

**$B_K$  in staggered chiral perturbation theory**

Ruth S. Van de Water\* and Stephen R. Sharpe†

*Physics Department, University of Washington, Seattle, Washington 98195-1560, USA*

(Received 13 August 2005; published 9 January 2006)

We calculate the kaon  $B$  parameter,  $B_K$ , to next-to-leading order in staggered chiral perturbation theory. We find expressions for partially quenched QCD with three sea quarks, quenched QCD, and full QCD with  $m_u = m_d \neq m_s$ . We extend the usual power counting to include the effects of using perturbative (rather than nonperturbative) matching factors. Taste breaking enters through the  $\mathcal{O}(a^2)$  terms in the effective action, through  $\mathcal{O}(a^2)$  terms from the discretization of operators, and through the truncation of matching factors. These effects cause mixing with several additional operators, complicating the chiral and continuum extrapolations. In addition to the staggered expressions, we present  $B_K$  at next-to-leading order in continuum PQ $\chi$ PT for  $N_f = 3$  sea quarks with  $m_u = m_d \neq m_s$ .

DOI: 10.1103/PhysRevD.73.014003

PACS numbers: 12.39.Fe, 12.38.Gc

**I. INTRODUCTION**

Experimental measurements of  $CP$  violation can be used to extract information about the Cabibbo-Kobayashi-Maskawa (CKM) matrix. In particular, the size of indirect  $CP$  violation in the neutral kaon system,  $\epsilon_K$ , combined with theoretical input, places an important constraint on the apex of the CKM unitarity triangle [1,2]. Because  $\epsilon_K$  is well known experimentally [3], the dominant source of error in this procedure is the uncertainty in the lattice determination of the nonperturbative constant  $B_K$ , which parametrizes the  $K^0 - \bar{K}^0$  matrix element. Because new physics would likely produce unitarity violation in the CKM matrix and additional  $CP$ -violating phases, a precise determination of  $B_K$  will help to constrain physics beyond the standard model.

Promising calculations with partially quenched staggered fermions are in progress [4,5].<sup>1</sup> Because staggered fermions are computationally cheaper than other standard discretizations, they allow simulations with the lightest dynamical quark masses currently available. Unfortunately, staggered fermions come with their own source of error—taste-symmetry breaking. Each lattice fermion flavor comes in four tastes, which are degenerate in the continuum. The nonzero lattice spacing,  $a$ , breaks the continuum taste symmetry at  $\mathcal{O}(a^2)$ , and the resulting discretization errors are numerically significant at present lattice spacings [7]. Thus one must use the chiral perturbation theory functional forms for *staggered* fermions in order to correctly perform the combined continuum and chiral extrapolations incorporating taste violations [8–10]. It is clear, then, that the potential improvement in the precision of the lattice determination of  $B_K$  is limited without the availability of the appropriate staggered  $\chi$ PT expression.

In this paper we calculate  $B_K$  to next-to-leading order (NLO) in an extended version of staggered chiral perturbation theory (S $\chi$ PT) that incorporates perturbative operator mixing. In the NLO S $\chi$ PT expressions for quantities calculated thus far, e.g.  $f_K$  and  $f_D$ , taste breaking primarily enters through additive corrections to the tree-level masses of pions inside loops [8,9,11]. The calculation of  $B_K$  in S $\chi$ PT provides a more complicated example of taste breaking in which new operators in the staggered chiral effective theory significantly change the result from that in continuum  $\chi$ PT, even at NLO. The bulk of our new work is in correctly enumerating all possible operators which contribute. It turns out that the NLO expression is in terms of 37 low-energy constants (21 for a single lattice spacing), five of which are already known from fitting other staggered lattice data. Because fitting so many parameters requires a large amount of data, we show how to extract additional coefficients using simple kaon matrix elements. Nevertheless, performing the continuum and chiral extrapolations of  $B_K$  will be challenging.

We aim to keep the body of this paper accessible to those wishing to use our NLO S $\chi$ PT expressions for calculations of  $B_K$  on the lattice. Our paper is therefore organized as follows. In Sec. II we define  $B_K$  in QCD, and discuss modifications to the quark-level operator necessary for its calculation with staggered quarks. We summarize the main results of our  $B_K$  operator enumeration in Sec. III. In Sec. IV we calculate  $B_K$  at 1 loop for a partially quenched (PQ) theory with three dynamical quark flavors (each with four tastes). We then add the NLO analytic terms and give results for  $B_K$  to NLO in quenched, partially quenched, and full QCD in Sec. V. Finally, in Sec. VI we explicitly show, for a single lattice spacing, how to determine six of the undetermined NLO coefficients, as well as how to use the S $\chi$ PT fit results to extract the renormalization group invariant quantity  $\hat{B}_K$  in the continuum. We conclude in Sec. VII. As a corollary to our staggered result, we present the expression for  $B_K$  at NLO in the *continuum* PQ $\chi$ PT with  $2 + 1$  sea quarks ( $m_u = m_d \neq m_s$ ) in Appendix A, since it is not available in the literature. Appendix B briefly

\*Electronic address: ruthv@fnal.gov

†Electronic address: sharpe@phys.washington.edu

<sup>1</sup>Calculations of  $B_K$  with dynamical domain wall fermions are also underway [6].

reviews the essentials of staggered chiral perturbation theory necessary for understanding the details of our NLO calculation and operator enumeration. Appendix C enumerates in detail all of the operators which contribute to  $B_K$  at NLO.

## II. $B_K$ WITH STAGGERED QUARKS

In this section we relate  $B_K$  in QCD to matrix elements of a continuum staggered theory in which one has taken the fourth root of the quark determinant but there are still four tastes per valence quark flavor.

The kaon  $B$  parameter is defined as a ratio of matrix elements:

$$\mathcal{M}_K = \langle \overline{K^0} | \mathcal{O}_K | K^0 \rangle = B_K \mathcal{M}_{\text{vac}}, \quad (1)$$

where  $\mathcal{O}_K$  is a weak operator and  $\mathcal{M}_{\text{vac}}$  is the result for  $\mathcal{M}_K$  in the vacuum saturation approximation:

$$\mathcal{O}_K = [\overline{s}_a \gamma_\mu (1 + \gamma_5) d_a] [\overline{s}_b \gamma_\mu (1 + \gamma_5) d_b], \quad (2)$$

$$\begin{aligned} \mathcal{M}_{\text{vac}} &= \frac{8}{3} \langle \overline{K^0} | [\overline{s}_a \gamma_\mu (1 + \gamma_5) d_a] | 0 \rangle \\ &\quad \times \langle 0 | [\overline{s}_b \gamma_\mu (1 + \gamma_5) d_b] | K^0 \rangle. \end{aligned} \quad (3)$$

Note that there are separate summations over color indices  $a$  and  $b$ . Thus the matrix element  $\mathcal{M}_K$  receives contributions from two different quark-level contractions, one of which produces a single color loop and the other of which has two color loops. The expression in Eq. (1) can be simplified using the fact that  $\mathcal{M}_{\text{vac}}$  is related to the square of the kaon decay constant:

$$\mathcal{M}_K = \frac{8}{3} B_K m_K^2 f_K^2, \quad (4)$$

where we use the normalization that  $f_K \approx 156$  MeV.

In order to calculate  $B_K$  with *staggered* fermions we must introduce the *taste* degree of freedom, both in the operators ( $\mathcal{O}_K$ ) and in the states ( $K^0$  and  $\overline{K^0}$ ). We choose the external staggered kaons to be taste  $P$ , by which we mean that the lattice meson operator contains the pseudo-scalar taste matrix  $\xi_5$ . Since this is the lattice Goldstone taste, its correlation functions satisfy  $U(1)_A$  Ward identities, so S $\chi$ PT expressions for its mass, decay constant, and other physical quantities are simpler than those for other tastes of pseudo-Goldstone bosons (PGBs). In addition, it is a local kaon on the lattice, and therefore relatively simple to implement numerically. Because the external kaons are taste  $P$ , the weak operator should also be taste  $P$ :

$$\begin{aligned} \mathcal{O}_K^{\text{naive}} &= [\overline{s}_a (\gamma_\mu (1 + \gamma_5) \otimes \xi_5) d_a] \\ &\quad \times [\overline{s}_b (\gamma_\mu (1 + \gamma_5) \otimes \xi_5) d_b]. \end{aligned} \quad (5)$$

However, when this operator is Fierz transformed, not only do its color indices change, but it mixes with other tastes. To remedy this, we introduce two types of valence quarks, 1 and 2, into the  $B_K$  operator:

$$\begin{aligned} \mathcal{O}_K^{\text{staggered}} &= 2\{[\overline{s}_1 (\gamma_\mu (1 + \gamma_5) \otimes \xi_5) d_{1a}] \\ &\quad \times [\overline{s}_2 (\gamma_\mu (1 + \gamma_5) \otimes \xi_5) d_{2b}] \\ &\quad + [\overline{s}_1 (\gamma_\mu (1 + \gamma_5) \otimes \xi_5) d_{1b}] \\ &\quad \times [\overline{s}_2 (\gamma_\mu (1 + \gamma_5) \otimes \xi_5) d_{2a}]\}, \end{aligned} \quad (6)$$

and take the matrix element between two types of kaons:

$$\mathcal{M}_K^{\text{staggered}} = \langle \overline{K_{1P}^0} | \mathcal{O}_K^{\text{staggered}} | K_{2P}^0 \rangle, \quad (7)$$

where  $K_1^0$  is an  $\overline{s}1d1$  meson and  $K_2^0$  is an  $\overline{s}2d2$  meson. The extra valence quarks require explicit inclusion of both color contractions in  $\mathcal{O}_K^{\text{staggered}}$ , while the overall factor of 2 ensures that  $\mathcal{M}_K$  and  $\mathcal{M}_K^{\text{staggered}}$  have the same total number of contractions. When  $\mathcal{O}_K^{\text{staggered}}$  is Fierz transformed, it has a new *flavor* structure so it does not contribute to  $\mathcal{M}_K^{\text{staggered}}$ . Thus  $\mathcal{M}_K^{\text{staggered}}$  cannot receive contributions from incorrect tastes. We choose that the two sets of valence quarks have equal masses, i.e.  $m_{d1} = m_{d2} \equiv m_x$  and  $m_{s1} = m_{s2} \equiv m_y$ , in order to simplify future expressions.

The continuum staggered theory differs from QCD in that it has four copies of each valence quark, so  $\mathcal{M}_K$  and  $\mathcal{M}_K^{\text{staggered}}$  are not precisely equal. Nevertheless they are simply related by overall normalization factors. Specifically,

$$\mathcal{M}_K = \frac{1}{N_t} \mathcal{M}_K^{\text{staggered}}, \quad (8)$$

$$\langle 0 | \overline{s} \gamma_\mu \gamma_5 d | K^0 \rangle = \sqrt{\frac{1}{N_t}} \langle 0 | \overline{s} (\gamma_\mu \gamma_5 \otimes \xi_5) d | K_P^0 \rangle, \quad (9)$$

where  $N_t = 4$  is the number of tastes per valence quark flavor [12]. We emphasize that these expressions assume that one has already taken the fourth root of the quark determinant and the continuum limit, and that the  $\sqrt[4]{\det}$  procedure is valid. We now reexpress  $B_K$  in terms of the staggered matrix elements calculated on the lattice:

$$B_K = \frac{\frac{1}{N_t} \langle \overline{K_{1P}^0} | \mathcal{O}_K^{\text{staggered}} | K_{2P}^0 \rangle}{\frac{8}{3} \sqrt{\frac{1}{N_t}} \langle \overline{K_P^0} | \overline{s} (\gamma_\mu \gamma_5 \otimes \xi_5) d | 0 \rangle \sqrt{\frac{1}{N_t}} \langle 0 | \overline{s} (\gamma_\mu \gamma_5 \otimes \xi_5) d | K_P^0 \rangle}. \quad (10)$$

Because the various factors of  $N_t$  cancel,  $B_K$  in QCD is equal to the value of the analogous ratio of staggered matrix elements.

Current calculations of  $B_K$  with staggered quarks use the partially quenched approximation. The quark content of the corresponding PQ theory is quite large. We have al-

ready seen that staggering requires two sets of  $d$  and  $s$  valence quarks. Partial quenching adds two corresponding sets of ghost quarks as well as three sea quarks, resulting in eleven total quark flavors, each of which comes in four tastes. In order to make this completely clear, we show the explicit form of the quark mass matrix:

$$M = \text{diag}\{\underbrace{m_u I, m_d I, m_s I}_{\text{sea}}, \underbrace{m_x I, m_y I}_{\text{valence 1}}, \underbrace{m_x I, m_y I}_{\text{valence 2}}, \underbrace{m_x I, m_y I}_{\text{ghost 1}}, \underbrace{m_x I, m_y I}_{\text{ghost 2}}\}, \quad (11)$$

where  $I$  is the  $4 \times 4$  identity matrix. Thus the generalization of  $B_K$  to PQ staggered quarks is relatively straightforward conceptually, although nontrivial to implement on the lattice.

### III. GENERALIZED STAGGERED CHIRAL PERTURBATION THEORY FOR $B_K$

The previous section described how to calculate  $B_K$  with staggered quarks in the continuum, but our expression for  $B_K$  in S $\chi$ PT must describe  $B_K$  at  $a \neq 0$  if it is to be used for continuum and chiral extrapolations of lattice data. It is therefore necessary to discuss in some detail how  $B_K$  is actually calculated on the lattice.

