $\mathrm{str}(T_6^2) < 0$, the i is

necessary in order for the kinetic energy of ϵ to be positive. In other words, ϵ is "ghost-like," rather than "quark-like." Like $\tilde{\phi}$, ϵ should be integrated along the entire real axis. Note that the only other linearly independent diagonal generator to span the quark-antiquark and ghost-antighost blocks is

$$T_7 = \frac{1}{\sqrt{3}}I = \frac{1}{\sqrt{3}}\mathrm{diag}(1, 1, 1, 1, 1, 1, 1), \quad (46)$$

where I is the identity matrix. Since T_7 is anomalous [$\mathrm{str}(T_7) = \sqrt{3}$], the corresponding meson (called Φ_0, or less precisely, η') is heavy and is integrated out of the chiral theory.

When including the θ angle, the most natural approach would be to remove the $\theta F\tilde{F}$ term by making an anomalous rotation using T_7 as a generator. At the chiral level, this would put a factor of $\exp(-i\theta/3)$ in front of the $\mathrm{str}(\mathcal{M}\Sigma)$ term in Eq. (40), and a factor of $\exp(i\theta/3)$ in front of the $\mathrm{str}(\mathcal{M}\Sigma^{-1})$ term. In other words, all quarks (sea and valence) and all ghosts would get the same θ phase. However, it is convenient to make an additional non-anomalous rotation with the generator

$$t = \mathrm{diag}(0, 0, 0, 1, 1, 1, 1), \quad (47)$$

which is a linear combination of T_6 and a generator in the quark-antiquark block, namely $\mathrm{diag}(1, 1, 1, -3/2, -3/2, 0, 0)$. This allows us to remove the θ-dependent phase from all mass terms of valence and ghost quarks, which makes the algebra somewhat simpler and has further advantages for the heavy-light case discussed in Sec. VI. The partially quenched chiral Lagrangian in the presence of θ is then

$$\mathcal{L}_{\chi,\mathrm{PQ},\theta} = \frac{f^2}{8}\mathrm{str}(\partial_\mu\Sigma\partial_\mu\Sigma^{-1})$$
$$- \frac{B_0 f^2}{4}\mathrm{str}(\mathcal{M}_B^*\Sigma + \mathcal{M}_B\Sigma^{-1}), \quad (48)$$

where

$$\mathcal{M}_B \equiv \mathrm{diag}(e^{i\theta/3}m, e^{i\theta/3}m, e^{i\theta/3}m_s, m_x, m_y, m_x, m_y). \quad (49)$$

In Appendix A, several choices for the mass matrix in the presence of θ are discussed; we include the subscript B on $\mathcal{M}_B$ for consistency with notation introduced there.

In the unitary theory, the absolute minimum of the potential energy term determines the vacuum state $\langle\Sigma\rangle$. Here, the potential energy V is complex. The authors of Ref. [9] argued that we should therefore find a saddle point of $|\exp(-V)| = \exp(-\mathrm{Re}(V))$ (deforming the $\tilde{\phi}$ and ϵ contours as needed), not a minimum. One issue that arises in the saddle-point analysis is how to choose the proper saddle point when the complex saddle-point equation,

---

[2]We follow the notation of Ref. [10] in naming the diagonal generators.

$V' = 0$ has multiple solutions, as it does here. Fortunately, de Bruijn [24] (referred to by Ref. [9]) gave a prescription for finding the unique useful saddle point for an analytic function like V: find 1), a point that is a solution of $V' = 0$, and 2), a deformation of the contour that goes through the point in the direction of steepest descent of $|\exp(-V)|$ [steepest ascent of $\mathrm{Re}(V)$] and satisfies the requirement that the point has the highest value of $|\exp(-V)|$ of any place on the deformed contour. There is at most one saddle point that satisfies these conditions, so once we find one such point, we are guaranteed to have found the unique solution, which determines $\langle\Sigma\rangle$.

The problem of solving $V' = 0$ is simplified by noting that, as in the unitary theory, $\langle\Sigma\rangle$ is diagonal. This can be proved by following the unitary-theory argument in Ref. [22]. For the case when all masses are nondegenerate, the argument goes through with only trivial modifications. Degeneracies among sea quarks, or between valence and sea quarks, also present no problem because Σ is a unitary matrix with c-number entries in these blocks, and can be diagonalized exactly as in the unitary theory. However, degeneracy between valence and ghost quarks must be considered because such degeneracies are built into the partially quenched theory. These degeneracies are different than those among quarks since the graded structure of the group is crucial. It is plausible that the block of $\langle\Sigma\rangle$ corresponding to the degenerate mass pair can be diagonalized by a vector similarity transformation, $\Sigma \to U\Sigma U^{-1}$, where U is an element of the graded symmetry group $SL(N + N_v|N_v)$ [i.e., $SL(5|2)$ here]. Such a transformation leaves the Lagrangian, including the mass (potential energy) term, unchanged. We have checked the diagonalization explicitly for the crucial twofold degeneracy of quarks and ghosts, and believe it must also be true if there is a higher degeneracy (e.g., $m_x = m_y$ also), but have not proved it.

Following Eqs. (42) through (44), we therefore parametrize $\langle\Sigma\rangle$ as

$$\langle\Sigma\rangle = \exp(i\,\mathrm{diag}(\alpha+\delta+i\epsilon, \alpha+\delta+i\epsilon, -2\alpha+\delta+i\epsilon,$$
$$\beta-3\delta/2+i\epsilon, -\beta-3\delta/2+i\epsilon, 5i\epsilon/2-i\gamma, 5i\epsilon/2+i\gamma)),$$
$$\tag{50}$$

where we have ensured that the exponent is straceless, and have used isospin symmetry ($m_u = m_d \equiv m$) to require that the first two entries along the diagonal be equal. We have chosen simpler normalization for the angles in Eq. (50) than we would need to use for the corresponding meson fields.

Since we will deform the contour for the ghost-antighost fields, the variables γ and ϵ may be complex. For convenience we define $\epsilon \equiv i\hat{\epsilon}$, $\gamma = i\hat{\gamma}$, where it will turn out that $\hat{\gamma}$ and $\hat{\epsilon}$ are in fact real at the saddle point. With this definition, the potential energy is

$$V = -\frac{B_0 f^2}{2}\{2m\cos(\theta/3 - \alpha - \delta + \hat{\epsilon})$$
$$+ m_s\cos(\theta/3 + 2\alpha - \delta + \hat{\epsilon})$$
$$+ m_x\cos(3\delta/2 - \beta + \hat{\epsilon}) + m_y\cos(3\delta/2 + \beta + \hat{\epsilon})$$
$$- m_x\cos(5\hat{\epsilon}/2 - \hat{\gamma}) - m_y\cos(5\hat{\epsilon}/2 + \hat{\gamma})\}. \tag{51}$$

From the requirement that V is stationary at the saddle point with respect to $\alpha, \beta, \hat{\gamma}, \delta$ and $\hat{\epsilon}$, respectively, we obtain the equations

$$m\sin(\theta/3 - \alpha - \delta + \hat{\epsilon}) - m_s\sin(\theta/3 + 2\alpha - \delta + \hat{\epsilon}) = 0, \tag{52}$$

$$m_x\sin(3\delta/2 - \beta + \hat{\epsilon}) - m_y\sin(3\delta/2 + \beta + \hat{\epsilon}) = 0, \tag{53}$$

$$m_x\sin(5\hat{\epsilon}/2 - \hat{\gamma}) - m_y\sin(\hat{\gamma} + 5\hat{\epsilon}/2) = 0, \tag{54}$$

$$2m\sin(\theta/3 - \alpha - \delta + \hat{\epsilon}) + m_s\sin(\theta/3 + 2\alpha - \delta + \hat{\epsilon})$$
$$- \frac{3}{2}m_x\sin(3\delta/2 - \beta + \hat{\epsilon}) - \frac{3}{2}m_y\sin(3\delta/2 + \beta + \hat{\epsilon}) = 0, \tag{55}$$

$$2m\sin(\theta/3 - \alpha - \delta + \hat{\epsilon}) + m_s\sin(\theta/3 + 2\alpha - \delta + \hat{\epsilon})$$
$$+ m_x\sin(3\delta/2 - \beta + \hat{\epsilon}) + m_y\sin(3\delta/2 + \beta + \hat{\epsilon})$$
$$- \frac{5}{2}m_x\sin(5\hat{\epsilon}/2 - \hat{\gamma}) - \frac{5}{2}m_y\sin(\hat{\gamma} + 5\hat{\epsilon}/2) = 0. \tag{56}$$

In the case $\theta = 0$, we have the standard perturbative solution $\alpha = \beta = \hat{\gamma} = \delta = \hat{\epsilon} = 0$, so $\langle\Sigma\rangle = I$. If the valence masses m_x and m_y are not too small, it is easy to check that this saddle point is the correct one to use because $\mathrm{Re}(V)$ increases monotonically away from the saddle on the original contours, on which ϵ and γ are real. However, Sharpe and Shoresh [10] found a lower bound on the valence masses, below which the real part of the squared-mass matrix of the ghost-like neutral fields is not positive definite. When the valence masses violate the bound, $\mathrm{Re}(V)$ decreases away from the saddle in a real direction on one of the contours. At first glance, this suggests that the perturbative vacuum is not the correct one in this case. We discuss the issue in detail in Appendix B, and show that the monotonic increase away from the perturbative saddle point is restored after one or more of the neutral quark-antiquark integrals are performed. This means that we may freely violate the Sharpe-Shoresh bound, and the perturbative vacuum is the correct one for any nonzero values of the valence-quark masses.

For θ not too far from 0, we expect that the proper saddle point is then the nearby one, where the magnitudes of the arguments of all the sine functions are less than $\pi/2$. If this were not true, it would invalidate the analysis that led to Eq. (10), since we assumed a smooth dependence on θ.

Nevertheless, to be sure our analysis is correct nonperturbatively, we will check this assumption below.

With the arguments bounded by assumption, two sine functions are equal if and only if their arguments are equal. Subtracting Eq. (55) from Eq. (56) to eliminate terms with m and m_s and then using Eqs. (53) and (54) to eliminate terms with m_x or m_y then implies $\hat{\epsilon} = \delta$ and $\hat{\gamma} = \beta$. The saddle-point value $\langle \Sigma \rangle$ simplifies to

$$\langle \Sigma \rangle = \exp(i\,\mathrm{diag}(\alpha, \alpha, -2\alpha, \beta - 5\delta/2, -\beta - 5\delta/2,$$
$$\beta - 5\delta/2, -\beta - 5\delta/2)). \tag{57}$$

This has two required features of a partially quenched theory: 1) the sea-quark sector is unaffected by the presence of valence quarks and ghosts, and 2) the vacuum expectation values of $\bar{q}q$ and $\bar{g}g$ are equal for corresponding quarks (q) and ghosts (g).