The  $B_K$  matrix element,  $\mathcal{M}_K^{\text{staggered}}$ , receives a contribution from the lattice version of  $\mathcal{O}_K^{\text{staggered}}$ , as well as from other lattice operators that are in the same representation of the symmetry group that maps a hypercube onto itself [13]:

$$\begin{aligned} \mathcal{O}_K^{\text{staggered,cont}} &= \mathcal{O}_K^{\text{staggered,lat}} + \frac{\alpha}{4\pi} [\text{taste } P \text{ ops.}] \\ &+ \frac{\alpha}{4\pi} [\text{other taste ops.}] + \alpha^2 [\text{all taste ops.}] \\ &+ \alpha^2 [\text{all taste ops.}] + \dots, \end{aligned} \quad (12)$$

where  $\alpha$  is the strong coupling constant. The 1-loop perturbative matching coefficients between  $\mathcal{O}_K^{\text{staggered}}$  in the continuum and four-fermion lattice operators are known, and are numbers of order unity times  $\alpha/4\pi$  [14]. However, the 2-loop matching coefficients have not been determined, so, in order to remain conservative, we consider them to be of order unity times  $\alpha^2$  without any factors of  $4\pi$ . Current numerical staggered calculations are in fact of the following matrix element:

$$\langle \overline{K}_{1P}^0 | \mathcal{O}_{1\text{-loop}}(\text{taste } P) | K_{2P}^0 \rangle \equiv \mathcal{M}_{\text{lat}}, \quad (13)$$

where the subscript ‘‘1-loop’’ and the argument ‘‘taste  $P$ ’’ indicate that one includes all staggered lattice operators with taste  $P$  that mix with the latticized  $B_K$  operator at  $\mathcal{O}(\alpha/4\pi)$ , i.e. those in the second term on the right-hand side of Eq. (12), using the appropriate matching coefficients. However, one neglects wrong-taste and higher-order perturbative mixing (terms three and four), as well as all operators which arise through discretization effects

(term five), in Eq. (12). Although the expression in Eq. (13) differs from the continuum matrix element, it reduces to the desired quantity in the continuum limit. Generically,

$$\mathcal{M}_{\text{lat}} = \mathcal{M}_{\text{cont}} + \frac{\alpha}{4\pi} \mathcal{M}' + \alpha^2 \mathcal{M}'' + \alpha^2 \mathcal{M}''' + \dots, \quad (14)$$

where  $\mathcal{M}_{\text{cont}}$  is the desired continuum result. The matrix element  $\mathcal{M}'$  comes from neglecting taste-violating 1-loop operator mixing, while  $\mathcal{M}''$  comes from neglecting 2-loop operator mixing. Both taste-breaking and taste-conserving discretization errors generate  $\mathcal{M}'''$ .

We are now in a position to determine the appropriate power-counting scheme for calculating  $B_K$  at next-to-leading order in S $\chi$ PT. Clearly it must incorporate  $a^2$ ,  $\alpha/4\pi$ , and  $\alpha^2$ . Continuum  $\chi$ PT is a low-energy expansion in both the PGB momentum and the quark masses; it assumes that  $p_{\text{PGB}}^2/\Lambda_\chi^2 \sim m_{\text{PGB}}^2/\Lambda_\chi^2$ , where  $m_{\text{PGB}}^2 \propto m_q$  and  $\Lambda_\chi \approx 4\pi f_\pi$  is the  $\chi$ PT scale. However, the numerical values of  $m_q$ ,  $a^2$ , and  $\alpha$  all depend on the particular parameters of a given lattice simulation. Current PQ staggered lattice simulations [7] use a range of PGB masses from  $m_{\text{PGB}}^2/\Lambda_\chi^2 = 0.04\text{--}0.2$ , so our S $\chi$ PT expression must apply throughout this range. Generic discretization errors are of the size  $a^2 \Lambda_{\text{QCD}}^2$ , which is approximately 0.04 for  $1/a \sim 2$  GeV and  $\Lambda_{\text{QCD}} \sim 400$  MeV, so they are comparable to the minimal  $m_{\text{PGB}}^2/\Lambda_\chi^2$  and should be included at the same order.<sup>2</sup> Taste-breaking discretization errors, on the other hand, are caused by exchange of gluons with momentum  $\pi/a$ , and therefore receive an additional factor of  $\alpha_V^2(\pi/a)$ , so their size must be considered separately.<sup>3</sup> At the lightest quark masses, the lattice Goldstone pion mass is comparable to the mass splittings among the other PGB tastes:  $m_{\text{PGB}}^2/\Lambda_\chi^2 \sim a^2 \alpha_V^2(q^* = \pi/a) \Lambda^2 \sim 0.04$ , where  $\Lambda$  is a QCD scale (distinct from  $\Lambda_{\text{QCD}}$ ) which turns out to be around 1200 MeV. Thus they are not suppressed

<sup>2</sup>Our generic discretization errors are not suppressed by  $\alpha$  because we consider unimproved operators. This is in contrast to the mass and decay constant calculations of Ref. [7], which use improved operators and therefore have smaller generic discretization errors of around 2%.

<sup>3</sup>Taste violations are suppressed by  $\alpha^2$  rather than  $\alpha$  because we assume an improved action.

relative to pure  $\mathcal{O}(a^2)$  discretization effects by the additional powers of  $\alpha$ , as naive power counting would suggest, because of the large scale  $\Lambda$  associated with the taste-breaking process at the quark level [15]. Standard S $\chi$ PT only includes discretization effects—we now consider additional errors from perturbative operator matching, which depend upon  $\alpha_V(q^*)$ . Generically, the choice of  $q^*$  is process dependent, and the value of  $\alpha_V(q^*)$  ranges from  $\sim 0.3$ – $0.55$  for  $q^* = \pi/a - 1/a$  at  $a = .125$  fm [16]. Thus both  $\alpha/4\pi$  and  $\alpha^2$  must be included at lowest order in our power counting.<sup>4</sup>

In light of this discussion, we adopt the following extended S $\chi$ PT power-counting scheme:

$$p^2 \sim m \sim a^2 \sim a_\alpha^2 \sim \alpha/4\pi \sim \alpha^2, \quad (15)$$

where  $a_\alpha^2 \equiv \alpha_V^2(\pi/a)a^2$ . We account for the fact that  $\alpha^2 a^2$  terms in the action are enhanced numerically by including them at the same order as simple discretization effects.<sup>5</sup> In fact, while it may seem *ad hoc*,  $p^2 \sim m \sim a_\alpha^2$  is the *standard* S $\chi$ PT power-counting scheme. It is simply not traditionally written as such because standard S $\chi$ PT calculations have only included  $\mathcal{O}(a_\alpha^2)$  taste-breaking discretization errors from the action, and have therefore not needed to contrast them with pure  $\mathcal{O}(a^2)$  discretization effects. We also use conservative power counting for the perturbative errors by assuming that 2-loop contributions

are not significantly smaller than those from 1-loop diagrams. We emphasize that our scheme is *phenomenologically based* on the particular parameter values of current staggered simulations—simulations using significantly lighter quark masses or smaller lattice spacings would require a different scheme and result in a different S $\chi$ PT expression for  $B_K$ .

We must determine all of the contributions at NLO in our power counting to  $\mathcal{M}_{\text{lat}}$ , including the effects of perturbative matching and discretization errors. We then find  $B_K$  by using  $\mathcal{M}_{\text{lat}}$  in the ratio given in Eq. (10). In S $\chi$ PT,  $\mathcal{M}_{\text{lat}}$  is simply the matrix element of a sum of operators with undetermined coefficients:

$$\mathcal{M}_{\text{lat}} = \langle \overline{K}_{1P}^0 | C_\chi^i \mathcal{O}_\chi^i | K_{2P}^0 \rangle, \quad (16)$$

so our goal is really to determine all *operators* in the chiral effective theory,  $\mathcal{O}_\chi^i$ , that contribute at NLO to the above matrix element. Once we do so, the full NLO S $\chi$ PT expression will contain the LO tree-level term, 1-loop corrections, and analytic terms of the same order as the 1-loop terms.<sup>6</sup> This task is not as daunting as it first seems given the extended power-counting scheme because it turns out that the only LO contribution to  $B_K$  is of  $\mathcal{O}(p^2)$ . Thus operators of the following orders in our power counting contribute to  $B_K$  at NLO:

$$\begin{aligned} \text{1-loop: } & \mathcal{O}(p^2), \mathcal{O}(m), \mathcal{O}(a^2), \mathcal{O}(a_\alpha^2), \mathcal{O}(\alpha/4\pi), \mathcal{O}(\alpha^2), \\ \text{NLO analytic: } & \mathcal{O}(p^4), \mathcal{O}(p^2 m), \mathcal{O}(p^2 a^2), \mathcal{O}(p^2 a_\alpha^2), \mathcal{O}(p^2 \alpha/4\pi), \mathcal{O}(p^2 \alpha^2), \mathcal{O}(m^2), \mathcal{O}(ma^2), \mathcal{O}(ma_\alpha^2), \\ & \mathcal{O}(m\alpha/4\pi), \mathcal{O}(m\alpha^2), \mathcal{O}(a^4), \mathcal{O}(a_\alpha^4), \mathcal{O}(\alpha^2/16\pi^2), \mathcal{O}(\alpha^4). \end{aligned} \quad (17)$$

In the rest of this section we give an overview of the operators that contribute to  $B_K$  at 1 loop in S $\chi$ PT. Appendix C contains the detailed operator enumeration as well as the determination of analytic terms.

Clearly the lattice version of the continuum quark-level operator  $\mathcal{O}_K^{\text{staggered}}$ , Eq. (6), will generate the dominant contribution to  $B_K$ , so we map  $\mathcal{O}_K^{\text{staggered}}$  onto chiral operators first. This is relatively straightforward; we follow the graded group-theory method of Refs. [10,17] to insure that we include all possible linearly independent operators. This results in two operators, Eqs. (C7) and (C8), the first of which is simply the staggered, partially quenched analog of the continuum ( $27_L, 1_R$ ) chiral operator, and the second of which arises because we use two sets of valence quarks. Because both chiral operators come from the same quark-

level operator, their coefficients are related, and turn out to be equal. This is crucial because  $B_K$  in the continuum  $\chi$ PT only depends on a single parameter [18], and our result must match the continuum one when  $a \rightarrow 0$ . Only the “standard” chiral operator contributes to  $B_K$  at tree level, and its contribution is of  $\mathcal{O}(p^2)$ .

Next we determine the chiral operators which contribute to  $B_K$  because of perturbative operator mixing. Four-fermion operators which mix with  $\mathcal{O}_K^{\text{staggered}}$  but are *not taste P* are unaccounted for in Eq. (13), so they introduce errors into  $B_K$  that are of  $\mathcal{O}(\alpha/4\pi)$ , as well as errors of  $\mathcal{O}(\alpha^2)$ .<sup>7</sup> We therefore map them onto chiral operators with two undetermined coefficients of  $\mathcal{O}(\alpha/4\pi)$  and  $\mathcal{O}(\alpha^2)$ , respectively. The resulting eight chiral operators are given in Eqs. (C16)–(C23). Taste *P* four-fermion operators are accounted for at  $\mathcal{O}(\alpha/4\pi)$  using the matching coefficients, but they still introduce errors into  $B_K$  at  $\mathcal{O}(\alpha^2)$ , through a new  $\mathcal{O}(\alpha^2)$  operator given in Eq. (C14).

<sup>4</sup>The quoted values of  $\alpha$  indicate that it is perhaps necessary to include  $\mathcal{O}(\alpha^3)$  errors as well; we assume that this is not the case.

<sup>5</sup>We assume that taste-conserving discretization errors are not enhanced since such behavior has not been observed in staggered lattice simulations.

<sup>6</sup>Note the distinction between  $B_K$  at NLO and  $B_K$  at 1 loop, which does not contain the analytic terms.

<sup>7</sup>Recall that  $\alpha/4\pi \sim \alpha^2$  in our power counting because we make no assumption about the numerical size of the  $\mathcal{O}(\alpha^2)$  matching coefficients.

To match the continuum regularized operator,  $\mathcal{O}_K^{\text{staggered,cont}}$ , we also need to include operators of higher dimension. From now on we will generically refer to this matching of the continuum onto lattice operators as “mixing.” In particular, we must include all dimension 8 operators that are in the same representation of the symmetry group that maps a hypercube onto itself. Because dimension 8 operators are explicitly suppressed by a factor of  $a^2$  relative to dimension 6 operators, they can map onto chiral operators with coefficients of  $\mathcal{O}(a^2)$  that contribute to  $B_K$  at NLO. There turn out to be nine such operators—the same as those which arise from perturbative operator mixing.

Finally, four-fermion operators from the staggered action that arise due to discretization effects can enter time-ordered products with the  $B_K$  operator  $\mathcal{O}_K^{\text{staggered}}$ , effectively correcting the  $B_K$  four-fermion vertex at  $\mathcal{O}(a_\alpha^2)$ . We must therefore determine the chiral operators which result from a combination of the  $B_K$  four-fermion operator and the inserted four-fermion operator from the action. We do so using the spurion method of Ref. [10]. This results in four new chiral operators, Eqs. (C26)–(C29), as well additional copies of seven operators which had appeared previously from operator mixing.