Plugging the results for $\hat{\epsilon}$ and $\hat{\gamma}$ into the saddle-point equations (52)–(56) then determines the remaining variables, α, β and δ. It is not necessary to obtain a closed-form solution. As in the unitary three-flavor case, we only need α', β' and δ', the derivatives of these angles with respect to θ at $\theta = 0$. By differentiating the saddle-point equations and solving, we find

$$\alpha' = \frac{m - m_s}{3(m + 2m_s)}, \tag{58}$$

$$\beta' = \frac{m_x - m_y}{2m_x m_y} \cdot \frac{m m_s}{m + 2m_s} = \hat{\gamma}', \tag{59}$$

$$\delta' = \frac{m_x + m_y}{5m_x m_y} \cdot \frac{m m_s}{m + 2m_s} = \hat{\epsilon}'. \tag{60}$$

As expected, the angle governing the sea-quark vacuum expectation value, α, obeys the same equation as in the unitary QCD case (33).

Before proceeding, we should check that the saddle we have found is the proper one to use. At the saddle point, $\hat{\epsilon}$ and $\hat{\gamma}$ (the imaginary parts of ϵ and γ) are equal to δ and β respectively, which are comparable to θ and hence small angles. If the Sharpe-Shoresh bound is satisfied, the steepest descent directions, in which $\mathrm{Re}(V)$ increases most rapidly away from the saddle point, are the real directions for ϵ and γ. As we continue the contours for ϵ and γ in these directions, $\mathrm{Re}(V)$ increases exponentially, dominated by one or both of the ghost terms that grow like $\cosh(\mathrm{Re}(5\epsilon/2 \pm \gamma))$. Far from the saddle point, it is then straightforward to see that we can bend the contours back to the real axis while keeping $\mathrm{Re}(V)$ large, i.e., much larger than at the saddle point. We have therefore found a proper saddle point and contours. Appendix B argues that the saddle and contours are still the correct ones when the Sharpe-Shoresh bound is violated.

Using Eqs. (58)–(60) to calculate the valence-meson mass and decay constant as in Sec. IV B, we find

$$M''_{xy} = -M_{xy}(0)\frac{m^2 m_s^2}{(m + 2m_s)^2}\frac{1}{2m_x m_y}, \tag{61}$$

$$f''_{xy} = -f_{xy}(0)\frac{m^2 m_s^2}{(m + 2m_s)^2}\frac{(m_x - m_y)^2}{4m_x^2 m_y^2}. \tag{62}$$

The results for the unitary pion and kaon, Eqs. (36), (37), (38) and (39), can be obtained from Eqs. (61) and (62) in the appropriate limits: $m_x = m_y = m$ for the pion and $m_x = m$, $m_y = m_s$ for the kaon.

Equations (61) and (62) have singular limits when m_x or m_y or both go to zero at fixed sea-quark masses. (For f_{xy} one has to keep $m_x \neq m_y$ if both go to zero to get a nonzero result.) Such mass singularities are typical for the partially quenched theory, but this is the only case we know of where they appear at tree level. On the other hand, the results vanish when either sea-quark mass goes to zero, as explained in Appendix A.

The results in Eqs. (61) and (62) agree with those computed by Aoki and Fukaya [7], who used the replica method for the partially quenched theory. Because the replica method has not been nonperturbatively justified, the methods of Refs. [9,10] seem preferable to us here, since the ground state of the theory is changing. The agreement of the two methods suggests, though, that this particular problem is essentially perturbative. This makes sense because in the end we only need derivatives of quantities at $\theta = 0$, so the dependence on θ is required only in an infinitesimal neighborhood of the perturbative vacuum. It has been a surprise to us that the more vexing nonperturbative issue in our analysis arises from the Sharpe-Shoresh bound, which already affects the $\theta = 0$ case.

VI. HEAVY-LIGHT MESONS

We now add a heavy quark Q to the theory. It is useful to consider the heavy quark in a partially quenched context: let its valence mass be m_Q and its sea mass be $m_{Q,\mathrm{sea}}$. In the presence of a nonzero θ, we put the θ dependence into the sea-quark (but not the valence-quark) mass matrix, as in $\mathcal{M}_B$ [Eq. (49)]. As both m_Q and $m_{Q,\mathrm{sea}}$ get large, the heavy sea quark decouples, and we are left with a theory of light sea quarks only. The valence heavy quark of course does not decouple, since it can appear in external states, but it carries no θ-dependent phases. Note that it does not make any physical difference how the θ dependence is put into the initial sea-quark mass matrix, i.e., whether or not the heavy sea-quark mass carries θ dependence. The end result after $m_{Q,\mathrm{sea}} \to \infty$ is always the same as if we had started with a theory of only light sea quarks. However the decoupling is indeed more subtle when the heavy sea quark carries θ dependence: we cannot simply delete the

heavy sea-quark terms from the Lagrangian. See Appendix C for a discussion of how decoupling works in that case.

The leading-order heavy-meson chiral Lagrangian is then exactly the standard one [25]:

$$\mathcal{L}_{\chi,\mathrm{HL}} = -is\mathrm{Tr}(\bar{H}Hv \cdot \overleftarrow{D}) + g_\pi s\mathrm{Tr}(\bar{H}H\gamma^\mu\gamma_5\mathbb{A}_\mu), \quad (63)$$

where H is the heavy-light meson field, composed of a pseudoscalar meson P and a vector meson P^*,

$$H_a = \frac{1 + \rlap{/}{v}}{2}[\gamma^\mu P^*_{\mu a} + i\gamma_5 P_a], \quad (64)$$

where v is the meson's velocity, and a is the flavor index of the light quark. In Eq. (63), $\overleftarrow{D}$ is the covariant derivative (acting to the left), sTr is a trace over Dirac indices and a supertrace over flavor indices,[3] and $\mathbb{A}_\mu$ is the light-quark axial current,

$$\mathbb{A}_\mu = \frac{i}{2}[\sigma^\dagger\partial_\mu\sigma - \sigma\partial_\mu\sigma^\dagger], \quad (65)$$

$$\sigma \equiv \sqrt{\Sigma}. \quad (66)$$

The leading-order left-handed current that destroys a heavy-light meson with light flavor b is [25]

$$j_L^{\mu,b} = \frac{\kappa}{2}\mathrm{tr}_D(\gamma^\mu(1 - \gamma_5)H)\sigma^\dagger\lambda^{(b)}, \quad (67)$$

where κ is a low-energy constant, tr_D is a trace over Dirac indices only, and $\lambda^{(b)}$ is a constant column vector that fixes the flavor of the light quark: $(\lambda^{(b)})_a = \delta_{ab}$. For the decay constant, we need the heavy-light axial current

$$\begin{aligned}j_5^{\mu,b} &= \frac{1}{2}(j_R^{\mu,b} - j_L^{\mu,b}) \\ &= \frac{\kappa}{4}[\mathrm{tr}_D(\gamma^\mu(1 + \gamma_5)H)\sigma - \mathrm{tr}_D(\gamma^\mu(1 - \gamma_5)H)\sigma^\dagger]\lambda^{(b)},\end{aligned}$$
$$(68)$$

where the right-handed current $j_R^{\mu,b}$ can be found from the left-handed current using parity ($H \to \gamma_0 H\gamma_0$, $\sigma \to \sigma^\dagger$). We obtain the decay constant, or more precisely $\Phi \equiv f\sqrt{M}$, from the relation

$$\langle 0|j_5^{\mu,a}|P_a\rangle = -iv^\mu\Phi_a \quad (\text{no sum on } a), \quad (69)$$

which implies $\Phi_a = \kappa$ to leading order when $\theta = 0$.

---

[3]The supertrace is used because the theory is partially quenched. In most cases, however, the difference between the trace and supertrace is irrelevant for the heavy-light part of the Lagrangian, since closed heavy-light meson loops are forbidden anyway.

When $\theta \neq 0$, it affects Φ through the expectation value of σ in Eq. (68): $\langle\sigma\rangle = \sqrt{\langle\Sigma\rangle} \neq I$. Using Eq. (57) for $\langle\Sigma\rangle$ gives

$$\Phi_x(\theta) = \Phi_x(\theta = 0)\cos\left(\frac{\beta}{2} - \frac{5\delta}{4}\right), \quad (70)$$

for a light valence quark x. Equations (59) and (60) then imply

$$\Phi_x'' = -\Phi_x(0)\frac{m^2 m_s^2}{(m + 2m_s)^2}\frac{1}{4m_x^2}. \quad (71)$$

This smoothly connects to the light quark result for f_{xy}'' [Eq. (62)] in the limit $m_y \to \infty$. (Note that the factor of $\sqrt{M}$ difference between Φ and f is not important here, since M has a small θ dependence.)

At leading order, the heavy-light meson mass is independent of θ, and has been removed from $\mathcal{L}_{\chi,\mathrm{HL}}$, as usual. However a dependence on θ can enter through the [next-to-leading-order (NLO)] light-quark mass contributions to the Lagrangian,

$$\begin{aligned}\mathcal{L}_{\chi,\mathrm{HL},m} &= 2\lambda_1 B_0 s\mathrm{Tr}(\bar{H}H\mathcal{M}^+) \\ &+ 2\lambda_1' B_0 s\mathrm{Tr}(\bar{H}H)s\mathrm{Tr}(\mathcal{M}^+),\end{aligned} \quad (72)$$

$$\mathcal{M}^+ \equiv \frac{1}{2}(\sigma\mathcal{M}_B^*\sigma + \sigma^\dagger\mathcal{M}_B\sigma^\dagger), \quad (73)$$

where $\mathcal{M}_B$ is the light-quark mass matrix given in Eq. (49), λ_1 and λ_1' are new low-energy constants (LECs), and B_0 (often omitted in definitions of λ_1, λ_1') is the light-quark LEC from Eq. (11). The dependence of the heavy-light meson mass M on the light valence mass is proportional to λ_1, while the sea-quark mass dependence comes from λ_1'.

Plugging $\langle\sigma\rangle$ into Eqs. (73) and (72), and adding the heavy-light mass in the chiral limit, M_0, which has been omitted from Eq. (63), gives

$$\begin{aligned}M_x(\theta) &= M_0 + 2\lambda_1 B_0 m_x \cos\left(\beta - \frac{5}{2}\delta\right) \\ &+ 2\lambda_1' B_0(2m\cos(\alpha - \theta/3) + m_s\cos(2\alpha + \theta/3)).\end{aligned}$$
$$(74)$$

From Eqs. (58), (59) and (60), we then obtain

$$M_x'' = -2B_0\lambda_1\frac{m^2 m_s^2}{(m + 2m_s)^2}\frac{1}{m_x} - 2B_0\lambda_1'\frac{mm_s}{m + 2m_s}. \quad (75)$$

Note that fractional changes in M_x with topology will be quite small (except in the limit when $m_x \ll m$), because M_x is dominated by the M_0 term, which is independent of θ. As in the light-light partially quenched case, our results for

both masses [Eq. (75)] and decay constants [Eq. (71)] vanish when either sea-quark mass goes to zero, consistent with the discussion in Appendix A.

To apply Eq. (75) to lattice data, we need the LECs λ_1, λ_1' and B_0. From the flavor splittings of B or D mesons, we can extract $\lambda_1 \approx 0.2$ (GeV)$^{-1}$ (see, for example, Ref. [26]). A more detailed analysis of the heavy-light lattice data from the Fermilab/MILC Collaboration [27] gives $\lambda_1 = 0.232(2)$ (GeV)$^{-1}$, where the error is statistical only. The same analysis implies $\lambda_1' = 0.042(4)$ (GeV)$^{-1}$; the error is again statistical. The smallness of λ_1' is not surprising, since it is suppressed by large-N_c counting. In addition, the λ_1' term does not blow up as $m_x \to 0$, unlike the λ_1 term. Its effects are thus negligible at the currently available statistical precision. To obtain B_0 we can use, for example, $M_{xy}^2 = B_0(m_x + m_y)$, where M_{xy} is the mass of the light pseudoscalar meson made of $\bar{x}$ and y.

The result for the mass in Eq. (75) does *not* join smoothly onto the corresponding light-light formula (61). As m_y gets heavy, M_{xy}'' falls like $1/\sqrt{m_y}$, where we have used $M_{xy}(0) = \sqrt{B_0(m_x + m_y)}$. However the heavy-light M_x'' is independent of the heavy-quark mass. The difference can be traced to the simple fact that light-light and heavy-light meson masses have different dependences on the masses of their valence quarks. It is still true, though, that in both cases M''/M vanishes in the limit of infinitely heavy quark mass.

VII. STAGGERED CORRECTIONS

It is not difficult to include the leading discretization corrections from taste violations with rooted staggered quarks. Each flavor becomes a staggered field with four tastes, and sea quarks are also replicated n_r times. Rooting is accomplished by taking $n_r \to 1/4$ at the end of the calculation [28,29].

We assume that the exact shift symmetry of staggered quarks [30] does not get spontaneously broken when θ becomes nonzero. At the level of the chiral theory, shift symmetry corresponds to the discrete taste symmetry [29]

$$\Sigma \to \xi_\mu \Sigma \xi_\mu \qquad (\text{no sum on } \mu), \tag{76}$$

where ξ_μ ($\mu = 1, \ldots, 4$) is any of the generators of the taste algebra. This symmetry is enough to guarantee that $\langle\Sigma\rangle$ has a trivial dependence on taste:

$$\langle\Sigma\rangle = \langle\Sigma\rangle_R \otimes \xi_I, \tag{77}$$

where $\langle\Sigma\rangle_R$ is a "reduced" diagonal matrix in flavor and replica space only, and ξ_I is the 4×4 identity matrix in taste space.

The determination of $\langle\Sigma\rangle_R$ in the $2 + 1$-flavor case is then very similar to the calculation of $\langle\Sigma\rangle$ in Sec. V. In analogy with Eq. (50) we parametrize $\langle\Sigma\rangle_R$ by

$$\langle\Sigma\rangle_R = \exp(i\,\mathrm{diag}(\alpha + \delta/n_r + i\epsilon/n_r, \ldots, \alpha + \delta/n_r + i\epsilon/n_r, \ldots,$$
$$-2\alpha + \delta/n_r + i\epsilon/n_r, \ldots, \beta - 3\delta/2 + i\epsilon, -\beta - 3\delta/2 + i\epsilon,$$
$$5i\epsilon/2 - i\gamma, 5i\epsilon/2 + i\gamma)), \tag{78}$$

where $\cdots$ stands for the replication of the preceding entry $n_r - 1$ times, and the explicit factors of $1/n_r$ compared with Eq. (50) are necessary here for the stracelessness of the exponent.

Aside from the replication of the sea-quark flavors, the main difference with the continuum calculation is the presence of the taste-violating contribution to the potential, $a^2\mathcal{V}$ [31,32]. Dependence on θ arises in some terms in $a^2\mathcal{V}$ both explicitly, through the anomalous chiral rotation that removes the $\theta F\tilde{F}$ term (see Appendix A), and implicitly, through the expectation value of Σ. However, because of the simple taste structure of $\langle\Sigma\rangle$, the contributions of $a^2\mathcal{V}$ combine with those of the quark mass term and produce terms proportional to squared taste-singlet meson masses. In direct correspondence with Eqs. (58)–(60), we find

$$\alpha' = \frac{M_{\pi,I}^2 - M_{S,I}^2}{12n_r(M_{\pi,I}^2 + 2M_{S,I}^2)}, \tag{79}$$

$$\beta' = \frac{M_{X,I}^2 - M_{Y,I}^2}{8n_r M_{X,I}^2 M_{Y,I}^2} \cdot \frac{M_{\pi,I}^2 M_{S,I}^2}{M_{\pi,I}^2 + 2M_{S,I}^2} = \hat{\gamma}' \equiv -i\gamma', \tag{80}$$

$$\delta' = \frac{M_{X,I}^2 + M_{Y,I}^2}{20n_r M_{X,I}^2 M_{Y,I}^2} \cdot \frac{M_{\pi,I}^2 M_{S,I}^2}{M_{\pi,I}^2 + 2M_{S,I}^2} = \hat{\epsilon}' \equiv -i\epsilon', \tag{81}$$

where the taste-singlet meson masses are given by

$$M_{\pi,I}^2 = 2B_0 m + a^2\Delta_I, \tag{82}$$

$$M_{S,I}^2 = 2B_0 m_s + a^2\Delta_I, \tag{83}$$

$$M_{X,I}^2 = 2B_0 m_x + a^2\Delta_I, \tag{84}$$

$$M_{Y,I}^2 = 2B_0 m_y + a^2\Delta_I. \tag{85}$$

Here $a^2\Delta_I$ is the splitting of the taste-singlet mesons from the corresponding pseudo-Goldstone (taste-ξ_5) mesons. Note that Eqs. (79)–(81) reduce to Eqs. (58)–(60) after taking the continuum limit and the rooting limit ($n_r \to 1/4$).

We may then find the θ dependence of the valence xy meson mass and decay constant by following at tree level the staggered $\theta = 0$ calculations of Refs. [32,33]. We choose taste ξ_5 for the meson to correspond with the choice made in the simulations. We obtain, after rooting,

$$M''_{xy,5} = -\frac{1}{4M_{xy,5}} \frac{M_{\pi,I}^4 M_{S,I}^4}{(M_{\pi,I}^2 + 2M_{S,I}^2)^2} \left[\frac{M_{X,5}^2}{M_{X,I}^4} + \frac{M_{Y,5}^2}{M_{Y,I}^4} \right], \quad (86)$$

$$f''_{xy,5} = -f_{xy,5} \frac{M_{\pi,I}^4 M_{S,I}^4}{(M_{\pi,I}^2 + 2M_{S,I}^2)^2} \frac{(M_{X,I}^2 - M_{Y,I}^2)^2}{4M_{X,I}^4 M_{Y,I}^4}, \quad (87)$$

where all quantities on the right-hand sides are evaluated at $\theta = 0$. The subscript 5 indicates taste ξ_5. The masses of these pseudo-Goldstone mesons are

$$M_{xy,5}^2 = B_0(m_x + m_y), \quad (88)$$

$$M_{X,5}^2 = 2B_0 m_x, \quad (89)$$

$$M_{Y,5}^2 = 2B_0 m_y. \quad (90)$$

Note that the partially quenched singularities as m_x or m_y go to zero are now cut off at nonzero lattice spacing by the taste-singlet splitting.

Paralleling what occurs for the vacuum angles, the $f''_{xy,5}$ result corresponds precisely to the continuum result, Eq. (62), with the simple replacement of each quark mass by the squared mass of the associated taste-singlet meson. The same simple correspondence between the leading-order continuum and staggered results also occurs for the topological susceptibility [34]. For the meson mass, however, the direct correspondence would occur only for the θ dependence of the taste-singlet mass. The taste-ξ_5 squared mass gets an explicit factor of each valence-quark mass, which appear without the singlet splitting $a^2\Delta_I$, thereby producing the $M_{X,5}^2$ and $M_{Y,5}^2$ terms in Eq. (86). The $M_{xy,5}$ term in the denominator arises simply from the fact that we give $M''_{xy,5}$ and not $(M_{xy,5}^2)''$.

It is straightforward to extend these calculations to heavy-light systems. In the heavy-light chiral Lagrangian, staggered discretization effects appear only at NLO [35]. This is in contrast to the light-light Lagrangian, where the taste-violating potential $a^2\mathcal{V}$ is LO in the usual power counting $m_q \sim a^2$, where m_q is a generic light-quark mass. For Φ, whose θ dependence starts at LO, this means that the result for $\Phi_x(\theta)$ in terms of the angles β and δ [Eq. (70)] remains valid. We just must use the staggered values for β' and δ', Eqs. (80) and (81), to find Φ''_x. We obtain

$$\Phi''_{x,5} = -\Phi_{x,5} \frac{M_{\pi,I}^4 M_{S,I}^4}{(M_{\pi,I}^2 + 2M_{S,I}^2)^2} \frac{1}{4M_{X,I}^4}, \quad (91)$$

where the subscript 5 indicates a taste-ξ_5 meson.