In total, we find 15 chiral operators; we will determine their contributions to  $B_K$  at 1 loop in the following section. We note, however, that many of these chiral operators arise in more than one way and correspond to more than one quark-level operator, so they actually have *more than one undetermined coefficient*. We list the  $\mathcal{O}(\alpha/4\pi)$ ,  $\mathcal{O}(\alpha^2)$ ,  $\mathcal{O}(a^2)$ , and  $\mathcal{O}(a_\alpha^2)$  operators and their coefficients in Table IV. For example, the operator  $\mathcal{O}_\chi^{1A}$ , which comes both from 1-loop perturbative matching and from insertions of operators in the staggered action, has three coefficients, one multiplied by  $\alpha/4\pi$ , one multiplied by  $\alpha^2$ , and one multiplied by  $a^2$ . The three coefficients can be separated, in principle, by carrying out simulations at more

than one lattice spacing. We emphasize that they *cannot* be lumped together in a fit to multiple lattice spacings. In the following sections, including the final expressions for  $B_K$  at NLO, we only associate a single coefficient,  $C_\chi^i$ , with each operator. This is simply to reduce the size of the final expressions. One must combine the expressions with Table IV in order to make them complete.

### $B_K$ AT 1 LOOP FOR 4 + 4 + 4 DYNAMICAL FLAVORS

We first calculate  $B_K$  at tree level, which receives a contribution from a single diagram, Fig. 1(a), and from a single chiral operator:

$$\mathcal{O}_\chi^K = \sum_\mu (\text{Str}(\Sigma \partial_\mu \Sigma^\dagger F_1) \text{Str}(\Sigma \partial_\mu \Sigma^\dagger F_2) + \text{p.c.}) \quad (18)$$

In the diagram, the  $B_K$  chiral operator is shown as two disconnected squares, each of which produces a kaon. This reflects the structure of  $\mathcal{O}_\chi^K$ , which has two supertraces in which the matrices  $F_1$  and  $F_2$  have the appropriate taste ( $\xi_5$ ) and flavor ( $s1\bar{d}1$  and  $s2\bar{d}2$ ) structures to produce a  $K_1^0$  and a  $K_2^0$ .  $\mathcal{O}_\chi^K$  is just the staggered analog of the continuum  $(27_L, 1_R)$   $B_K$  operator—it reduces to the standard form after removing the taste structure ( $\xi_5 \rightarrow \xi_I$ ) and the partial quenching ( $d1, d2 \rightarrow d$  and  $\text{Str} \rightarrow \text{Tr}$ ). The tree-level  $B_K$  matrix element is simple:

$$\begin{aligned} \mathcal{M}_K^{\text{LO}} &= -\frac{2C_\chi^K}{f^2} \sum_\mu \langle \bar{K}_1^0 | \text{Str}(\partial_\mu \pi F_1) \text{Str}(\partial_\mu \pi F_2) | K_2^0 \rangle \\ &= -\frac{32C_\chi^K}{f^2} \partial_\mu \bar{K}_1^0 \partial_\mu K_2^0 = \frac{32C_\chi^K}{f^2} m_{xy}^2, \end{aligned} \quad (19)$$

where  $m_{xy}^2 = \mu(m_x + m_y)$  is the tree-level kaon mass. The kaon  $B$  parameter itself is just the ratio  $\mathcal{M}_K/\mathcal{M}_{\text{vac}}$ ,

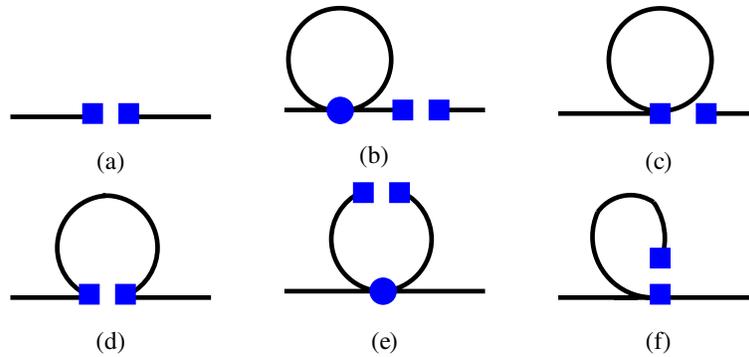


FIG. 1 (color online). Tree-level and 1-loop contributions to  $\mathcal{M}_K$ . In (a)–(e), one external PGB is a  $\bar{K}_1^0$  and the other is a  $K_2^0$ . (b) generates kaon wave function renormalization. The circle represents a vertex from the LO staggered chiral Lagrangian. The squares represent an insertion of the  $B_K$  operator. Each box “changes” the quark flavor from  $d \leftrightarrow s$ . (f) does not actually contribute in the staggered case because it has the wrong flavor structure to contract with two different external kaons.

where  $\mathcal{M}_{\text{vac}}$  is given in Eq. (3).<sup>8</sup> Because the matrix element  $\mathcal{M}_{\text{vac}}$  is proportional to the square of the axial current, it is proportional to  $f^2 p^2$  at tree level:

$$\mathcal{M}_{\text{vac}}^{\text{LO}} = \frac{8}{3} f^2 m_{xy_p}^2. \quad (20)$$

Therefore the kaon mass drops out of the ratio of matrix elements:

$$\left(\frac{\mathcal{M}_K}{\mathcal{M}_{\text{vac}}}\right)^{\text{LO}} = \frac{12}{f^4} C_X^K \equiv B_0, \quad (21)$$

and  $B_K$  is just a constant at lowest order in  $S\chi\text{PT}$ .

We are now ready to calculate  $B_K$  at 1 loop in  $S\chi\text{PT}$ . It is useful to simplify the form of  $B_K$  as much as possible *diagrammatically* before explicitly showing any expressions. In general,  $\mathcal{M}_K$  receives 1-loop contributions from the five diagrams shown in Figs. 1(b)–1(f) [19]. However, one can easily show that Fig. 1(f) cannot contribute in the staggered case because it is impossible to draw a corresponding quark-line diagram with two different external kaons. Moreover, because  $B_K$  is defined by the ratio  $\mathcal{M}_K/\mathcal{M}_{\text{vac}}$ , cancellations occur between matrix elements in the same manner as the kaon mass dropped out at tree level. It turns out that Figs. 1(b) and 1(c) cancel entirely in the  $B_K$  ratio, as we now show.

The 1-loop matrix element  $\mathcal{M}_K$  can be expressed as

$$\mathcal{M}_K = \frac{8}{3} B_0 f^2 m_{xy_p}^2 \{1 + X[\text{Figs. 1(b)–1(c)}]\} + X'[\text{Figs. 1(d)–(e)}], \quad (22)$$

where  $X$  and  $X'$  denote the results of the specified diagrams and  $m_{xy_p}^2$  is the 1-loop kaon mass squared.<sup>9</sup> The NLO expression for  $\mathcal{M}_{\text{vac}}$  has the same form as the tree-level one:

$$\mathcal{M}_{\text{vac}} = \frac{8}{3} m_{xy_p}^2 f_{xy_p}^2, \quad (23)$$

except that both  $m_{xy_p}^2$  and  $f_{xy_p}$  become the 1-loop quantities. Now consider Figs. 1(b) and 1(c). Visually, it is easy to see that they *factorize*—if one draws a vertical line between the boxes, the left halves of the two diagrams renormalize  $f_K$  at 1 loop while the right halves just produce  $f$  at tree level. It therefore seems natural that  $X[\text{Figs. 1(b)–(c)}]$  should be related to the 1-loop renormalization of the kaon decay constant, and, if one works out the details, one finds that

<sup>8</sup>Because we work only in the staggered theory for the rest of the paper, we drop the superscript “staggered” from both operators and matrix elements in order to reduce clutter in expressions.

<sup>9</sup>Note that  $X'$  is not within the curly braces because it comes from operators other than  $\mathcal{O}_X^K$  and is therefore not proportional to  $B_0$ .

$$X[\text{Figs. 1(b)–(c)}] = 2 \frac{\delta f_{\text{NLO}}}{f}, \quad (24)$$

where the factor of 2 results from the fact that the loop can be on either leg. Thus Figs. 1(b) and 1(c) are exactly those necessary to turn the tree-level  $f^2$  into the 1-loop  $f_{xy_p}^2$ <sup>10</sup>:

$$\mathcal{M}_K = \frac{8}{3} B_0 f_{xy_p}^2 m_{xy_p}^2 + X'[\text{Figs. 1(d)–(e)}], \quad (25)$$

and  $B_K$  only depends on Figs. 1(d) and 1(e):

$$B_K^{1\text{-loop}} = B_0 + \frac{3}{8} \frac{X'[\text{Figs. 1(d)–(e)}]}{f_{xy_p}^2 m_{xy_p}^2}. \quad (26)$$

The diagrams in Figs. 1(d) and 1(e) receive contributions from all of the enumerated  $B_K$  chiral operators, which we list here for convenience:

$$\begin{aligned} \mathcal{O}_X^K &= \sum_{\mu} (\text{Str}(\Sigma \partial_{\mu} \Sigma^{\dagger} F_1) \text{Str}(\Sigma \partial_{\mu} \Sigma^{\dagger} F_2) \\ &\quad + \text{Str}(\Sigma^{\dagger} \partial_{\mu} \Sigma F_1) \text{Str}(\Sigma^{\dagger} \partial_{\mu} \Sigma F_2)), \\ \mathcal{O}_X^N &= \sum_{\mu} (\text{Str}(\Sigma \partial_{\mu} \Sigma^{\dagger} F_1 \Sigma \partial_{\mu} \Sigma^{\dagger} F_2) \\ &\quad + \text{Str}(\Sigma^{\dagger} \partial_{\mu} \Sigma F_1 \Sigma^{\dagger} \partial_{\mu} \Sigma F_2)), \\ \mathcal{O}_X^{1P} &= \text{Str}(F_1 \Sigma F_2 \Sigma^{\dagger}), \\ \mathcal{O}_X^{2P} &= \text{Str}(\xi_5 \Sigma F_1 \Sigma^{\dagger} \xi_5 \Sigma F_2 \Sigma^{\dagger}) + \text{p.c.}, \\ \mathcal{O}_X^{3P} &= \text{Str}(\xi_5 \Sigma F_1 \Sigma^{\dagger}) \text{Str}(\xi_5 \Sigma F_2 \Sigma^{\dagger}) + \text{p.c.}, \\ \mathcal{O}_X^{1I} &= \text{Str}(F_{1I} \Sigma F_{2I} \Sigma^{\dagger}), \\ \mathcal{O}_X^{1T} &= \sum_{\mu \neq \nu} \text{Str}(F_{1T} \Sigma F_{2T} \Sigma^{\dagger}), \\ \mathcal{O}_X^{2T} &= \sum_{\mu \neq \nu} \text{Str}(\xi_{\mu\nu} \Sigma F_1 \Sigma^{\dagger} \xi_{\nu\mu} \Sigma F_2 \Sigma^{\dagger}) + \text{p.c.}, \\ \mathcal{O}_X^{3T} &= \sum_{\mu \neq \nu} \text{Str}(\xi_{\mu\nu} \Sigma F_1 \Sigma^{\dagger}) \text{Str}(\xi_{\nu\mu} \Sigma F_2 \Sigma^{\dagger}) + \text{p.c.}, \\ \mathcal{O}_X^{1V} &= \sum_{\mu} (\text{Str}(F_{1V} \Sigma F_{2V} \Sigma) + \text{Str}(F_{1V} \Sigma^{\dagger} F_{2V} \Sigma^{\dagger})), \\ \mathcal{O}_X^{2V} &= \sum_{\mu} (\text{Str}(F_{1V} \Sigma) \text{Str}(F_{2V} \Sigma) + \text{Str}(F_{1V} \Sigma^{\dagger}) \text{Str}(F_{2V} \Sigma^{\dagger})), \\ \mathcal{O}_X^{3V} &= \sum_{\mu} \text{Str}(F_{1V} \Sigma) \text{Str}(F_{2V} \Sigma^{\dagger}), \\ \mathcal{O}_X^{1A} &= \sum_{\mu} (\text{Str}(F_{1A} \Sigma F_{2A} \Sigma) + \text{Str}(F_{1A} \Sigma^{\dagger} F_{2A} \Sigma^{\dagger})), \\ \mathcal{O}_X^{2A} &= \sum_{\mu} (\text{Str}(F_{1A} \Sigma) \text{Str}(F_{2A} \Sigma) + \text{Str}(F_{1A} \Sigma^{\dagger}) \text{Str}(F_{2A} \Sigma^{\dagger})), \\ \mathcal{O}_X^{3A} &= \sum_{\mu} \text{Str}(F_{1A} \Sigma) \text{Str}(F_{2A} \Sigma^{\dagger}). \end{aligned} \quad (27)$$

Appendix C explains the details of our operator notation, but we note here the important features. In general, each

<sup>10</sup>Note that the NLO analytic contributions to  $f_{xy_p}$  simply change the coefficients of the NLO analytic corrections to  $B_K$ .