The calculation is a bit more complicated for the heavy-light meson mass because θ dependence first appears at NLO. At this order, there are also a large number of a^2 terms in the heavy-light chiral Lagrangian, which were catalogued in Ref. [35], and appear in the terms $\mathcal{L}_{2,a^2}^{A1}$, $\mathcal{L}_{2,a^2}^{B1}$,

$\mathcal{L}_{2,a^2}^{A2}$, and $\mathcal{L}_{2,a^2}^{B2}$ defined there. Although the majority of these terms do not contribute to θ dependence, there are ten terms that do, both explicitly and implicitly, as in the light-light potential $a^2\mathcal{V}$ discussed above. Unlike what happens in the light-light case, however, the a^2 terms do not combine with the quark-mass terms to form taste-singlet light-light meson masses, because the heavy-light LECs are independent of the light-light ones. We find

$$M''_{x,5} = -\lambda_1 \frac{M_{\pi,I}^4 M_{S,I}^4}{(M_{\pi,I}^2 + 2M_{S,I}^2)^2} \frac{(M_{X,5}^2 + a^2\Delta_{\mathrm{HL}}^{\mathrm{val}})}{M_{X,I}^4} +$$
$$-\lambda_1' \frac{2(M_{\pi,5}^2 + a^2\Delta_{\mathrm{HL}}^{\mathrm{sea}})M_{S,I}^4 + (M_{S,5}^2 + a^2\Delta_{\mathrm{HL}}^{\mathrm{sea}})M_{\pi,I}^4}{(M_{\pi,I}^2 + 2M_{S,I}^2)^2}. \quad (92)$$

Here $\Delta_{\mathrm{HL}}^{\mathrm{val}}$ is a linear combination of the LECs $K_{1,3}^{A1}$, $K_{1,4}^{A1}$, $K_{1,2}^{A2}$, $K_{1,3}^{A2}$, $K_{1,7}^{A2}$, $K_{1,8}^{A2}$, $K_{1,1}^{B2}$, and $K_{1,2}^{B2}$ from Ref. [35] (divided by λ_1), and $\Delta_{\mathrm{HL}}^{\mathrm{sea}}$ is a linear combination of the LECs $K_{2,3}^{A1}$ and $K_{2,4}^{A1}$ (divided by λ_1'). We have not bothered to work out the coefficients in these linear combinations since the relations are unlikely to be useful, but it is straightforward to find them if they are ever needed. One can easily check that Eqs. (91) and (92) reduce to Eqs. (71) and (75), respectively, in the continuum limit.

There are also other, "generic," discretization effects with staggered quarks that have nothing to do with the (partial) violation of chiral symmetry that results in taste splittings. Such generic effects are of order $\alpha_s a^2$ in a tree-level improved staggered action; the two-stage smearing in the HISQ action further suppresses these effects by a numerical factor. Analyses of various physical quantities in HISQ simulations typically give subpercent generic discretization errors for the range of lattice spacings ($a \lesssim 0.09$ fm) considered here. For example, f_K/f_π varies from its continuum limit by about 0.3% over these lattice spacings [13,36].

VIII. COMPARISON TO SIMULATION RESULTS

The calculation of meson masses and decay constants on the HISQ ensembles was described in Ref. [36]. To find the dependence on the topological charge, we use the results of a single-elimination jackknife analysis of these quantities together with the time histories of topological charge shown above. To estimate B'' we rearrange Eq. (10), using $\langle Q^2 \rangle = \chi_t V$, as

$$B|_{Q,V} = \left(B + \frac{1}{2\langle Q^2 \rangle} B'' \right) - \left(\frac{1}{2\langle Q^2 \rangle^2} B'' \right) Q^2. \quad (93)$$

This lets us find B'' by fitting $B|_{Q,V}$ to a constant plus a term quadratic in Q. In fact what is available is a list of single-elimination jackknife averages for B—that is, values for B

obtained by omitting one lattice from the analysis. We can effectively "undo" the jackknife using

$$B_j - \bar{B} = -(N-1)(\bar{B}_j - \bar{B}) \qquad (94)$$

where $\bar{B} = \frac{1}{N}\sum_i B_i$ is the full sample average and $\bar{B}_j = \frac{1}{N-1}\sum_{i\neq j} B_i$ is the jth jackknife sample. In practice we simply fit the jackknife averages, and supply the factor of $-(N-1)$ later. To estimate the error on B'' we assign an error equal to the error in our sample average $\bar{B}$ to each data point, which results in a fit with $\chi^2 \approx N$, and then use the error on the fit parameters found in the standard way. (This actually neglects the part of the variance of B coming from its dependence on Q^2, but in practice this turns out to be a small part of the variance.)

Although the results are barely statistically significant, they are consistent with the χPT predictions. Statistically significant signals are found in the $m_l = m_s/5$ ensembles, since these have much smaller physical volumes than the physical light quark mass ensembles. The effects that we observe, in Eq. (10), all have an overall factor of $1/V$, even if the parenthesized factor $\left(1 - \frac{Q^2}{\chi_T V}\right)$ covers a range of order one. For example, in the $a \approx 0.06$ fm ensembles the $m_l = m_s/5$ lattices have a volume of 180 (fm)^4, while the physical m_l lattices have a volume of 1920 (fm)^4. Also, Eq. (62) shows that the derivatives of the masses and decay constants have a partially quenched divergence when m_x or m_y goes to zero with m_l fixed, and for the $m_l = m_s/5$ ensembles we have used valence-quark masses smaller than m_l, in some cases as small as the physical m_l.

Figure 4 shows $\frac{\partial^2 M}{\partial \theta^2}$ for the $m_s/5$ ensembles for degenerate valence-quark masses, $m_x = m_y$. The black line in the figure is the PQχPT prediction in Eq. (61), which we emphasize is a prediction with no free parameters. Obviously the statistical errors are large, but they are consistent with the prediction, and the divergence at small valence-quark mass is clearly seen. The green and red lines in this figure show the PQχPT prediction including the taste-breaking effects (86). To calculate these effects we need to know the taste splitting Δ_I in Eqs. (82)–(85), which is expected to be proportional to α_s^2. Since α_s changes significantly over this range of lattice spacings, we simply use the values of Δ_I computed directly from $M_I^2 - M_5^2$ at $a = 0.09$ and $a = 0.06$ fm. We do not have a direct measurement of taste breaking on the 0.042 fm ensemble, but we expect the result there to be very close to the continuum result, as is clear from the fact that $a = 0.06$ fm (red) curve is already barely distinguishable from the continuum (black) curve.

Since $\frac{\partial^2 F}{\partial \theta^2}$ vanishes for degenerate valence quarks, we plot this quantity along different lines in Fig. 5. The left panel shows $\frac{\partial^2 F}{\partial \theta^2}$ as a function of one valence-quark mass, m_x, with the other fixed at the strange-quark mass, together
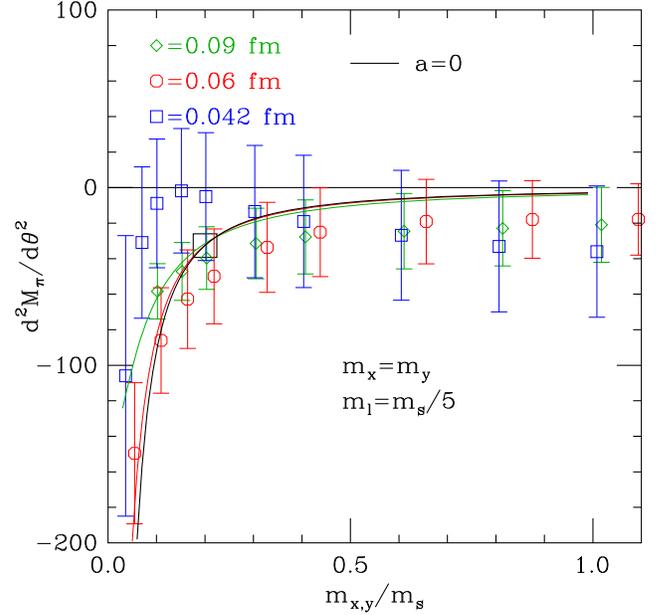

FIG. 4. $\frac{\partial^2 M}{\partial \theta^2}$ on ensembles with $m_l = m_s/5$ as a function of m_x, with $m_y = m = m_x$. The black line is the PQχPT prediction at $a = 0$, or without taste-breaking effects (no free parameters). The green and red lines show the PQχPT prediction including taste breaking [Eq. (86)] for the 0.09 and 0.06 fm ensembles. The black square marks the unitary point, with the valence-quark mass equal to the sea-quark mass.

with the χPT prediction. The right panel shows $\frac{\partial^2 F}{\partial \theta^2}$ along lines where m_y is held fixed at the lightest valence-quark mass available in each ensemble. The vanishing of $\frac{\partial^2 F}{\partial \theta^2}$ when the valence quarks are degenerate is particularly striking in this plot. In the left panel, with $m_y = m_s$, the black line shows the continuum result, without taste breaking, and the green and red lines show the results for 0.09 and 0.06 fm including the taste breaking. In the right panel, since each of these three ensembles has a different lightest valence mass, the colored dotted lines show the prediction without including taste breaking for each of the three ensembles. Because one of the valence-quark masses, m_y, is held fixed at its lightest value in the right-hand panel, the effects of taste violation are large for all values of m_x. As a fraction of the continuum value, the staggered effects in this graph do not decrease as much as might be expected when a changes from 0.09 to 0.06 fm because the relative size of $a^2\Delta_I$ and $2B_0 m_y$ in $M_{Y,I}$ in Eqs. (85) and (87) is what matters, and m_y has changed from $m_s/10$ to $m_s/20$.

Our statistical errors on the heavy-light masses and decay constants are much larger than on the light quark quantities, so we are unable to test our data against the chiral perturbation theory for the heavy-light quantities.

Knowing the dependence of masses and decay constants on the average Q^2, we can correct our simulation results to account for the difference of the average in our simulation,
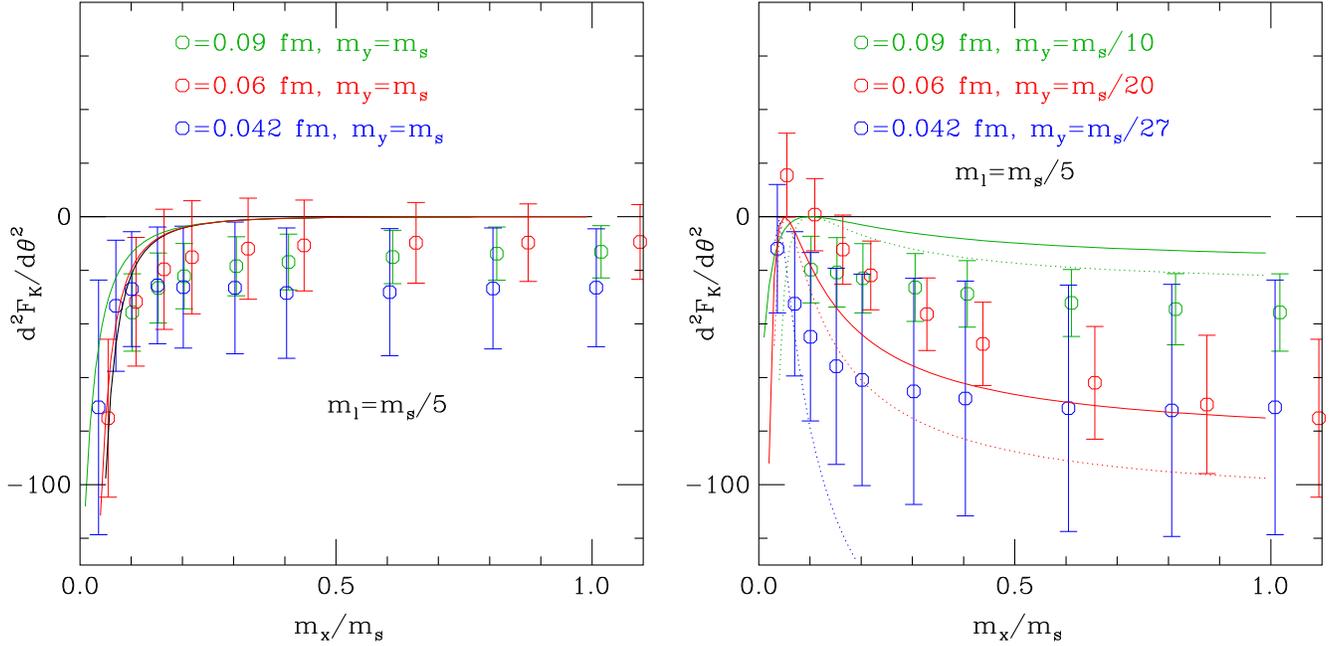
FIG. 5. $\frac{\partial^2 F}{\partial \theta^2}$ on ensembles with $m_l = m_s/5$. The left panel shows $\frac{\partial^2 F}{\partial \theta^2}$ as a function of one valence quark mass, m_x along the line $m_y = m_s$. The black line is the PQχPT prediction (no free parameters), which vanishes for degenerate quarks. The red ($a = 0.06$ fm) and green ($a = 0.09$ fm) lines include the effects of taste breaking from Eq. (87). The right panel shows the quantity with m_y fixed to the smallest available value. The dotted lines are PQχPT predictions ignoring taste breaking, or at $a = 0$; there are three separate lines because the smallest valence-quark mass is different in each ensemble: $0.1 m_s$, $0.05 m_s$ and $0.037 m_s$ for the 0.09, 0.06 and 0.042 fm ensembles respectively. Again, the solid red and green lines include the effects of taste breaking from Eq. (87).