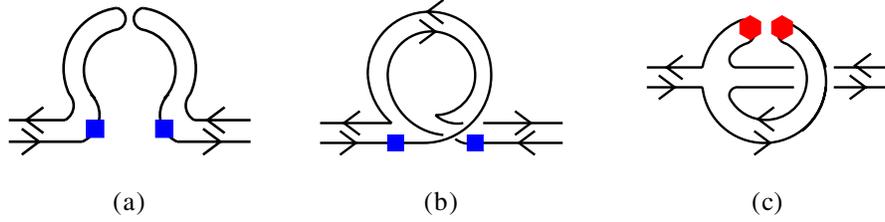


FIG. 2 (color online). Quark-line diagram contributions to  $B_K$  at 1 loop. One external meson is a  $\overline{K}_1^0$  and the other is a  $K_2^0$ . The two boxes (hexagons) represent an insertion of the  $B_K$  operator. Each box changes the valence quark flavor from ( $d1 \leftrightarrow s1$ ) or ( $d2 \leftrightarrow s2$ ). Each hexagon changes the valence quark flavor from ( $d1 \leftrightarrow s2$ ) or ( $d2 \leftrightarrow s1$ ). (a) and (b) contribute to Fig. 1(d) while (c) contributes to Fig. 1(e).

operator contains an  $F_1$  and an  $F_2$ , which are taste matrices with different, nontrivial *flavor* structures such that  $\mathcal{O}_\chi^K$  produces the desired kaons at tree level. In particular,  $F_1$  and  $F_2$  with no additional labels are taste  $\xi_5$ , whereas the remaining  $F$ 's have “incorrect” tastes, e.g.  $F_{1T}$  is taste  $\xi_{\mu\nu}$ . The fact that many operators with tastes other than  $P$  contribute to  $B_K$  illustrates the importance of incorporating taste-symmetry breaking when calculating quantities with staggered fermions on the lattice.

Figure 1(d), in which one of the many  $B_K$  operators produces a four-meson vertex and two of the legs are then contracted to form a loop, corresponds to two quark-level diagrams, shown in Figs. 2(a) and 2(b). The two boxes in each diagram are similar, but not identical: one changes a  $d1$  quark into an  $s1$  quark while the other changes a  $d2$  into an  $s2$ . Figure 2(a) *must* be disconnected at the quark level because the two external kaons contain different kinds of valence quarks, so it only receives contributions from double-supertrace operators<sup>11</sup>:

$$\begin{aligned} \mathcal{M}^{(a)} = & \frac{8C_\chi^K}{f^4} \int \frac{d^4q}{(2\pi)^4} (p^2 + q^2) \{D_{xx}^I(q) + D_{yy}^I(q) - 2D_{xy}^I(q)\} - \frac{16(2C_\chi^{2V} + C_\chi^{3V})}{f^4} \int \frac{d^4q}{(2\pi)^4} \{D_{xx}^A(q) + D_{yy}^A(q) - 2D_{xy}^A(q)\} \\ & - \frac{16(2C_\chi^{2A} + C_\chi^{3A})}{f^4} \int \frac{d^4q}{(2\pi)^4} \{D_{xx}^V(q) + D_{yy}^V(q) - 2D_{xy}^V(q)\}. \end{aligned} \quad (28)$$

The symbols  $D^I$ ,  $D^V$ , and  $D^A$  represent the taste  $I$ ,  $V$ , and  $A$  disconnected propagators, respectively—we will show their explicit forms later as needed. Figure 2(b), on the other hand, is connected and only receives contributions from the seven single-supertrace operators<sup>12</sup>:

$$\begin{aligned} \mathcal{M}^{(b)} = & \sum_{B'} \frac{C_\chi^K f^{B'}}{f^4} \int \frac{d^4q}{(2\pi)^4} \left\{ \frac{2q^2}{3} \left[ \frac{1}{q^2 + m_{xy_{B'}}^2} \right] - \frac{(q^2 + p^2)}{4} \left[ \frac{1}{q^2 + m_{xx_{B'}}^2} + \frac{1}{q^2 + m_{yy_{B'}}^2} \right] \right\} \\ & + \sum_{B, B'} \frac{C_\chi^{1B} g^{BB'}}{f^4} \int \frac{d^4q}{(2\pi)^4} \left\{ \frac{1}{3} \left[ \frac{1}{q^2 + m_{xy_{B'}}^2} \right] - \frac{1}{4} \left[ \frac{1}{q^2 + m_{xx_{B'}}^2} + \frac{1}{q^2 + m_{yy_{B'}}^2} \right] \right\} \\ & + \sum_{B, B'} \frac{C_\chi^{2B} h^{BB'}}{f^4} \int \frac{d^4q}{(2\pi)^4} \left\{ \frac{2}{3} \left[ \frac{1}{q^2 + m_{xy_{B'}}^2} \right] - \frac{1}{4} \left[ \frac{1}{q^2 + m_{xx_{B'}}^2} + \frac{1}{q^2 + m_{yy_{B'}}^2} \right] \right\}, \end{aligned} \quad (29)$$

where the coefficients  $f^{B'}$ ,  $g^{BB'}$ , and  $h^{BB'}$  are matrices in which  $B = I, P, V, A, T$  labels the contributing operator and  $B' = I, P, V, A, T$  indicates the loop meson taste:

$$f^{B'} = -8(1 \ 1 \ 4 \ 4 \ 6), \quad (30)$$

$$g^{BB'} = -4 \begin{pmatrix} 1 & 1 & -4 & -4 & 6 \\ 1 & 1 & 4 & 4 & 6 \\ -8 & 8 & -16 & 16 & 0 \\ -8 & 8 & 16 & -16 & 0 \\ 12 & 12 & 0 & 0 & -24 \end{pmatrix}, \quad (31)$$

$$h^{BB'} = 128 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 6 & 12 \end{pmatrix}. \quad (32)$$

<sup>11</sup>Double-supertrace operators  $\mathcal{O}_\chi^{3P}$  and  $\mathcal{O}_\chi^{3T}$  cannot contribute to  $B_K$  at NLO because the loop mesons must be connected through a hairpin propagator, and can only be tastes  $I$ ,  $V$ , or  $A$ . For a review of  $S\chi$ PT and a brief discussion of hairpin (quark disconnected) propagators see Appendix B.

<sup>12</sup>Recall that  $\mathcal{O}_\chi^K$  and  $\mathcal{O}_\chi^N$  have the same coefficient, which we call  $C_\chi^K$ .

The sum over  $B'$  accounts for how many times a particular

loop meson taste contributes to the above expression with a given coefficient—this is important because a staggered meson’s mass depends on its taste as well as its flavor.<sup>13</sup> The pattern  $\{1, 1, 4, 4, 6\}$  in Eq. (30) counts the number of lattice PGBs with each taste—thus the coefficient  $f^{B'}$  averages over all PGB tastes. Note that the rows of zeros in Eq. (32) have no deep physical meaning, but simply correspond to the absence of certain operators, e.g.  $h^{0B} = (0, 0, 0, 0, 0)$  indicates that  $\mathcal{O}_\chi^{2I}$  does not exist. Nevertheless we include them so as to allow both  $B$  and  $B'$  to run over the same set of indices.

Figure 1(e), in which an operator from the staggered chiral Lagrangian generates the four-meson vertex and one of the  $B_K$  operators is inserted in the loop, corresponds to the third quark-level diagram, Fig. 2(c). Because the LO staggered chiral Lagrangian contains three types of operators—the continuum two-derivative term, the continuum mass term, and the  $\mathcal{O}(a^2)$  potential—there are many contributions to this diagram. Nevertheless, although the staggered potential contains both single- and double-supertrace operators, we only show a connected vertex in Fig. 2(c).<sup>14</sup> This is because it turns out that, for all disconnected 4-PGB vertices from the LO staggered potential in which two of the legs have the desired flavors of external kaons, the various Wick contractions cancel and the vertex is identically zero. Thus we need only consider connected 4-PGB vertices. Because the two external kaons have entirely different flavors, they cannot share any quark lines without a  $B_K$  operator insertion (box or hexagon), so Fig. 2(c) shows the only possible quark flow. However, it is now easy to see that the  $B_K$  operator must change both the quark flavor and the valence set, so we use hexagons rather than boxes to indicate this. Finally, at the 2-PGB level, double-supertrace  $B_K$  operators (by construction) change flavors within a valence set, whereas single-supertrace operators change flavors between sets of valence quarks. Thus the last diagram only receives contributions from single-supertrace operators:

$$\begin{aligned} \mathcal{M}^{(c)} = & \sum_{B'} \int \frac{d^4 q}{(2\pi)^4} \left[ -C_\chi^K f^{B'} q^2 + \sum_B (C_\chi^{1B} g^{BB'}) \right. \\ & \left. - C_\chi^{2B} h^{BB'} \right] \left[ \frac{p^2 + q^2 - 2m_{xy}^2 + m_{xy'}^2}{6f^4} \right] \\ & \times \left[ \frac{1}{q^2 + m_{xy'}^2} \right]^2. \end{aligned} \quad (33)$$

Note that it depends on the *same* coefficients,  $f$ ,  $g$ , and  $h$  as  $\mathcal{M}^{(b)}$ —this will allow us to combine the two in the following section.

<sup>13</sup>See Ref. [8] for the staggered meson masses at tree level.

<sup>14</sup>The staggered potential is shown explicitly in Appendix B.

## V. NEXT-TO-LEADING-ORDER $B_K$ RESULTS

The 1-loop matrix elements determined in the previous section apply to a PQ theory with four tastes of each sea quark flavor. To go from four tastes per sea quark to one taste per sea quark (from the so-called “4 + 4 + 4 theory” to the “1 + 1 + 1 theory” when  $N_{\text{sea}} = 3$ ), we must multiply every sea quark loop by 1/4. At first glance, the quark-level diagrams which contribute to  $B_K$  at 1 loop, Figs. 2(a)–2(c), do not appear to contain any sea quark loops. However, sea quark loops implicitly enter the disconnected hairpin propagators when one rediagonalizes the mass matrices of the flavor-neutral, taste  $V$ ,  $A$ , and  $I$ , sectors. In general, the forms of quenched, PQ, and full QCD matrix elements only differ for diagrams with sea quark loops, so the final expression for Eq. (28) will change in the three cases, but those for Eqs. (29) and (33) will remain the same.

### A. $B_K$ at NLO in the PQ theory

We first consider  $B_K$  in the 1 + 1 + 1 PQ theory ( $m_u \neq m_d \neq m_s$ ), as it is the most general. We label the valence quarks by  $x$  and  $y$ , and reserve the labels  $u$ ,  $d$ , and  $s$  for the sea quarks. Note that the sea quark masses only show up in the PQ results *implicitly* through the masses of the flavor-neutral PGBs:  $\pi^0$ ,  $\eta$ , and  $\eta'$ . Thus the 1 + 1 + 1 result can be turned into the 2 + 1 ( $m_u = m_d \neq m_s$ ) result simply by changing the expressions for the  $\pi^0$ ,  $\eta$ , and  $\eta'$  masses.

The matrix elements that come from *connected* quark-level diagrams are the most simple to calculate, so we consider them first. Neither Fig. 2(b) nor Fig. 2(c) contains sea quark loops, so we do not have to insert any factors of 1/4 by hand—we need only perform the integrals. To simplify the resulting expressions, we use a condensed notation for the chiral logarithms. As in Refs. [8,9], we define two functions,  $\ell$  and  $\tilde{\ell}$ :

$$\int \frac{d^4 q}{(2\pi)^4} \frac{1}{q^2 + m^2} \equiv \frac{1}{16\pi^2} \ell(m^2), \quad (34)$$

$$\int \frac{d^4 q}{(2\pi)^4} \frac{1}{(q^2 + m^2)^2} \equiv \frac{1}{16\pi^2} \tilde{\ell}(m^2). \quad (35)$$

The remaining integrals can all be expressed in terms of  $\ell$  and  $\tilde{\ell}$ :

$$\int \frac{d^4 q}{(2\pi)^4} \frac{q^2}{q^2 + m^2} = \frac{-m^2}{16\pi^2} \ell(m^2), \quad (36)$$

$$\int \frac{d^4 q}{(2\pi)^4} \frac{q^2}{(q^2 + m^2)^2} = \frac{1}{16\pi^2} (\ell(m^2) - m^2 \tilde{\ell}(m^2)), \quad (37)$$

$$\int \frac{d^4 q}{(2\pi)^4} \frac{q^4}{(q^2 + m^2)^2} = \frac{m^2}{16\pi^2} (m^2 \tilde{\ell}(m^2) - 2\ell(m^2)). \quad (38)$$

Because finite-volume (FV) corrections to  $B_K$  only alter the chiral logarithms by turning integrals into sums, one can use either the infinite volume or FV expressions for  $\ell$  and  $\tilde{\ell}$  as desired:

$$\begin{aligned} \ell(m^2) &= m^2 \left( \ln \frac{m^2}{\mu_{\text{D.R.}}^2} + \delta_1^{FV}(mL) \right), \\ \delta_1^{FV}(mL) &= \frac{4}{mL} \sum_{\vec{r} \neq 0} \frac{K_1(|\vec{r}|mL)}{|\vec{r}|} \end{aligned} \quad (39)$$

$$\begin{aligned} \tilde{\ell}(m^2) &= - \left( \ln \frac{m^2}{\mu_{\text{D.R.}}^2} + 1 \right) + \delta_3^{FV}(mL), \\ \delta_3^{FV}(mL) &= 2 \sum_{\vec{r} \neq 0} K_0(|\vec{r}|mL), \end{aligned} \quad (40)$$

where  $L^3$  is the lattice spatial volume and  $K_0$  and  $K_1$  are Bessel functions of imaginary argument.<sup>15</sup> Thus our final expression for  $B_K$  will be completely general. In terms of  $\ell$  and  $\tilde{\ell}$ , the 1-loop matrix element contribution to  $B_K$  from connected quark-level diagrams,  $\mathcal{M}^{(b)} + \mathcal{M}^{(c)}$ , is

$$\begin{aligned} \mathcal{M}_{\text{conn}} &= \frac{1}{64\pi^2 f^4} \sum_{B'} \left\{ C_\chi^K f^{B'} [m_{X_{B'}}^2 \ell(m_{X_{B'}}^2) + m_{Y_{B'}}^2 \ell(m_{Y_{B'}}^2) - 2m_{xy_{B'}}^2 \ell(m_{xy_{B'}}^2)] \right. \\ &\quad - \sum_B (C_\chi^{1B} g^{BB'} + C_\chi^{2B} h^{BB'}) [\ell(m_{X_{B'}}^2) + \ell(m_{Y_{B'}}^2) - 2\ell(m_{xy_{B'}}^2)] + C_\chi^K f^{B'} m_{xy_p}^2 [\ell(m_{X_{B'}}^2) + \ell(m_{Y_{B'}}^2) + 2\ell(m_{xy_{B'}}^2)] \\ &\quad \left. - 2m_{xy_{B'}}^2 \tilde{\ell}(m_{xy_{B'}}^2) - \sum_B (C_\chi^{1B} g^{BB'} - C_\chi^{2B} h^{BB'}) m_{xy_p}^2 [2\tilde{\ell}(m_{xy_{B'}}^2)] \right\}. \end{aligned} \quad (41)$$

We use a condensed notation for the flavor-neutral mesons, in which  $m_X^2 \equiv m_{xx}^2$  and so forth. Note that the first two lines vanish when  $m_x = m_y$ . Note also that the second and third terms in  $\mathcal{M}_{\text{conn}}$  are proportional to the mass of the external kaons because we have let  $p^2 = -m_{xy_p}^2$ .

Next we consider the matrix element that comes from the *disconnected* diagram in Fig. 2(a), Eq. (28), which contains tastes  $I$ ,  $V$ , and  $A$  hairpin propagators. Because the 2-PGB disconnected vertices are dimension 2, an additional hairpin vertex can be canceled by an additional propagator in the  $S_\chi$ PT power counting, so one must

resum the flavor-neutral propagators with all possible numbers of sea quark loops between the external hairpins. The staggered, flavor-neutral, full propagators were determined in Ref. [8] using the general method outlined in Appendix A of Ref. [20], but adding factors of 1/4 for every sea quark loop in the series. We will just quote the results.

Generically, hairpin propagators contain poles at the mass eigenstates ( $\pi^0$ ,  $\eta$ , and  $\eta'$ ) and the flavor eigenstates ( $X$  and  $Y$ ).<sup>16</sup> These multiple poles can be rewritten as sums of single poles times their residues, e.g.,

$$D_{xy}^{V,\text{PQ}}(q) = -a^2 \delta'_V \frac{1}{(q^2 + m_{X_V}^2)(q^2 + m_{Y_V}^2)} \frac{(q^2 + m_{U_V}^2)(q^2 + m_{D_V}^2)(q^2 + m_{S_V}^2)}{(q^2 + m_{\pi^0_V}^2)(q^2 + m_{\eta_V}^2)(q^2 + m_{\eta'_V}^2)} \quad (42)$$

$$= -a^2 \delta'_V \sum_{j=X,Y,\pi^0,\eta,\eta'} \frac{1}{q^2 + m_{jV}^2} R_{jV}^{[5,3]}(\{m_X, m_Y, m_{\pi^0}, m_\eta, m_{\eta'}\}; \{m_U, m_D, m_S\}), \quad (43)$$

where  $R_{jV}^{[5,3]}$  is the residue of the single pole at  $q^2 = -m_{jV}^2$ , e.g.,

$$R_{X_V}^{[5,3]} = \frac{(m_{U_V}^2 - m_{X_V}^2)(m_{D_V}^2 - m_{X_V}^2)(m_{S_V}^2 - m_{X_V}^2)}{(m_{Y_V}^2 - m_{X_V}^2)(m_{\pi^0_V}^2 - m_{X_V}^2)(m_{\eta_V}^2 - m_{X_V}^2)(m_{\eta'_V}^2 - m_{X_V}^2)}. \quad (44)$$

We use the notation of Refs. [8,9], which is most clearly described in the Appendix of [9]. The expression in Eq. (42) is valid for three sea quarks with any combination of masses; an analogous expression holds for  $D_{xy}^{A,\text{PQ}}(q)$ . Because the *overall* flavor-taste singlet decouples from the PQ theory, when  $m_0 \rightarrow \infty$  the expression for  $D_{xy}^{I,\text{PQ}}(q)$  simplifies further:

<sup>15</sup>Note that we use dimensional regularization and the same renormalization scheme as Refs. [8,9].

<sup>16</sup>The staggered factors of 1/4 shift the mass eigenvalues but do not introduce additional poles. Expressions for the  $\pi^0$ ,  $\eta$ , and  $\eta'$  masses in a theory with 2 + 1 sea quarks are given in Ref. [8]. The generalization to the 1 + 1 + 1 case is straightforward.

$$D_{xy}^{I,\text{PQ}}(q) = -\frac{4m_0^2}{3} \frac{1}{(q^2 + m_{X_i}^2)(q^2 + m_{Y_i}^2)} \frac{(q^2 + m_{U_i}^2)(q^2 + m_{D_i}^2)(q^2 + m_{S_i}^2)}{(q^2 + m_{\pi_0}^2)(q^2 + m_{\eta_i}^2)(q^2 + m_{\eta'_i}^2)} \quad (45)$$

$$\xrightarrow{m_0 \rightarrow \infty} -\frac{4}{3} \frac{1}{(q^2 + m_{X_i}^2)(q^2 + m_{Y_i}^2)} \frac{(q^2 + m_{U_i}^2)(q^2 + m_{D_i}^2)(q^2 + m_{S_i}^2)}{(q^2 + m_{\pi_0}^2)(q^2 + m_{\eta_i}^2)} \quad (46)$$

$$= -\frac{4}{3} \sum_{j=X,Y,\pi^0,\eta} \frac{1}{q^2 + m_{j_i}^2} R_{j_i}^{[4,3]}(\{m_X, m_Y, m_{\pi^0}, m_{\eta}\}; \{m_U, m_D, m_S\}). \quad (47)$$

Some of the taste  $I$  disconnected propagators are multiplied by  $q^2$ , so the residues of their poles are changed:

$$q^2 D_{xy}^{I,\text{PQ}}(q) \sim \frac{4}{3} \sum_{j=X,Y,\pi^0,\eta} \frac{m_{j_i}^2}{q^2 + m_{j_i}^2} R_{j_i}^{[4,3]}(\{m_X, m_Y, m_{\pi^0}, m_{\eta}\}; \{m_U, m_D, m_S\}), \quad (48)$$

where “ $\sim$ ” indicates that this relationship only holds within the integral.

Using the definitions of the disconnected propagators it can be shown that

$$D_{xx} + D_{yy} - 2D_{xy} = (m_X^2 - m_Y^2)^2 \frac{\partial}{\partial m_X^2} \frac{\partial}{\partial m_Y^2} \{D_{xy}\}. \quad (49)$$

This greatly simplifies the 1-loop matrix element contribution to  $B_K$  from disconnected quark-level diagrams:

$$\begin{aligned} \mathcal{M}_{\text{disc}}^{\text{PQ}} &= \frac{2C_X^K}{3\pi^2 f^4} (m_{X_i}^2 - m_{Y_i}^2)^2 \frac{\partial}{\partial m_{X_i}^2} \frac{\partial}{\partial m_{Y_i}^2} \left\{ \sum_{j=X,Y,\pi^0,\eta} \ell(m_{j_i}^2) (m_{xy_p}^2 + m_{j_i}^2) R_{j_i}^{[4,3]} \right\} + \frac{(2C_X^{2V} + C_X^{3V}) a^2 \delta'_A}{\pi^2 f^4} (m_{X_A}^2 - m_{Y_A}^2)^2 \frac{\partial}{\partial m_{X_A}^2} \\ &\times \frac{\partial}{\partial m_{Y_A}^2} \left\{ \sum_{j=X,Y,\pi^0,\eta,\eta'} \ell(m_{j_A}^2) R_{j_A}^{[5,3]} \right\} + \frac{(2C_X^{2A} + C_X^{3A}) a^2 \delta'_V}{\pi^2 f^4} (m_{X_V}^2 - m_{Y_V}^2)^2 \frac{\partial}{\partial m_{X_V}^2} \frac{\partial}{\partial m_{Y_V}^2} \left\{ \sum_{j=X,Y,\pi^0,\eta,\eta'} \ell(m_{j_V}^2) R_{j_V}^{[5,3]} \right\}, \end{aligned} \quad (50)$$

where we have set  $p^2 = -m_{xy_p}^2$  and the arguments of the  $R$ 's are as in Eqs. (43) and (47). This compact notation emphasizes that the disconnected matrix element vanishes quadratically as  $(m_x - m_y) \rightarrow 0$ .

Finally,  $B_K$  at NLO in a PQ theory with 1 + 1 + 1 sea quarks is

$$\begin{aligned} B_K^{\text{PQ}} &= B_0 + \frac{3}{8} \left( \frac{\mathcal{M}_{\text{conn}} + \mathcal{M}_{\text{disc}}^{\text{PQ}}}{f_{xy_p}^2 m_{xy_p}^2} + \frac{A}{f_{xy_p}^2} + B \frac{m_{xy_p}^2}{f_{xy_p}^2} \right. \\ &\left. + C \frac{(m_x - m_y)^2}{f_{xy_p}^2 m_{xy_p}^2} + D \frac{(m_u + m_d + m_s)}{f_{xy_p}^2} \right), \end{aligned} \quad (51)$$

where  $B_0$  is given in Eq. (21) and  $\mathcal{M}_{\text{conn}}$  and  $\mathcal{M}_{\text{disc}}^{\text{PQ}}$  are given in Eqs. (41) and (50). Here we have reintroduced the analytic terms from Appendix C 3. We have checked explicitly that all renormalization scale dependence in the logarithms can be absorbed by the analytic terms.

## B. $B_K$ at NLO in the quenched theory

We now find  $B_K$  in the quenched theory. Because of the absence of sea quark loops, additional factors of 1/4 to reduce the number of tastes per flavor are unnecessary. Moreover, the flavor-neutral propagators need not be resummed, so their forms are quite simple, e.g.,

$$D_{xy}^{V,\text{Quench}}(q) = -a^2 \delta'_V \frac{1}{(q^2 + m_{X_V}^2)(q^2 + m_{Y_V}^2)}, \quad (52)$$

and likewise for  $D_{xy}^{A,\text{Quench}}(q)$ . However, the flavor-taste singlet ( $\eta'_i$ ) does not decouple in the quenched theory, so we cannot take  $m_0 \rightarrow \infty$ :

$$D_{xy}^{I,\text{Quench}}(q) = -\frac{4}{3} \frac{m_0^2 + \alpha q^2}{(q^2 + m_{X_i}^2)(q^2 + m_{Y_i}^2)}, \quad (53)$$

where  $\alpha$  is an additional chiral parameter that is only present in the quenched theory, not the strong coupling constant. As before, the nondegenerate propagators can be

further simplified into sums of poles times residues:

$$D_{xy}^{V,\text{Quench}}(q) = -a^2 \delta'_V \frac{1}{m_{Y_V}^2 - m_{X_V}^2} \left[ \frac{1}{q^2 + m_{X_V}^2} - \frac{1}{q^2 + m_{Y_V}^2} \right],$$

$$D_{xy}^{I,\text{Quench}}(q) = \left( \frac{-4}{3} \right) \frac{1}{m_{Y_I}^2 - m_{X_I}^2} \left[ \frac{m_0^2 - \alpha m_{X_I}^2}{(q^2 + m_{X_I}^2)} - \frac{m_0^2 - \alpha m_{Y_I}^2}{(q^2 + m_{Y_I}^2)} \right]. \quad (54)$$

Because the connected quark-level diagrams that contribute to  $B_K$  at 1 loop, Figs. 2(b) and 2(c), do not contain sea quark loops, the corresponding matrix elements remain unchanged from the PQ theory. The matrix elements from disconnected quark-level diagrams, after setting  $p^2 = -m_{xy_p}^2$ , are

$$\begin{aligned} \mathcal{M}_{\text{disc}}^{\text{Quench}} &= \frac{2C_X^K}{3\pi^2 f^4} (m_{X_I}^2 - m_{Y_I}^2)^2 \frac{\partial}{\partial m_{X_I}^2} \frac{\partial}{\partial m_{Y_I}^2} \left\{ \sum_{j=X,Y} \ell(m_{j_I}^2) (m_{xy_p}^2 + m_{j_I}^2) (m_0^2 - \alpha m_{j_I}^2) R_{j_I}^{[2,0]} \right\} \\ &+ \frac{(2C_X^{2V} + C_X^{3V}) a^2 \delta'_A}{\pi^2 f^4} (m_{X_A}^2 - m_{Y_A}^2)^2 \frac{\partial}{\partial m_{X_A}^2} \frac{\partial}{\partial m_{Y_A}^2} \left\{ \sum_{j=X,Y} \ell(m_{j_A}^2) R_{j_A}^{[2,0]} \right\} \\ &+ \frac{(2C_X^{2A} + C_X^{3A}) a^2 \delta'_V}{\pi^2 f^4} (m_{X_V}^2 - m_{Y_V}^2)^2 \frac{\partial}{\partial m_{X_V}^2} \frac{\partial}{\partial m_{Y_V}^2} \left\{ \sum_{j=X,Y} \ell(m_{j_V}^2) R_{j_V}^{[2,0]} \right\}, \end{aligned} \quad (55)$$

where

$$R_{X_V}^{[2,0]} \equiv R_{X_V}^{[2,0]}(\{m_X, m_Y\}) = \frac{1}{(m_{Y_V}^2 - m_{X_V}^2)} = -R_{Y_V}^{[2,0]}. \quad (56)$$

Therefore  $B_K$  at NLO in the quenched theory is

$$\begin{aligned} B_K^{\text{Quench}} &= B_0 + \frac{3}{8} \left( \frac{\mathcal{M}_{\text{conn}} + \mathcal{M}_{\text{disc}}^{\text{Quench}}}{f_{xy_p}^2 m_{xy_p}^2} + \frac{A}{f_{xy_p}^2} + B \frac{m_{xy_p}^2}{f_{xy_p}^2} \right. \\ &\left. + C \frac{(m_x - m_y)^2}{f_{xy_p}^2 m_{xy_p}^2} \right), \end{aligned} \quad (57)$$

where  $B_0$  is given in Eq. (21) and  $\mathcal{M}_{\text{conn}}$  is given in Eq. (41). We emphasize that the undetermined coefficients,  $C_X^i$  and  $A-C$ , are *different* in the quenched theory than in QCD, and that there is no  $D$  term.