$\langle Q^2 \rangle_{\rm sample}$, and the correct $\langle Q^2 \rangle$. To estimate this correct $\langle Q^2 \rangle$ we use the lowest-order staggered χPT result [34]

$$\chi_T = \frac{f_\pi^2}{4(2/M_{\pi,I}^2 + 1/M_{S,I}^2)} \qquad (95)$$

where the taste-singlet masses $M_{\pi,I}^2$ and $M_{S,I}^2$ are given in Eqs. (82) and (83). The χPT results are shown in Fig. 3. In computing the χPT form for the ensembles with $a > 0.042$ fm, we use measured values for the taste-singlet pion and $s\bar{s}$ pseudoscalar masses on each ensemble. For the 0.042 fm physical quark-mass ensemble we estimate taste-breaking effects by scaling the taste splittings from the 0.06 fm physical quark-mass ensemble, and for the 0.042 and 0.03 fm, $m_s/5$ ensembles, where we expect the taste-breaking effect to be negligible (and certainly not measurable with our statistics), we use the Goldstone pseudoscalar masses. For large a, the deviation from the lowest-order χPT result is due to lattice artifacts, probably mostly higher-order taste-breaking effects, but for $a = 0.042$ and 0.03 fm we expect the χPT results to be pretty good.

For examples of the size of this effect in our simulations, we look at the two ensembles with $a \approx 0.042$ fm. For example, to adjust the decay constants, we rearrange Eq. (10) as

$$f_{\rm corrected} = f_{\rm sample} - \frac{1}{2\chi_T V} f'' \left(1 - \frac{\langle Q^2 \rangle_{\rm sample}}{\chi_T V} \right). \qquad (96)$$

Table I shows the size of the topology adjustment for selected quantities, together with the central value and statistical errors. The sign of the adjustment differs between the two ensembles because, as can be seen in Fig. 3, the difference between the sample average Q^2 and the chiral perturbation theory prediction is different in the two cases. The effects are larger in the ensemble with $m_l/m_s = 0.2$, since these lattices have much smaller volume and partial quenching divergences. It is amusing to note that for the physical quark mass ensemble the topology adjustment for f_K/f_π is a factor of ~6 smaller than the "conventional" finite-size effects from pions propagating around the periodic lattice, estimated in NLO staggered χPT, of 0.0009. (Conventional finite-size effects on the heavy quark quantities are quite small.)

We note that this strategy is in the same spirit as our treatment of conventional finite-size effects. We use χPT to estimate the effects and correct our results, and estimates of the effects of higher-order χPT and/or uncertainties in the χPT parameters should be included in the systematic error budget.

TABLE I.   Examples of topological adjustments for pseudoscalar masses and decay constants in the 0.042 fm HISQ ensembles. Each field contains the unshifted value for the quantity, its statistical error in parentheses, and the topology adjustment in square brackets. These quantities are evaluated at valence strange and charm quark masses equal to the sea quark masses, which have small tuning errors. The light valence mass was $1/27.4$ times the sea strange quark mass. For the heavy-light masses we used $B_0 = 3.4 \text{GeV}$, $\lambda_1 = 0.232(\text{GeV})^{-1}$ and $\lambda_1' = 0.042(\text{GeV})^{-1}$ [27]. Note that since the Q^2 in the second line is the Q^2 averaged over our sample, there is no statistical error associated with it.

	$m_l = m_s/5$	$m_l = \text{physical}$
$Q^2_{\text{sample}}/Q^2_{\chi\text{PT}}$	1.30	0.65
f_K/f_π	1.20508 (0.00250) [−0.01271]	1.19680 (0.00114) [0.00015]
aM_π	0.031147 (0.000172) [−0.000707]	0.028964 (0.000020) [0.000008]
af_D	0.048858 (0.000261) [−0.000552]	0.045389 (0.000245) [0.000006]
aM_D	0.409786 (0.000391) [−0.000044]	0.400678 (0.000258) [0.000001]
af_{D_s}	0.054828 (0.000068) [−0.000001]	0.053582 (0.000025) [0.000000]
aM_{D_s}	0.430966 (0.000116) [−0.000004]	0.422041 (0.000037) [0.000000]

IX. CONCLUSIONS

Our key χPT results are given in Eqs. (61) and (62) for partially quenched light-light meson masses and decay constants, and in Eqs. (75) and (71) for the corresponding quantities for heavy-light mesons. In the light-light case, these results reproduce those of Aoki and Fukaya [7], but are now computed within a nonperturbatively justified partially quenched chiral theory. Results with the leading staggered discretization effects included are given in Eqs. (86), (87), (91) and (92). The results can be compared with simulation data for the quantities as a function of topological charge Q using Eq. (10). We have done this in Figs. 4 and 5; there are large statistical errors but the qualitative agreement is good. Discretization effects are generally quite small at the fine lattice spacings where these results are likely to be used. For very light valence-quark masses in comparison to sea-quark masses, however, the staggered taste splitting cuts off the partially quenched singularity and can thus be numerically important. One can also use the results in conjunction with Eq. (96) to adjust quantities for incorrect sampling of Q, as discussed at the end of Sec. VIII. The corrections in the cases we have looked at are very small (much less than the statistical errors) for our physical-mass ensemble at $a \approx 0.042$. For the ensemble with $m_l = m_s/5$ at the same lattice spacing, however, the corrections are often statistically significant, with the largest ($\sim$5 times the statistical sigma) occurring for f_K/f_π, which has very small statistical errors. Both the smaller volume of this ensemble and the partial quenching divergences play a role here. In view of the fact that the corrections are calculated only to leading order in χPT, one should take a relatively large fraction of the correction as the associated systematic error. Nevertheless, our analysis suggests that it will be possible to satisfactorily control the systematic errors due to nonequilibrated topological charge distributions, even at rather small lattice spacings.

ACKNOWLEDGMENTS

APPENDIX A: VANISHING SEA-QUARK MASS

The usual expectation is that all θ dependence should disappear when one (or more) sea-quark masses vanish. Indeed that is the reason that a zero value for the up-quark mass would solve the strong CP problem [38]. However, we will see in this appendix that the absence of θ dependence is in general true only for spectral quantities in QCD, i.e., quantities that are entirely determined by the QCD Lagrangian. While the second derivatives with respect to θ should vanish for all meson masses in the limit of a vanishing sea-quark mass, this is not necessarily true for decay constants, which depend on external (axial) currents. In particular, the results for decay constants in unitary theories, Eqs. (28) and (39), do not vanish when one sea-quark mass goes to zero. On the other hand, as long as the valence masses remain nonzero, derivatives of the partially quenched decay constant, Eq. (62), do vanish in this limit.

To understand what is going on, it is helpful to look in detail at the chiral transformations that have (implicitly) been used to put the Lagrangian in various convenient forms. We work here in the partially quenched context so that the results will apply to all the calculations described above. Under a chiral transformation,

$$\Sigma \to L\Sigma R^{-1}, \qquad \Sigma^{-1} \to R\Sigma^{-1}L^{-1}. \qquad \text{(A1)}$$

The axial current transforms as

$$\mathcal{A}_\mu^{ij} = i\frac{f^2}{4}(\partial_\mu\Sigma\Sigma^{-1} + \Sigma^{-1}\partial_\mu\Sigma)^{ij}$$

$$\to i\frac{f^2}{4}(L\partial_\mu\Sigma\Sigma^{-1}L^{-1} + R\Sigma^{-1}\partial_\mu\Sigma R^{-1})^{ij}. \qquad \text{(A2)}$$

In the partially quenched Lagrangian, the chiral transformation is equivalent to leaving Σ unchanged and transforming the mass matrix as

$$\mathcal{M} \to L^{-1}\mathcal{M}R, \qquad \mathcal{M}^* \to R^{-1}\mathcal{M}^*L. \qquad \text{(A3)}$$

Note that this is the inverse of the fake (spurion) transformation on $\mathcal{M}$ that would leave the Lagrangian invariant.

The first chiral transformation we consider (call it "A") is the anomalous, flavor-singlet one that removes the $\theta F\tilde{F}$ term from the QCD Lagrangian, and puts a uniform phase in the mass matrix, as in Eq. (11):

$$R_A = L_A^{-1} = \exp\left(\frac{i\theta}{6}I\right),$$

$$\mathcal{M} \to \mathcal{M}_A \equiv \exp\left(\frac{i\theta}{3}\right)\mathcal{M}, \qquad \text{(A4)}$$

where I is the identity matrix given in Eq. (46), and we have specialized to $N = 3$ and $N_v = 2$. All axial currents are invariant under this flavor-singlet transformation.

The second chiral transformation ("B"), is the one that removes the θ dependence from the valence- and ghost-quark masses. This is the nonanomalous transformation

$$R_B = L_B^{-1} = \exp\left(\frac{-i\theta}{6}t\right), \qquad \text{(A5)}$$

with t given in Eq. (47). Transformation B produces the mass matrix $\mathcal{M}_B$ given in Eq. (49) and used in the calculations of Sec. V. It leaves valence-valence currents (or sea-sea currents) unchanged, so f''_{xy} in Eq. (62) is correct for the axial current defined as usual, with no θ dependence.

It is straightforward to check that we get the same answer for f''_{xy} using the mass matrix $\mathcal{M}_A$ [Eq. (A4)] instead of $\mathcal{M}_B$. As a further check, we may calculate the θ dependence of the (strange) sea-valence decay constant, f''_{sx}. In this case it is crucial to include the nontrivial θ dependence of the axial current induced by transformation B via Eq. (A2). Once this is done, the results in the A and B cases agree, and agree with the result for f''_{xy} [Eq. (62)] when $m_x = m_s$. The meson mass is of course insensitive to the currents, so there are no subtleties in the calculation, and the results in cases A and B are again identical.

We can now turn to the question of θ dependence when a sea-quark mass vanishes. For definiteness in our $N = 3$ example, let us take $m_s \to 0$. The trick here is to make a third chiral transformation ("C") that is nonanomalous and

puts all θ dependence into the m_s term in the quark mass matrix:

$$R_C = L_C^{-1} = \exp\left(\frac{-i\theta}{6}\lambda\right), \quad [\lambda \equiv \mathrm{diag}(1,1,-2,0,0,0,0)],$$

(A6)

$$\mathcal{M}_B \to \mathcal{M}_C = \mathrm{diag}(m, m, e^{i\theta}m_s, m_x, m_y, m_x, m_y).$$ (A7)

Now all θ dependence disappears for spectral quantities in the limit $m_s \to 0$. The valence-valence meson mass provides an example: M''_{xy} in Eq. (61) vanishes in this limit, for any (fixed, nonzero) values of the other masses. It is important not to take a valence mass to zero before $m_s \to 0$; the limits are not interchangeable because of the PQ singularities. We can, however, put $m_x = m_s$, $m_y = m$ or $m_x = m_y = m$ to get the full theory M''_K or M''_π, respectively, Eqs. (37) and (36), which again vanish in the $m_s \to 0$ limit.

Transformation C does not affect the valence-valence axial current, so f''_{xy} [Eq. (62)] also vanishes as $m_s \to 0$. However, this is not true of the full theory f''_K [Eq. (39)]. In this case, the axial current $\mathcal{A}^{13}_\mu$ is changed by the transformation. Indeed, all θ dependence in the limit $m_s \to 0$ comes from the current, and we find $f_K(\theta) = f\cos(\theta/2)$. This gives the nonvanishing result $f''_K = -f_K(0)/4$ in this limit, in agreement with Eq. (39).

We can similarly check the $m \to 0$ limit. In this case, we should put the θ dependence equally into the up- and down-quark entries of the mass matrix, so as not to spoil isospin invariance, which was assumed in the calculations of Sec. V. We then find $f_K(\theta) = f\cos(\theta/4)$ in this limit, giving $f''_K = -f_K(0)/16$, in agreement with the limit of Eq. (39). Further, M''_K and M''_π should vanish in this limit, in agreement with Eqs. (37) and (36). (For M''_π, we need to use the fact that M_π itself vanishes in the limit.)

Finally, a note of warning. The various limits of vanishing quark mass are subtle, and it is easy to go astray. This is already clear in the full theory from the fact that the limits $m \to 0$ and $m_s \to 0$ do not commute for f''_K [Eq. (39)]. Another interesting example is the limit $m_x \to 0$ for a valence mass, with sea masses and other valence masses held fixed. We expect a partially quenched singularity in this case, and Eq. (61) shows this for the spectral quantity M''_{xy}. However, we can also make a plausible-sounding argument that M''_{xy} should vanish in this limit! Starting from case B, suppose we make a nonanomalous transformation ("D")

$$R_D = L_D^{-1} = \exp\left(\frac{-i\theta}{6}\tilde{\lambda}\right), \quad [\tilde{\lambda} \equiv \mathrm{diag}(1,1,1,-3,0,0,0)],$$

(A8)

$$\mathcal{M}_B \to \mathcal{M}_D = \mathrm{diag}(m, m, m_s, e^{i\theta}m_x, m_y, m_x, m_y).$$ (A9)

Now all θ dependence, for spectral quantities, is in the valence-quark x term in $\mathcal{M}_D$, so it would appear that M''_{xy} vanishes in the $m_x \to 0$ limit.

The problem with this argument is that the partially quenched singularity is so strong as $m_x \to 0$ that it overwhelms the reduction in spectral quantities coming directly from the $e^{i\theta}m_x$ term in the quark mass matrix. Keeping $m_x \neq 0$ and repeating the calculational steps in Sec. V, we in fact reproduce Eq. (61) for M''_{xy}. This is not unexpected because we have simply made a nonanomalous chiral transformation, which should not affect physical quantities. Indeed, if we are careful about the phase introduced in the axial current by transformation D, we also reproduce f''_{xy} in Eq. (62). Further, despite the apparent breaking of quark-ghost symmetry by $\mathcal{M}_D$, it is still true that $\langle\Sigma\rangle_{xx} = \langle\Sigma\rangle_{\tilde{x}\tilde{x}}$, just as in the cases where we preserve quark-ghost symmetry explicitly in the mass matrix.

One may still wonder whether we can accept that there is a discontinuity at $m_x = 0$ and simply set $m_x = 0$ from the start. This is not allowed, however, because the ghost integral needs a nonzero mass term for convergence.[4]

APPENDIX B: CONVERGENCE OF THE NEUTRAL GHOST-ANTIGHOST INTEGRALS

The bosonic path integrals over the field $\tilde{\phi}$ for ghost-antighost mesons [Eq. (44)] and the field ϵ for the "ghost-like" neutral meson [Eqs. (43) and (45)] must be convergent in order for PQχPT to be nonperturbatively well defined. Further, in order for the perturbative vacuum defined by $\tilde{\phi} = 0$, $\epsilon = 0$ to be the correct one, it seems that $\mathrm{Re}(V)$ should have a minimum at this point for real ϵ and Hermitian $\tilde{\phi}$. These conditions apparently require that the real part of the Lagrangian, expanded to quadratic order in the fields, should be positive definite. Indeed, the factors of i in Eqs. (43) and (44) are inserted to ensure that $\tilde{\phi}$ and ϵ have positive kinetic energy terms for $p^2 \neq 0$. The same factor of i also guarantees that the real part of the mass term of $\tilde{\phi}$ is positive definite as long as no valence masses vanish.

The ghost-like field ϵ, however, presents problems. Because it is a linear combination of quark-antiquark and ghost-antighost fields, there is a competition in its mass term between ghost masses, which give positive terms, and quark masses, which give negative terms. Each ghost term always wins over the corresponding valence term, because ϵ has more support in the ghost sector than in the valence sector. In contrast, the competition with the sea masses can go either way. For low enough valence and ghost masses compared to sea masses, the real part of the ϵ mass term will become negative, putting into doubt the convergence of the ϵ path integral, or at least the validity of the perturbative vacuum at $\epsilon = 0$. This leads to

---

[4] We thank Maarten Golterman for this point.

the Sharpe-Shoresh lower bound [10] on the valence masses. Because ϵ can mix with the neutral component of $\tilde{\phi}$, corresponding for example to the parameter γ in Eq. (50), Ref. [10] requires that the full neutral ghost and ghost-like mass matrix be positive definite, resulting in the bound

$$N_v \overline{\chi_v^{-1}}(N_v \overline{\chi_v} + N\bar{\chi}) < (N + N_v)^2, \quad \text{(B1)}$$

where $\bar{\chi}$ and $\overline{\chi_v}$ are the average sea-quark mass and average valence-quark mass, respectively, and $\overline{\chi_v^{-1}}$ is the average inverse valence-quark mass. Sharpe and Shoresh suspect that this is some kind of artifact of the chiral theory, since the underlying partially quenched QCD has no apparent problem at or below this bound. Nor is there any evidence from the calculation of standard perturbative quantities within partially quenched χPT, or their comparison with simulations, that things go wrong when the bound is violated. Nevertheless, since much of the simulation data that we analyze violates this bound, the apparent lack of convergence of the chiral theory is disconcerting. We certainly cannot claim our chiral results to be nonperturbatively correct in the region where the bound is violated unless we can show that the bound itself is not actually an obstacle to using the theory around the standard vacuum.

We work primarily in the case $\theta = 0$; nonzero but small θ does not present any significant additional problems. We also start by considering a simpler theory than that of Sec. V, with $N = 2$, $N_v = 1$ and degenerate sea-quark masses. After showing in this simple model that a violation of the Sharpe-Shoresh bound does not lead to any problem with convergence or with the perturbative vacuum, we will be able to use a shortcut to arrive at a similar conclusion for the case of interest, $N = 3$ and $N_v = 2$ with nondegenerate valence and sea masses. We can easily generalize from there to arbitrary N with arbitrary sea-quark masses. We will not attempt to prove the result for arbitrary $N_v > 2$, but will argue that it is probably true in that case too.

In the neutral sector of the $N = 2$, $N_v = 1$ model, there are three mesons, π, δ, and ϵ, with the meson field Φ of Eq. (42) given by

$$\Phi = \frac{1}{\sqrt{6}}\text{diag}(\sqrt{3}\pi + \delta + i\epsilon, -\sqrt{3}\pi + \delta + i\epsilon, -2\delta + i\epsilon, 3i\epsilon),$$
$$\text{(B2)}$$

where entries are ordered (sea, valence, ghost). The quark mass matrix is $\mathcal{M} = \text{diag}(m, m, m_x, m_x)$. Expanding the Lagrangian in momentum space to quadratic order in these fields, we find

$$\mathcal{L}_{\text{quad}} = \frac{1}{2}(p^2 + M_\pi^2)\pi^2 + \frac{1}{2}(p^2 + M_\delta^2)\delta^2 + \frac{1}{2}(p^2 + M_\epsilon^2)\epsilon^2$$
$$+ i\alpha\delta\epsilon, \quad \text{(B3)}$$

where

$$M_\pi^2 = B_0(2m), \qquad M_\delta^2 = B_0\left(\frac{4}{3}m_x + \frac{2}{3}m\right),$$

$$M_\epsilon^2 = B_0\left(\frac{8}{3}m_x - \frac{2}{3}m\right), \qquad \alpha = B_0\left(\frac{2}{3}m - \frac{2}{3}m_x\right). \quad \text{(B4)}$$

We see that M_ϵ^2 is only positive for $m_x > m/4$, which is precisely the Sharpe-Shoresh lower bound for this case. However, there is also an imaginary (hence not Hermitian) mixing term between ϵ and δ. (There is no mixing with π in this model because of the exact sea-quark isospin symmetry.)