### C. $B_K$ at NLO in the full (2 + 1) theory

Finally we determine  $B_K$  in full QCD, i.e.  $m_x = m_d$  and  $m_y = m_s$ , with 2 + 1 sea quarks ( $m_u = m_d \neq m_s$ ). Although one can, in principle, just take the limit of the final PQ expression, Eq. (51), it is easier to start from the expressions for the 1-loop matrix elements in terms of integrals.

Because QCD is a physical theory, the disconnected propagators only have single poles. It is easy to show this explicitly for the 2 + 1 theory, in which  $m_{\pi^0}^2 = m_U^2 = m_D^2$  and parts of the numerator and denominator cancel. For example, the taste  $V$  (and  $A$ ) disconnected propagators are

$$\begin{aligned} D_{dd}^{V,2+1}(q) &= -a^2 \delta'_V \left[ \frac{1}{q^2 + m_{D_V}^2} \right. \\ &\times \left( \frac{(m_{S_V}^2 - m_{D_V}^2)}{(m_{\eta_V}^2 - m_{D_V}^2)(m_{\eta'_V}^2 - m_{D_V}^2)} \right) \\ &\left. + (D \rightarrow \eta \rightarrow \eta' \rightarrow D) \right], \end{aligned} \quad (58)$$

$$\begin{aligned} D_{ss}^{V,2+1}(q) &= -a^2 \delta'_V \left[ \frac{1}{q^2 + m_{S_V}^2} \right. \\ &\times \left( \frac{(m_{D_V}^2 - m_{S_V}^2)}{(m_{\eta_V}^2 - m_{S_V}^2)(m_{\eta'_V}^2 - m_{S_V}^2)} \right) \\ &\left. + (S \rightarrow \eta \rightarrow \eta' \rightarrow S) \right], \end{aligned} \quad (59)$$

$$\begin{aligned} D_{ds}^{V,2+1}(q) &= -a^2 \delta'_V \left[ \frac{1}{q^2 + m_{\eta_V}^2} \frac{1}{(m_{\eta'_V}^2 - m_{\eta_V}^2)} \right. \\ &\left. + (\eta \leftrightarrow \eta') \right], \end{aligned} \quad (60)$$

where the arrows indicate all possible permutations of masses. As in the PQ theory, the taste singlet  $\eta'$  decouples, so we can use the explicit large- $m_0$  expressions for the masses,

$$m_{\pi^0}^2 = m_{U_I}^2 = m_{D_I}^2, \quad (61)$$

$$m_{\eta_I}^2 = \frac{m_{D_I}^2}{3} + \frac{2m_{S_I}^2}{3}, \quad (62)$$

$$m_{\eta'_I}^2 = m_0^2, \quad (63)$$

to drastically simplify the taste  $I$  disconnected propagators:

$$D_{dd}^{I,2+1}(q) \xrightarrow{m_0 \rightarrow \infty} -2 \frac{1}{q^2 + m_{D_I}^2} + \frac{2}{3} \frac{1}{q^2 + m_{\eta_I}^2}, \quad (64)$$

$$D_{ss}^{I,2+1}(q) \xrightarrow{m_0 \rightarrow \infty} -4 \frac{1}{q^2 + m_{S_I}^2} + \frac{8}{3} \frac{1}{q^2 + m_{\eta_I}^2}, \quad (65)$$

$$D_{ds}^{I,2+1}(q) \xrightarrow{m_0 \rightarrow \infty} -\frac{4}{3} \frac{1}{q^2 + m_{\eta_I}^2}. \quad (66)$$

In this section we only show the 1-loop matrix elements that come from disconnected quark-level diagrams:

$$\begin{aligned} \mathcal{M}_{\text{disc}}^{2+1} = & \frac{C_\chi^K}{\pi^2 f^4} \{ (m_{ds_p}^2 + m_{D_I}^2) \ell(m_{D_I}^2) + 2(m_{ds_p}^2 + m_{S_I}^2) \ell(m_{S_I}^2) - 3(m_{ds_p}^2 + m_{\eta_I}^2) \ell(m_{\eta_I}^2) \} \\ & - \frac{(2C_\chi^{2V} + C_\chi^{3V}) a^2 \delta'_A}{\pi^2 f^4} \left\{ 2 \left[ \ell(m_{\eta_A}^2) \frac{1}{(m_{\eta'_A}^2 - m_{\eta_A}^2)} + (\eta \leftrightarrow \eta') \right] \right. \\ & - \left[ \ell(m_{D_A}^2) \left( \frac{(m_{S_A}^2 - m_{D_A}^2)}{(m_{\eta_A}^2 - m_{D_A}^2)(m_{\eta'_A}^2 - m_{D_A}^2)} \right) + (D \rightarrow \eta \rightarrow \eta' \rightarrow D) \right] \\ & \left. - \left[ \ell(m_{S_A}^2) \left( \frac{(m_{D_A}^2 - m_{S_A}^2)}{(m_{\eta_A}^2 - m_{S_A}^2)(m_{\eta'_A}^2 - m_{S_A}^2)} \right) + (S \rightarrow \eta \rightarrow \eta' \rightarrow S) \right] \right\} \\ & - \frac{(2C_\chi^{2A} + C_\chi^{3A}) a^2 \delta'_V}{\pi^2 f^4} \left\{ 2 \left[ \ell(m_{\eta_V}^2) \frac{1}{(m_{\eta'_V}^2 - m_{\eta_V}^2)} + (\eta \leftrightarrow \eta') \right] \right. \\ & - \left[ \ell(m_{D_V}^2) \left( \frac{(m_{S_V}^2 - m_{D_V}^2)}{(m_{\eta_V}^2 - m_{D_V}^2)(m_{\eta'_V}^2 - m_{D_V}^2)} \right) + (D \rightarrow \eta \rightarrow \eta' \rightarrow D) \right] \\ & \left. - \left[ \ell(m_{S_V}^2) \left( \frac{(m_{D_V}^2 - m_{S_V}^2)}{(m_{\eta_V}^2 - m_{S_V}^2)(m_{\eta'_V}^2 - m_{S_V}^2)} \right) + (S \rightarrow \eta \rightarrow \eta' \rightarrow S) \right] \right\}. \quad (67) \end{aligned}$$

The matrix elements which come from connected diagrams are *minimally* changed from the PQ ones in that  $x \rightarrow d$  and  $y \rightarrow s$ . Therefore  $B_K$  at NLO in full QCD with 2 + 1 sea quarks is

$$\begin{aligned} B_K^{2+1} = & B_0 + \frac{3}{8} \left( \frac{\mathcal{M}_{\text{conn}}(x \rightarrow d, y \rightarrow s) + \mathcal{M}_{\text{disc}}^{2+1}}{f_{ds_p}^2 m_{ds_p}^2} \right) \\ & + \frac{3}{8} \left( \frac{A}{f_{ds_p}^2} + B \frac{m_{ds_p}^2}{f_{ds_p}^2} + C \frac{(m_d - m_s)^2}{f_{ds_p}^2 m_{ds_p}^2} \right. \\ & \left. + D \frac{(2m_d + m_s)}{f_{ds_p}^2} \right), \quad (68) \end{aligned}$$

where  $B_0$  is given in Eq. (21) and  $\mathcal{M}_{\text{conn}}$  is given in Eq. (41).

## VI. GUIDE FOR LATTICE DETERMINATION OF $B_K$

Any lattice determination of  $B_K$  necessarily entails two steps: first, fit the  $\chi$ PT expression to the lattice data and determine the unknown coefficients, and, second, use the

resulting function to extrapolate  $B_K$  to its value at the physical quark masses in the continuum limit. The large number of undetermined constants in the staggered  $\chi$ PT expression for  $B_K$  at NLO could make the fitting procedure quite difficult, and certainly necessitates a lot of data. Our goal is therefore to find alternative ways to access some of these coefficients. Once measured, they can be used in the full NLO S $\chi$ PT  $B_K$  fit.

We begin by reviewing explicitly all of the parameters in the expression for  $B_K$  at NLO. We presume that  $B_K$  will be fit as a function of the kaon mass squared, so we do not consider the latter to be a fit parameter. In order to properly count the number of undetermined coefficients, we must combine the PQ expression for  $B_K$ , Eq. (51), with Tables IV and V. Only 13 of the 15 operators listed in Eq. (27) in fact appear in Eq. (51). Operators  $\mathcal{O}_\chi^{3P}$  and  $\mathcal{O}_\chi^{3T}$  turn out not to contribute to  $B_K$  at NLO. Operators  $\mathcal{O}_\chi^{2V}$  and  $\mathcal{O}_\chi^{3V}$  (and similarly  $\mathcal{O}_\chi^{2A}$  and  $\mathcal{O}_\chi^{3A}$ ) turn out to enter the expression for  $B_K$  as a single linear combination, and can therefore be associated with a single coefficient. Thus we can see from Table IV that there are 23 coefficients associated with NLO chiral operators, in addition to  $C_\chi^K$ .

Similarly, Table V shows that there are six linearly independent analytic terms. Finally, certain parameters appear in all  $S\chi$ PT calculations: the four tree-level mass splittings, the two hairpin parameters, and  $f_{xy_p}$ . Thus the expression for  $B_K$  at NLO is given in terms of 37 undetermined coefficients. Fortunately, some of these parameters have already been determined on the MILC 2 + 1 dynamical field configurations [7]. In particular, the taste Goldstone meson decay constant, as well as the mass splittings between the other PGB tastes, are already known.<sup>17</sup> However, this still leaves 32 *undetermined fit parameters*.

An obvious way to reduce the number of parameters is to perform a fit to a single lattice spacing. This reduces the coefficients in Table IV to one per operator (or per linear combination), i.e. from 23 to 9, and the analytic terms in Table V from 6 to 4. Clearly one must eventually fit to multiple lattice spacings. We suggest, however, that is preferable to fit  $B_K$  in a multistage process. First, perform separate NLO  $B_K$  fits at multiple lattice spacings with this reduced set of parameters. As we will show, a single lattice spacing fit of  $B_K$ , in combination with tree-level fits of other matrix elements, can be used to separate *almost* all of the perturbative and discretization errors from the desired continuum result. Second, remove these errors at each lattice spacing. Third, fit the results after error subtraction simultaneously to multiple lattice spacings using the appropriately simplified  $S\chi$ PT expression. Last, send the remaining errors, which have now been obtained, to zero, and calculate a value for the continuum parameter  $\hat{B}_K$ .<sup>18</sup>

We now discuss step one: fitting the NLO expression for  $B_K$  at a single lattice spacing. Because this still involves 16 parameters, it is qualitatively useful to first consider the case of degenerate valence quarks, i.e.  $m_x = m_y$ . Recall from the previous sections that the NLO expression for  $B_K$  greatly simplifies in this limit. The matrix element from the disconnected diagram vanishes, thereby removing coefficients  $C_\chi^{2V} - C_\chi^{3A}$  and the hairpin parameters, as do half of the terms in the matrix element from the connected diagrams. The degenerate mass case is therefore the appropriate testing ground for whether the NLO expression describes the data at all, i.e. if one is in the chiral regime, before moving on to the more complex, but physical, situation of  $m_x \neq m_y$ .