The ϵ-δ mixing term has an important effect: if we treat it as an iterated two-point interaction, it generates the expected double poles for neutral particles in a partially quenched theory. In other words, it plays the role that the anomalous Φ_0 mass term plays in the case where the limit of infinite Φ_0 mass is postponed until after the computation of the neutral propagators. Note that the poles (whether single or double) of a partially quenched neutral propagator occur at the squared masses either of physical sea-sea neutral mesons, or of unmixed valence-valence meson masses, which here would be proportional to m_x. All these squared masses are positive (for nonzero quark masses), which suggests that the apparent problem of $M_\epsilon^2 < 0$ can be avoided if we treat the mixing term $i\alpha\delta\epsilon$ on the same footing as the other mass terms. We can accomplish this by performing the path integral over δ *before* the ϵ integral. The δ integral is convergent since $M_\delta^2 > 0$.

By completing the square, we may integrate over δ in the partition function

$$Z = \int D\pi D\epsilon D\delta \, e^{-\int d^4x \mathcal{L}_{\text{quad}}} \quad \text{(B5)}$$

and obtain

$$Z = \int D\pi D\epsilon \, e^{-\frac{1}{2}\int d^4x((p^2 + M_\pi^2)\pi^2 + F(p, M_\epsilon, M_\delta, \alpha)\epsilon^2)}, \quad \text{(B6)}$$

where

$$F(p, m_\epsilon, m_\delta, \alpha) = p^2 + M_\epsilon^2 + \frac{\alpha^2}{p^2 + M_\delta^2}$$

$$= \frac{(p^2 + M_X^2)^2}{p^2 + M_\delta^2}, \quad \text{(B7)}$$

$$M_X^2 \equiv B_0(2m_x). \quad \text{(B8)}$$

Here M_X is the mass of the valence $\bar{x}x$ meson. Note that F is positive definite, so the ϵ integral converges, independent of whether m_x is above or below the Sharpe-Shoresh lower bound.

The ϵ propagator, $G_{\epsilon\epsilon} = F^{-1}$ has a characteristic double pole at M_X^2:

$$G_{\epsilon\epsilon} = \frac{p^2 + M_\delta^2}{(p^2 + M_X^2)^2} = \frac{1}{3}\left[\frac{2}{p^2 + M_X^2} + \frac{p^2 + M_\pi^2}{(p^2 + M_X^2)^2}\right], \quad (B9)$$

where the second form will simplify the comparison to more complicated cases. We can also obtain this result for $G_{\epsilon\epsilon}$ by iterating the $i\alpha\delta\epsilon$ term in Eq. (B3) or by returning to the theory with Φ_0 included, iterating the Φ_0 mass term, and then taking the mass to infinity. Repeating the derivation of Eq. (B6) with sources for the δ and ϵ fields allows us to find the δ propagator $G_{\delta\delta}$ and the mixed propagator $G_{\epsilon\delta}$, which of course also agree with the results found using the other methods.

We have thus obtained a convergent ϵ path integral by integrating over δ first. The path integral, as given by Eq. (B5), is however only conditionally convergent for $m_x < m/4$. The integral of the absolute value is not convergent, since the imaginary mixing term is lost and the ϵ integral is then divergent. Alternatively, it is clear that the result is ill defined for $m_x < m/4$ if the ϵ integral in Eq. (B5) is performed first.

The lack of absolute convergence of Eq. (B5) suggests that the partially quenched chiral theory for $m_x < m/4$ is a delicate object for which the usual manipulations of perturbation theory are suspect and must be checked carefully. This suggestion is however incorrect! The reason is that the exponent in Eq. (B5) is only the quadratic action, and not the full action. With the full action, the ϵ integral is convergent even if it is performed first. That is because the ghost-ghost component of Φ in Eq. (B2) gives rise to a term in the potential that goes like $+m_x \cosh(\sqrt{6}\epsilon/f)$, which dominates for large $|\epsilon/f|$ over the negative term $-(2m + m_x)\cosh(\sqrt{2/3}\epsilon/f)$ for all nonzero values of m_x and m. Thus if we make the quadratic approximation of Eqs. (B3) and (B5), we should consider the ϵ integral as cut off at large but finite $|\epsilon|$. (We might, for example, add a term $\lambda\epsilon^4$ to the Lagrangian, where λ is small and positive.) In that case the full path integral is absolutely convergent, and by Fubini's theorem [39] the order of integration does not affect the answer. It is true, though, that the simplest way to evaluate the partition function is by first integrating over δ. Indeed, we do not know how to show that the final result even has a limit when $\lambda \to 0$, except by the use of Fubini's theorem.

In the presence of nonzero, but small, θ, these results change only slightly. The saddle point is now at $\pi = 0$, $\delta = \delta_0$, $\epsilon = i\delta_0$, where δ_0 is of order θ. For $m_\epsilon^2 > 0$ the steepest descent directions from the saddle point are the real directions for all three fields, and we keep those directions even when $m_\epsilon^2 < 0$. Expanding around the saddle, there are of course no linear terms in the fields, and the quadratic terms are changed slightly by cosines of the angles θ and

δ_0. The end result is that Eqs. (B6) and (B7) are still valid, but with redefined values:

$$M_X^2 = 2B_0\sqrt{m_x^2 - m^2\sin^2(\theta/2)},$$

$$M_\delta^2 = B_0\left(\frac{2}{3}m\cos(\theta/2) + \frac{4}{3}\sqrt{m_x^2 - m^2\sin^2(\theta/2)}\right). \quad (B10)$$

We may restrict ourselves to infinitesimal θ values in order to find the derivatives with respect to θ at $\theta = 0$. The quadratic action of ϵ therefore remains positive definite for any nonzero quark masses. On the other hand, for θ small but finite, a new singularity would develop for very small but nonzero $m_x < m\sin(\theta/2)$. We do not concern ourselves further with this interesting, but to us irrelevant, singularity.

We now turn to the case discussed in Sec. V: $N = 3$ (masses $m_u = m_d = m$, m_s) and $N_v = 2$ (masses m_x, m_y). In this case, there are two neutral ghost-type fields, corresponding to ϵ and γ in Eq. (50), and three neutral quark-type fields that they may mix with, corresponding to α, β, and δ. Because the algebra involved in doing the convergent integrals over α, β, and δ is rather messy, we resort to a short cut. As noted in the discussion of the toy model above, Gaussian integration (with possible linear terms) is equivalent to perturbation theory. Hence we can simply find the ϵ-ϵ, γ-γ and ϵ-γ propagators using perturbation theory, and deduce the positivity of the ϵ-γ quadratic action from the propagator matrix.

The easiest way to obtain the needed propagators is to restore the anomalous Φ_0 field (mass m_0) and use the diagonal basis for neutral fields, $\Phi = \text{diag}(U, D, S, X, Y, \tilde{X}, \tilde{Y})$. Quark-line connected and disconnected propagators among the diagonal-basis fields are the standard ones of PQχPT, and the limit $m_0 \to \infty$ may be taken. We then just need to write ϵ and γ as linear combinations of the diagonal-basis fields, and take the corresponding linear combinations of the propagators. With

$$\epsilon = -i\sqrt{\frac{2}{15}}\left(U + D + S + X + Y + \frac{5}{2}\tilde{X} + \frac{5}{2}\tilde{Y}\right), \quad (B11)$$

$$\gamma = -i\sqrt{\frac{1}{2}}(-\tilde{X} + \tilde{Y}), \quad (B12)$$

$$M_X^2 = 2B_0 m_x, \qquad M_Y^2 = 2B_0 m_y, \qquad M_\pi^2 = 2B_0 m,$$

$$M_S^2 = 2B_0 m_s, \qquad M_\eta^2 = B_0(2m/3 + 4m_s/3), \quad (B13)$$

we find

$$G_{\epsilon\epsilon} = \frac{1}{10}\left[\frac{3}{p^2 + M_X^2} + \frac{3}{p^2 + M_Y^2} + \frac{(p^2 + M_\pi^2)(p^2 + M_S^2)}{p^2 + M_\eta^2}\left(\frac{1}{p^2 + M_X^2} + \frac{1}{p^2 + M_Y^2}\right)^2\right],$$

$$G_{\gamma\gamma} = \frac{1}{6}\left[\frac{3}{p^2 + M_X^2} + \frac{3}{p^2 + M_Y^2} + \frac{(p^2 + M_\pi^2)(p^2 + M_S^2)}{p^2 + M_\eta^2}\left(\frac{1}{p^2 + M_X^2} - \frac{1}{p^2 + M_Y^2}\right)^2\right],$$

$$G_{\epsilon\gamma} = \frac{1}{\sqrt{60}}\left[\frac{3}{p^2 + M_Y^2} - \frac{3}{p^2 + M_X^2} + \frac{(p^2 + M_\pi^2)(p^2 + M_S^2)}{p^2 + M_\eta^2}\right.$$
$$\left. \times\left(\frac{1}{(p^2 + M_Y^2)^2} - \frac{1}{(p^2 + M_X^2)^2}\right)\right]. \tag{B14}$$

We write the ϵ-γ propagator matrix as

$$G = \begin{pmatrix} G_{\epsilon\epsilon} & G_{\epsilon\gamma} \\ G_{\epsilon\gamma} & G_{\gamma\gamma} \end{pmatrix}. \tag{B15}$$

Then the ϵ-γ quadratic action matrix, after integrating out all the neutral quark-antiquark fields that have quadratic interactions with ϵ,[5] must be G^{-1}. Because G and G^{-1} are real symmetric matrices, they are diagonalizable. This means that if one of them is positive definite, the other must be too. So we need only prove that G is positive definite. We can do that by showing that its eigenvalues are both positive, which simply requires $\mathrm{tr}(G) > 0$ and $\det(G) > 0$. From Eq. (B14), $\mathrm{tr}(G)$ is clearly positive. After some algebra, we find

$$\det(G) = \frac{1}{5(p^2 + M_X^2)(p^2 + M_Y^2)}$$
$$\times\left[3 + \frac{(p^2 + M_\pi^2)(p^2 + M_S^2)}{p^2 + M_\eta^2}\right.$$
$$\left. \times\left(\frac{1}{p^2 + M_X^2} + \frac{1}{p^2 + M_Y^2}\right)\right], \tag{B16}$$

which is also positive.

Thus, the quadratic action of the neutral ghost-antighost fields in the $N = 3$, $N_v = 2$ case is positive definite after integration of the neutral quark-antiquark fields at quadratic order. As in the $N = 2$, $N_v = 1$ case, it is also true that here the quadratic ghost-antighost integrals may be considered to be cut off at large field values by higher terms in the full action. The cutoff terms, which come from contributions to the potential from the ghost-ghost block, grow like $m_x \cosh((\sqrt{10/3}\epsilon - \sqrt{2}\gamma)/f) + m_y \cosh((\sqrt{10/3}\epsilon + \sqrt{2}\gamma)/f)$ and dominate the negative quark-quark contributions in all

---

[5]No such fields have quadratic interactions with γ.

real directions of the ϵ-γ plane. So, once again, the path integral around the trivial vacuum to quadratic order is absolutely convergent. The order of integration has no effect, except in the practical sense that integrating the neutral quark-antiquark fields first is much easier.