In fact, the degenerate mass case is also quantitatively useful. Recall that the dominant contribution to  $B_K$  at NLO comes from the tree-level matrix element of  $\mathcal{O}_\chi^K$ , and is proportional to  $C_\chi^K$ :

$$\frac{\langle \bar{K}_{1P}^0 | \mathcal{O}_{1\text{-loop}}(\text{taste } P) | K_{2P}^0 \rangle}{\mathcal{M}_{\text{vac}}} = 6 \frac{C_\chi^K}{f^4} + \dots \quad (69)$$

Because we wish to determine  $B_K$  (and consequently  $C_\chi^K$ ) to NLO accuracy, the above expression is insufficient.<sup>19</sup> An NLO fit to  $B_K$  with degenerate valence quark masses, however, can potentially be used to determine  $C_\chi^K$  to NLO accuracy. In this limit,  $B_K$  at NLO has the following form:

$$\begin{aligned} B_K = & 12 \frac{C_\chi^K}{f^4} \left\{ 1 + \frac{1}{512\pi^2 f_{xy_p}^2} \right. \\ & \times \sum_{B'} f^{B'} \left( \ell(m_{\bar{K}_{B'}}^2) - \frac{1}{2} m_{\bar{K}_{B'}}^2 \tilde{\ell}(m_{\bar{K}_{B'}}^2) \right) \Big\} \\ & - \frac{3}{256\pi^2 f_{xy_p}^2} \sum_{B'} \tilde{\ell}(m_{\bar{K}_{B'}}^2) \sum_B \left( \frac{C_\chi^{1B}}{f^4} g^{BB'} - \frac{C_\chi^{2B}}{f^4} h^{BB'} \right) \\ & + A \frac{3}{8f_{xy_p}^2} + B \frac{3m_{xy_p}^2}{8f_{xy_p}^2} + D \frac{3(m_u + m_d + m_s)}{8f_{xy_p}^2}. \quad (70) \end{aligned}$$

It depends on  $C_\chi^K$ , now at NLO, three analytic terms, and five linear combinations of other operator coefficients,  $\sum_B (C_\chi^{1B} g^{BB'} - C_\chi^{2B} h^{BB'})$  for  $B' = I, P, V, A, T$ . As we show below, these five combinations can be determined to the required accuracy using other matrix elements. Once they are found, we would like to use a fit of Eq. (70) to extract  $C_\chi^K$  to NLO accuracy. Unfortunately, this turns out not to be possible; separation of the NLO part of  $C_\chi^K$  from the analytic term  $A$  requires a simultaneous fit to multiple lattice spacings. We can, however, consider their sum,  $(C_\chi^K/f^4 + 3A/16f_{xy_p}^2)$ , to be a single fit parameter that *can* be determined to NLO accuracy. We can also determine the remaining analytic terms,  $B$  and  $D$ , using Eq. (70) at a single lattice spacing.

The other simple handle that one has in lattice simulations, besides the quark masses, is the flavors and tastes of the external states. At tree level, double-supertrace operators have nonvanishing matrix elements between a  $\bar{K}_1^0$  ( $s1\bar{d}1$ ) and a  $K_2^0$  ( $\bar{s}2d2$ ). In contrast, single-supertrace operators have nonvanishing matrix elements between “mixed” kaons:  $s1\bar{d}2$  and  $\bar{s}2d1$ . Both single- and double-supertrace operators can have nonvanishing matrix elements between all five possible tastes:  $I, P, V, A$ , and  $T$ . Thus, by evaluating the matrix element of  $\mathcal{O}_{1\text{-loop}}(\text{taste } P)$  between different external states, we can naively hope to determine as many as nine independent coefficients for use in the larger NLO  $S\chi$ PT fit. In essence, while the two sets of valence quarks and multiple tastes introduce many new

<sup>17</sup>We note that the vector hairpin is poorly determined from the MILC data, and that effects of the vector and axial-vector hairpins are, in practice, difficult to disentangle in  $S\chi$ PT fits.

<sup>18</sup>The explicit expression for  $\hat{B}_K$  in terms of  $B_K$  is given in Eq. (82).

<sup>19</sup>In fact, at first glance, we also need  $f$  at NLO. However, this is really unnecessary because  $C_\chi^K$  enters both the LO and NLO contributions to  $B_K$  as the combination  $C_\chi^K/f^4$ . Thus we can treat the quantity  $(C_\chi^K/f^4)$  as the undetermined fit parameter, and likewise for the other NLO coefficients.

operators into staggered  $\chi$ PT, they also allow the independent determination of more operator coefficients.

First consider the double-supertrace operators, of which there are four:  $\mathcal{O}_\chi^{2V}$ ,  $\mathcal{O}_\chi^{3V}$ ,  $\mathcal{O}_\chi^{2A}$ , and  $\mathcal{O}_\chi^{3A}$ . Clearly, to extract the coefficient of  $\mathcal{O}_\chi^{2V}$  (or  $\mathcal{O}_\chi^{2A}$ ) we need to choose taste vector (or axial) kaons for our external states. The resulting matrix elements are

$$\langle \overline{K}_{1V}^0 | \mathcal{O}_{1\text{-loop}}(\text{taste } P) | K_{2V}^0 \rangle = \frac{-16}{f^2} (2C_\chi^{2V} - C_\chi^{3V}), \quad (71)$$

$$\langle \overline{K}_{1A}^0 | \mathcal{O}_{1\text{-loop}}(\text{taste } P) | K_{2A}^0 \rangle = \frac{-16}{f^2} (2C_\chi^{2A} - C_\chi^{3A}), \quad (72)$$

so we can only determine the linear combination of coefficients  $(2C_\chi^{2V} - C_\chi^{3V})$ . However, in the 1-loop contributions to  $B_K$ , these coefficients enter in a different linear combination:  $(2C_\chi^{2V} + C_\chi^{3V})$ . Thus we cannot in any simple way separately extract the coefficients of the two supertrace operators. Nevertheless, calculation of these matrix elements should still be done, as it provides an estimate of the size of the NLO coefficients. Moreover, because the matrix element of  $\mathcal{O}_{1\text{-loop}}(\text{taste } P)$  between taste  $I$  and  $T$  kaons should be zero at this order,

$$\langle \overline{K}_{1I}^0 | \mathcal{O}_{1\text{-loop}}(\text{taste } P) | K_{2T}^0 \rangle = 0(+\text{NNLO}), \quad (73)$$

$$\langle \overline{K}_{1T}^0 | \mathcal{O}_{1\text{-loop}}(\text{taste } P) | K_{2I}^0 \rangle = 0(+\text{NNLO}), \quad (74)$$

they can be used to gauge the size of the higher-order terms, thereby testing the validity of our phenomenological power-counting scheme and determining whether or not it is necessary, as is done in Ref. [7], to introduce a small subset of next-to-next-to leading order (NNLO) terms.

Next consider the single-supertrace operators. Extracting many of their coefficients turns out to be feasible, although not completely trivial. There are eight such operators— $\mathcal{O}_\chi^N$ ,  $\mathcal{O}_\chi^{1P}$ ,  $\mathcal{O}_\chi^{2P}$ ,  $\mathcal{O}_\chi^{1T}$ ,  $\mathcal{O}_\chi^{1I}$ ,  $\mathcal{O}_\chi^{2T}$ ,  $\mathcal{O}_\chi^{1V}$ , and  $\mathcal{O}_\chi^{1A}$ —and five possible matrix elements. We can combine them into a single equation using the same coefficients as in Eq. (70):

$$\begin{aligned} & \frac{\langle \overline{K}_{1B'}^{\text{mix}} | \mathcal{O}_{1\text{-loop}}(\text{taste } P) | K_{2B'}^{\text{mix}} \rangle}{\mathcal{M}_{\text{vac}}} \\ &= \frac{3}{8} \frac{1}{N^{B'} f^4 m_{K_p}^2} \left[ C_\chi^K f^{B'} m_{K_{B'}}^2 + \sum_B (C_\chi^{1B} g^{BB'} - C^{2B} h^{BB'}) \right], \end{aligned} \quad (75)$$

where

$$N^{B'} = (-1 \quad -1 \quad 4 \quad 4 \quad -6). \quad (76)$$

The coefficient  $C_\chi^K$  can be separated through its quark mass dependence, but this is not useful because it is only determined at LO in this tree-level matrix element. However, the numerical values of the five linear combinations of coefficients corresponding to  $B' = I, P, V, A, T$  are needed,

along with the degenerate mass  $B_K$  result, Eq. (70), to determine the combination of coefficients  $(12C_\chi^K/f^4 + 3A/8f_{xy}^2)$  to NLO accuracy. The five matrix elements in Eq. (75) also aid in the true,  $m_x \neq m_y$ ,  $B_K$  determination because we can solve for five coefficients in terms of the remaining two, and substitute the resulting functions into the NLO S $\chi$ PT expression before fitting.

Thus, after exploiting both degenerate mass and mixed-flavor matrix elements, we are left with a NLO fit at a single lattice spacing to only *seven parameters*:  $\delta'_V$ ,  $\delta'_A$ ,  $(2C_\chi^{2V} + C_\chi^{3V})$ ,  $(2C_\chi^{2A} + C_\chi^{3A})$ ,  $C$ , and two additional NLO coefficients of our choice from Eq. (75). This should certainly be feasible. Once this is done, the next step is to remove as much of the discretization and perturbative error as possible *before* combining the data from multiple lattice spacings. We do this by setting all coefficients that are multiplied by  $a^2$ ,  $a_\alpha^2$ ,  $\alpha/4\pi$ , and  $a^2$  to zero.<sup>20</sup> One can see now why it is preferable to fit to one lattice spacing at a time—it avoids having to give each operator in Table IV a separate coefficient for each expansion parameter. Doing so would clearly be a waste of effort, since one only needs the total error, not the individual coefficients, in order to remove it. We emphasize, however, that there is one coefficient that cannot be removed at this stage— $A$ . It is tied up in the linear combination  $(12C_\chi^K/f^4 + 3A/8f_{xy}^2)$ . We must therefore remove all of the coefficients of operators in Table IV and then perform a new fit to multiple lattice spacings.<sup>21</sup> Such a fit will only contain the continuum parameter  $C_\chi^K$  and  $A$ , the latter of which must now be separated into three pieces of  $\mathcal{O}(a^2)$ ,  $\mathcal{O}(a_\alpha^2)$ , and  $\mathcal{O}(a^2)$ . It therefore has *four parameters*, which is a clear improvement over the 37 parameters that would have originally been needed in a fit to multiple lattice spacings. Once we have separated the three contributions to  $A$  from the coefficient  $C_\chi^K$  at NLO, we will remove them as well.

Finally, assuming that we have successfully carried out the fitting procedure outlined above, we are ready to turn the result into a continuum value for  $\hat{B}_K$ . We therefore review the procedure for matching lattice and continuum operators. Along the way, we clarify some of the points and justify some of the assumptions made previously.

After removal of discretization errors, multiple lattice regularized operators match onto a given continuum operator in the following manner:

$$\mathcal{O}^{\text{cont}}(\mu) = Z(\mu, a) \mathcal{O}^{\text{lat}}(a), \quad (77)$$

where  $\mathcal{O}^{\text{cont}}(\mu)$  is a single operator,  $\mathcal{O}^{\text{lat}}(a)$  is a vector of

<sup>20</sup>This is obviously the correct way to remove discretization errors. The right way to remove perturbative errors, however, is less apparent, since we have only *partially* matched the continuum operator at 1 loop. We will later show that it is, in fact, correct to simply set their coefficients to zero.

<sup>21</sup>Strictly speaking we should fit to the quantity  $\hat{B}_K$ , which is independent of lattice spacing, as we explain below.

operators, and the vector  $Z$  relates the two. The 1-loop expression for  $Z(\mu, a)$  is typically written in the following way:

$$Z(\mu, a) = 1 + \frac{\alpha}{4\pi} c(\mu a), \quad (78)$$

where  $c$  is a vector that contains the 1-loop matching coefficients. However, Eq. (78) clearly has a problem: what is the appropriate scale at which to evaluate  $\alpha$ ? This ambiguity can be avoided by using the exact perturbative formula for  $Z$  given by Ji in Ref [21]:

$$Z(\mu, a) = \exp\left(-\int_0^{g(\mu)} dg'' \frac{\gamma_{\text{cont}}(g'')}{\beta_{\text{cont}}(g'')} T_{g''}\right) \times \exp\left(-\int_{g(a)}^0 dg' \frac{\gamma_{\text{lat}}(g')}{\beta_{\text{lat}}(g')}\right), \quad (79)$$

where  $T_{g'}$  indicates a  $g'$ -ordered product<sup>22</sup> and  $\gamma$  and  $\beta$  are the anomalous dimension and beta function, respectively.<sup>23</sup> Equation (79) is easy to understand physically; it shows that matching is a three-step process. First, run the lattice operators calculated at a scale  $q^* = 1/a$  up to  $a = 0$ . Next, match the lattice operators at this scale onto the continuum operator at a scale  $\mu = \infty$ . Finally, run the continuum operator down to a scale  $\mu$ , which is typically 2 GeV. This procedure cleanly separates the lattice scale,  $a$ , from the continuum scale,  $\mu$ .

Mathematically, Eq. (79) can be used to derive the following relationship between the continuum and lattice operators:

$$\begin{aligned} \alpha(\mu)^{-\gamma_0/2\beta_0} \left[ 1 + \frac{\alpha(\mu)}{4\pi} J_{\text{cont}} + \mathcal{O}(\alpha^2) \right] \mathcal{O}^{\text{cont}}(\mu) \\ = \alpha(q^*)^{-\gamma_0/2\beta_0} \left[ 1 + \frac{\alpha(q^*)}{4\pi} J_{\text{cont}} + \mathcal{O}(\alpha^2) \right] \\ \times \left[ 1 + \frac{\alpha(q^*)}{4\pi} c(q^*a) + \mathcal{O}(\alpha^2) \right] \mathcal{O}^{\text{lat}}(a), \end{aligned} \quad (80)$$

where

$$J = \frac{\beta_1 \gamma_0}{2\beta_0^2} - \frac{\gamma_1}{2\beta_0}, \quad (81)$$

and the vector  $c$  contains the standard 1-loop matching coefficients. This formulation of perturbative matching is extremely useful. Because the matching coefficients only appear on the ‘‘lattice side’’ of the equation, and no longer depend on the continuum scale  $\mu$ , one can choose any reasonable scale at which to calculate the matching coefficients—a natural choice is  $q^* = 1/a$ . One must, of course, then evaluate all  $\alpha$ ’s in the NLO S $\chi$ PT expression for  $B_K$  that arise due to truncation of perturbative matching factors at this same scale.

<sup>22</sup>Note that there is no  $T_{g''}$  because we are considering the case in which there is only one continuum operator and no mixing on the continuum side.