For small but nonzero θ the analysis follows the procedure described above for $N = 2$, $N_v = 1$. The saddle point occurs at imaginary values of the neutral ghost-antighost fields. We expand around the saddle point in the real directions of these fields. Compared to the $\theta = 0$ case, the quadratic terms are changed slightly by cosines of linear combinations of θ and the saddle-point angles, which can make small differences in the meson masses, as in Eq. (B10). The theory therefore remains positive except possibly for very small values of the valence masses of order $\bar{m}\theta$ (where $\bar{m}$ is the average sea-quark mass). Such possibilities do not pose any difficulties for our analysis in the body of this paper, since we only need to consider infinitesimal θ values to find the derivatives at $\theta = 0$.

It is quite easy to generalize our results to an arbitrary number $N \geq 2$ of sea-quark flavors, and arbitrary sea-quark masses. The only changes in Eqs. (B14) and (B16) will be (1) an adjustment of the relative normalization of the single pole and double pole terms and (2) a replacement of the factor $(p^2 + M_\pi^2)(p^2 + M_S^2)/p^2 + M_\eta^2)$ in each equation with the corresponding sea-meson factor that multiplies the disconnected neutral propagator in the given theory. Since this factor "goes along for the ride" in all the manipulations that led to Eqs. (B14) and (B16), the quadratic action in the ghost-ghost sector will remain positive definite.

Since $N_v = 2$ is the most useful case for analyzing simulations, we have not tried very hard to generalize to the more complicated cases with $N_v > 2$. There are however some indications that the quadratic action of the ghost-antighost sector remains positive definite after integration of the neutral quark-antiquark fields. First of all, it is clear that $G_{\epsilon\epsilon} > 0$, since the sum of two valence-mass poles in Eq. (B14) will just be replaced by the sum of N_v valence mass poles (and again the relative normalization of single and double pole terms may change). We can see this change explicitly by comparing with the $N_v = 1$ case, Eq. (B9).

Second, when the valence masses are degenerate, ϵ does not interact in the quadratic action with any other ghost-type fields (which themselves have positive masses and do not interact with the neutral quark-type fields). Therefore the positivity of $G_{\epsilon\epsilon}$ is all we need for the positivity of G. Any nonpositivity in the nondegenerate case would have to come from large effects of interactions with the other ghost-type fields. Such effects can occur before integration over the quark-type fields, as in the Sharpe-Shoresh bound. In that case M_ϵ^2 can be positive but small, and then interactions can generate a negative eigenvalue in the complete ghost-type mass matrix. However, this seems less likely to occur after integration over the quark-type fields. Based on our $N_v = 2$ example in Eq. (B14), it looks difficult to choose

masses such that $G_{\epsilon\epsilon}$ is small, but interaction terms (such as $G_{\epsilon\gamma}$) get large enough to change the sign of an eigenvalue. In any case it is clear that positivity will be guaranteed in some neighborhood of the degenerate point.

APPENDIX C: DECOUPLING

If we take one quark mass to infinity and look at a meson made of light quarks, the decoupling paradigm implies that the meson should be unaffected by the heavy quark. The results in Secs. IV A and IV B give an example of this. Taking $m_s \to \infty$ in the pion mass of Eq. (36), gives $M_\pi'' = -M_\pi(0)/8$. This agrees with the result for the degenerate ($m_u = m_d$) pion in $N = 2$, Eq. (27). The agreement is nontrivial, since the light-light part of the chiral Lagrangian starts with different θ dependences in the two cases: $\exp(i\theta/3)$ versus $\exp(i\theta/2)$. The decay constant results trivially agree with decoupling, since they vanish in both cases for degenerate light quarks.

For a more robust check of decoupling, we look at the case of arbitrary $N \geq 3$ quarks, with $k < N$ light degenerate quarks of mass m and $N - k$ degenerate heavy quarks of mass m_h. In the first instance, we proceed exactly as in Secs. IV A and IV B, starting with a factor of $e^{i\theta/N}$ multiplying all quark masses, both light and heavy, as in Eq. (11). Let $\langle \Sigma \rangle$ be given by a generalization of Eq. (29), with $\exp(i\alpha)$ in the first k diagonal entries, and $\exp(-ik\alpha/(N - k))$ in the last $N - k$ diagonal entries. Following the same steps as before, we obtain

$$\alpha' = \frac{m - m_h}{Nm + Nkm_h/(N - k)}. \qquad \text{(C1)}$$

As always, the decay constant for the degenerate quarks vanishes, while the mass obeys

$$M_\pi'' = -\frac{M_\pi(0)}{2} \frac{m_h^2}{((N - k)m + km_h)^2}. \qquad \text{(C2)}$$

Alternatively, following the discussion in Appendix A, we may make a nonanomalous transformation at the start, so that we have a factor of $e^{i\theta/k}$ multiplying the light quark masses only. The vacuum expectation value $\langle \Sigma \rangle$ has the same form in terms of α as above, but now

$$\alpha' = \frac{m}{km + k^2 m_h/(N - k)}. \qquad \text{(C3)}$$

As expected, however, Eq. (C2) is unchanged.

In either case, taking $m_h \to \infty$ gives $M_\pi'' = -M_\pi(0)/(2k^2)$, independent of the existence or number of heavy quarks. The same answer would be obtained if we had set $k = N$ from the beginning, i.e., if we had only (degenerate) light quarks.[6] Thus, no matter how the problem is set up, the heavy quarks decouple when $m_h \to \infty$. The result in that limit is the same as it would be if the heavy quarks were not there at all. Note, however, that the decoupling is subtle: solving the problem with the heavy quarks present and then taking their mass to infinity is not a matter of simply deleting all terms in the original Lagrangian that involve the heavy quarks. This is particularly clear when we set up the problem with $e^{i\theta/N}$ multiplying all quark masses. Simply deleting the heavy quark terms after the setup would give the incorrect result $M_\pi'' = -M_\pi(0)/(2N^2)$.

---

[6]In that case we must immediately put $\alpha = 0$ because a nonzero α is inconsistent with the dual requirements of flavor symmetry and $\det(\Sigma) = 1$.

---

[1] L. D. Debbio, H. Panagopoulos, P. Rossi, and E. Vicari, Phys. Rev. D **65**, 021501 (2001); J. High Energy Phys. 01 (2002) 009; 08 (2002) 044.

[2] M. Lüscher, J. High Energy Phys. 08 (2010) 071; 03 (2014) 092(E); *Proc. Sci.*, LATTICE2010 (2010) 015, arXiv:1009.5877; M. Lüscher and S. Schaefer, J. High Energy Phys. 07 (2011) 036.

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

[4] S. Schaefer, R. Sommer, and F. Virotta, Nucl. Phys. **B845**, 93 (2011).

[5] G. McGlynn and R. D. Mawhinney, Phys. Rev. D **90**, 074502 (2014).

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

[7] S. Aoki and H. Fukaya, Phys. Rev. D **81**, 034022 (2010).

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

[9] M. Golterman, S. Sharpe, and R. Singleton, Jr., Phys. Rev. D **71**, 094503 (2005).

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

[11] C. Bernard and D. Toussaint, *Proc. Sci.* LATTICE2016 (2016) 189, arXiv:1611.04522.

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

[13] A. Bazavov et al., arXiv:1712.09262.

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

[15] A. Hasenfratz and F. Knechtli, Phys. Rev. D **64**, 034504 (2001).

[16] T. A. DeGrand, A. Hasenfratz, and T. G. Kovacs, Nucl. Phys. **B505**, 417 (1997); A. Hasenfratz and C. Nieter, Phys. Lett. B **439**, 366 (1998).

[17] C. Alexandrou *et al.*, arXiv:1708.00696.
[18] E. Witten, Nucl. Phys. **B156**, 269 (1979).
[19] G. Veneziano, Nucl. Phys. **B159**, 213 (1979).
[20] H. Leutwyler and A. V. Smilga, Phys. Rev. D **46**, 5607 (1992).
[21] S. Aoki, H. Fukaya, S. Hashimoto, and T. Onogi, Phys. Rev. D **76**, 054508 (2007).
[22] J. Gasser and H. Leutwyler, Nucl. Phys. **B250**, 465 (1985).
[23] C. W. Bernard and M. F. L. Golterman, Phys. Rev. D **49**, 486 (1994).
[24] N. G. de Bruijn, *Asymptotic Methods in Analysis* (Dover, New York, 1981).
[25] See, for example, A. Manohar and M. Wise, *Heavy Quark Physics* (Cambridge University Press, Cambridge, England, 2000).
[26] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **85**, 114506 (2012).
[27] J. Komijani, in *Proceedings of Lattice 2017* (to be published); A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations) (to be published).
[28] C. Aubin and C. Bernard, Nucl. Phys. B, Proc. Suppl. **129–130**, 182 (2004); C. Bernard, Phys. Rev. D **73**, 114503 (2006).

[29] C. Bernard, M. Golterman, and Y. Shamir, Phys. Rev. D **77**, 074505 (2008) and references therein.
[30] C. van den Doel and J. Smit, Nucl. Phys. **B228**, 122 (1983); M. F. L. Golterman and J. Smit, Nucl. Phys. **B245**, 61 (1984).
[31] W.-J. Lee and S. R. Sharpe, Phys. Rev. D **60**, 114503 (1999).
[32] C. Aubin and C. Bernard, Phys. Rev. D **68**, 034014 (2003).
[33] C. Aubin and C. Bernard, Phys. Rev. D **68**, 074011 (2003).
[34] B. Billeter, C. E. Detar, and J. Osborn, Phys. Rev. D **70**, 077502 (2004).
[35] C. Bernard and J. Komijani, Phys. Rev. D **88**, 094017 (2013).
[36] A. Bazavov *et al.*, Phys. Rev. D **90**, 074509 (2014).
[37] J. Towns *et al.*, Comput. Sci. Eng. **16**, 62 (2014).
[38] G. 't Hooft, Phys. Rev. Lett. **37**, 8 (1976); Phys. Rev. D **14**, 3432 (1976); R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37**, 172 (1976); C. Callan, R. Dashen, and D. Gross, Phys. Lett. B **63**, 334 (1976); D. B. Kaplan and A. V. Manohar, Phys. Rev. Lett. **56**, 2004 (1986).
[39] See, for instance, https://en.wikipedia.org/wiki/Fubini's_theorem.