<sup>23</sup>Note also that the coupling  $g(a)$  can be chosen to be a continuumlike coupling such as  $\overline{\text{MS}}$  [22].

Equation (80) also justifies our treatment of errors associated with perturbative matching. Recall that the standard matching procedure is to *only match to taste P operators* at 1 loop. In terms of the above matching expression, this means that one calculates the matrix element of  $[1 + c(\alpha/4\pi)]\mathcal{O}_P^{\text{lat}}$ , where the vector  $\mathcal{O}_P^{\text{lat}}$  contains only taste  $P$  operators. However, this is not correct, as lattice operators of *all tastes* are required to correctly match the continuum operator. We account for this error by including ‘‘wrong-taste’’ operators in the S $\chi$ PT fit with coefficients of  $\mathcal{O}(\alpha/4\pi)$ . Conversely, if we *were* to correctly match onto the continuum  $B_K$  operator at 1 loop, we would *not need* to include such operators. Thus we conclude that in order to extract the correct result for  $B_K$  in the continuum we must set the coefficients of  $\mathcal{O}(\alpha/4\pi)$  in Table IV to zero. One can use the same logic to show that *all* of the perturbative matching errors should be dealt with in this manner, i.e. we must also set the  $\mathcal{O}(\alpha^2)$  coefficients in Tables IV and V to zero. This result is fortunate, as it allows treatment of all errors—both from discretization and perturbative matching—in the same, simple manner.

Finally, Eq. (80) shows that, once one has successfully fit the S $\chi$ PT expression for  $B_K$  and removed the discretization and perturbative matching errors at each lattice spacing, constructing the renormalization group invariant quantity,  $\hat{B}_K$  is trivial. This is because, to 1-loop order [23],

$$\hat{B}_K = \alpha(\mu)^{-\gamma_0/2\beta_0} \left[ 1 + \frac{\alpha(\mu)}{4\pi} J_{\text{cont}} \right] B_K^{\overline{\text{MS}}}(\mu), \quad (82)$$

so

$$\hat{B}_K = \alpha(q^*)^{-\gamma_0/2\beta_0} \left[ 1 + \frac{\alpha(q^*)}{4\pi} J_{\text{cont}} \right] B_K^{\text{lat,subtracted}}, \quad (83)$$

where  $B_K^{\text{lat,subtracted}}$  is the lattice value for  $B_K$  at the scale  $q^*$  with discretization and perturbative matching errors subtracted. Conversely, we can use Eq. (83) to remove the remaining errors from  $\hat{B}_K$  in the following way. Once we have values for  $B_K$  with all errors but those from  $A$  removed at each lattice spacing, we turn them into values for  $\hat{B}_K$  by multiplying by  $\alpha(q^*)^{-\gamma_0/2\beta_0} [1 + J_{\text{cont}}\alpha(q^*)/4\pi]$ . We then simultaneously fit  $\hat{B}_K$  at multiple lattice spacings to separate the continuum parameter  $C_\chi^K$  from  $A$ . Next, we subtract  $A$  to get a continuum value for  $\hat{B}_K$ . Last, we extrapolate the continuum result to the physical quark masses using the continuum  $\chi$ PT expression.

Even with the implementation of all of these suggestions, fitting the staggered  $B_K$  data will be highly non-trivial. One would clearly like to somehow decrease the number of operators, or at least undetermined coefficients, that contribute to  $B_K$  at NLO. One way to do this is by using improved links. This drastically reduces the 1-loop perturbative mixing between  $\mathcal{O}_{1\text{-loop}}$ (taste  $P$ ) and certain wrong-taste operators. In particular, for the case of hypercubic fat (HYP) links, the  $\mathcal{O}(\alpha/4\pi)$  mixing with taste  $S$  and taste  $T$  operators is so small that we can consider it to

be of higher order in our power counting [14]. This leaves the operators  $\mathcal{O}_\chi^{1I}$  and  $\mathcal{O}_\chi^{1T}$  each with only a single coefficient of  $\mathcal{O}(\alpha^2)$ , thereby reducing the total number of fit parameters by two.<sup>24</sup>

The biggest source of potential improvement, however, is in better perturbative matching. A good first step would therefore be to match to *all tastes of lattice operators* at 1 loop, since the requisite matching coefficients are known. While it has been thought that wrong-taste operator contributions to  $B_K$  would be less important than those from taste  $P$  operators, our power counting and operator enumeration show that this is not the case, in general. Complete 1-loop matching would eliminate the six coefficients of  $\mathcal{O}(\alpha/4\pi)$  in the first column Table IV. Moreover, 2-loop matching of all tastes, or better yet fully nonperturbative matching, would completely remove the first column in Table IV and eliminate the operators  $\mathcal{O}_\chi^{1I}$  and  $\mathcal{O}_\chi^{1T}$  altogether at this order, reducing the total number of fit parameters from 37 to 24.<sup>25</sup>

## VII. CONCLUSIONS

We have calculated  $B_K$  to NLO in  $S\chi$ PT for quenched, PQ, and full QCD, the form of which is necessary for correct continuum and chiral extrapolation of staggered lattice data. These results apply to both infinite and finite spatial volume. The most general expression is that for a PQ theory with three sea quarks in which all valence and sea masses are different.

This is the first calculation of a hadronic weak matrix element that is not a conserved current in staggered chiral perturbation theory. It illustrates the generic drawback of staggered fermions—large numbers of operators, both on the lattice and in the chiral effective theory, due to the additional taste degree of freedom. For example, while  $B_K$  at 1 loop in continuum chiral perturbation theory is proportional to a single operator coefficient, the same quantity in staggered  $\chi$ PT comes from 13 operators, many of which have more than one distinguishable coefficient. To address this concern we show explicitly how to determine six of these coefficients (at a single lattice spacing) from other simple matrix elements of the same lattice operator, thereby reducing the number of undetermined fit parameters. Furthermore, we suggest two ways to reduce the large number of coefficients—use of improved links, e.g. HYP or Asqtad, and of nonperturbative matching coefficients.

<sup>24</sup>We presume that a similar simplification occurs for Asqtad links, although their perturbative matching coefficients have yet to be determined.

<sup>25</sup>Recall that, generically, use of the matching procedure outlined in Eq. (79) to a given order in perturbation theory requires knowledge of the anomalous dimension at one higher order. For example, 1-loop matching requires the 2-loop anomalous dimension. Thus, while 2-loop matching would greatly reduce the number of coefficients in the NLO  $S\chi$ PT fit, the 3-loop anomalous dimension must also be calculated.

As usual, partial quenching helps to provide more handles in the fits, and will likely be essential here.

Use of our expression for  $B_K$  at NLO in  $S\chi$ PT, in combination with sufficient lattice data, should allow a precise determination of  $B_K$  with staggered quarks.

## ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy under Grant No. DE-FG02-96ER40956.

## APPENDIX A: $B_K$ IN CONTINUUM PQ $\chi$ PT WITH 2 + 1 SEA QUARKS

The continuum partially quenched result for  $B_K$  for  $N_{\text{sea}}$  degenerate sea quarks has been given by Golterman and Leung [24], and checked for  $N_{\text{sea}} = 2$  by Becirevic and Villadoro [25]. The result for “2 + 1” flavors of sea quarks does not, however, exist in the literature. It can be obtained by taking the appropriate limit of our staggered PQ expression, Eq. (51), and here we present the result.

To go from Eq. (51) to the continuum theory, we first set all coefficients associated with lattice errors, i.e. those multiplied by factors of  $a^2$ ,  $a_\alpha^2$ ,  $\alpha/4\pi$ , and  $\alpha^2$ , equal to zero. This eliminates all operator coefficients but  $C_\chi^K$ . It also removes the splittings among the 16 tastes of PGBs; we are therefore free to remove the taste subscripts. Next we take the ratio of  $B_K$  to  $B_0$  so that all factors of 4 due to traces in taste space drop out. Finally we set  $m_u = m_d \neq m_s$ . The next-to-leading-order result for  $B_K$  is then

$$\left(\frac{B_K}{B_0}\right)^{\text{PQ},2+1} = 1 + \frac{1}{48\pi^2 f^2 m_{xy}^2} [I_{\text{conn}} + I_{\text{disc}} + b m_{xy}^4 + c(m_X^2 - m_Y^2)^2 + d m_{xy}^2 (2m_D^2 + m_S^2)], \quad (\text{A1})$$

where  $f \approx 130$  MeV. The coefficients  $b$ ,  $c$ , and  $d$  are unknown dimensionless low-energy constants proportional to the constants  $B$ ,  $C$ , and  $D$  in Eq. (51). Note that the  $d$  term is proportional to the sum of the sea quark masses, and, to this order, effectively renormalizes  $B_0$  in a fit at a given sea quark mass. Note also that  $c$  is absent for degenerate valence quarks,  $m_x = m_y$ .

The chiral logarithms are given by  $I_{\text{conn}}$  and  $I_{\text{disc}}$ . We have broken them into two contributions since the latter vanishes when  $m_x = m_y$  and is also the only place where sea quark masses enter. The result from the quark connected diagrams is

$$I_{\text{conn}} = 6m_{xy}^4 \tilde{\ell}(m_{xy}^2) - 3\ell(m_X^2)(m_{xy}^2 + m_X^2) - 3\ell(m_Y^2)(m_{xy}^2 + m_Y^2), \quad (\text{A2})$$

while that for the diagrams involving a hairpin vertex is

$$I_{\text{disc}} = I_X + I_Y + I_\eta, \quad (\text{A3})$$

$$I_X = \tilde{\ell}(m_X^2) \frac{(m_{xy}^2 + m_X^2)(m_D^2 - m_X^2)(m_S^2 - m_X^2)}{(m_\eta^2 - m_X^2)} - \ell(m_X^2) \left[ \frac{(m_{xy}^2 + m_X^2)(m_D^2 - m_X^2)(m_S^2 - m_X^2)}{(m_\eta^2 - m_X^2)^2} + \frac{2(m_{xy}^2 + m_X^2)(m_D^2 - m_X^2)(m_S^2 - m_X^2)}{(m_Y^2 - m_X^2)(m_\eta^2 - m_X^2)} + \frac{(m_D^2 - m_X^2)(m_S^2 - m_X^2) - (m_{xy}^2 + m_X^2)(m_S^2 - m_X^2) - (m_{xy}^2 + m_X^2)(m_D^2 - m_X^2)}{(m_\eta^2 - m_X^2)} \right], \quad (\text{A4})$$

$$I_Y = I_X(X \leftrightarrow Y), \quad (\text{A5})$$

$$I_\eta = \ell(m_\eta^2) \frac{(m_X^2 - m_Y^2)^2 (m_{xy}^2 + m_\eta^2)(m_D^2 - m_\eta^2)(m_S^2 - m_\eta^2)}{(m_X^2 - m_\eta^2)^2 (m_Y^2 - m_\eta^2)^2}. \quad (\text{A6})$$

The expression appears singular, but in fact is not—the poles all cancel when  $m_X^2, m_Y^2 \rightarrow m_\eta^2$  or  $m_X^2 \rightarrow m_Y^2$ . Nevertheless, we have not found a simpler expression. We have also checked that this expression becomes equal to the two sea quark result (as given most compactly in Becirevic and Villadoro [25]) in the limit that  $m_S^2 \gg m_D^2, m_X^2, m_Y^2, m_{xy}^2$  (so that  $m_\eta^2 = 2m_S^2/3$ ).

### APPENDIX B: REVIEW OF STAGGERED CHIRAL PERTURBATION THEORY

Here we recall the leading-order effective staggered chiral Lagrangian, whose construction is discussed in much greater detail in Ref. [8], and discuss some of its general consequences. The next-to-leading-order Lagrangian, determined in Ref. [10], turns out not to be necessary for  $B_K$  at NLO.

For  $n$  staggered quark flavors, spontaneous breakdown of the approximate  $SU(4n)$  chiral symmetry by the vacuum,

$$SU(4n)_L \times SU(4n)_R \rightarrow SU(4n)_V, \quad (\text{B1})$$

leads to  $16n^2 - 1$  PGBs. We collect them into an  $SU(4n)$  matrix,

$$\Sigma = \exp(i\Phi/f), \quad (\text{B2})$$

where  $\Phi$  is a traceless  $4n \times 4n$  matrix with  $4 \times 4$  submatrices:

$$\Phi = \begin{pmatrix} U & \pi^+ & K^+ & \cdots \\ \pi^- & D & K^0 & \cdots \\ K^- & \bar{K}^0 & S & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (\text{B3})$$

$$U = \sum_{a=1}^{16} U_a T_a, \text{ etc.} \quad (\text{B4})$$

and  $f$  is normalized such that  $f_\pi \approx 132$  MeV. We use Euclidean gamma matrices as the  $SU(4)$  generators:

$$T_a = \{\xi_5, i\xi_{\mu 5}, i\xi_{\mu\nu}, \xi_\mu, \xi_I\}, \quad (\text{B5})$$

where  $\xi_I$  is the  $4 \times 4$  identity matrix. The quark mass matrix is also of size  $4n \times 4n$ , but has trivial (singlet) taste structure:

$$\mathcal{M} = \begin{pmatrix} m_u I & 0 & 0 & \cdots \\ 0 & m_d I & 0 & \cdots \\ 0 & 0 & m_s I & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (\text{B6})$$

Under chiral symmetry transformations,  $\Sigma$  transforms as

$$\Sigma \rightarrow L \Sigma R^\dagger, \quad (\text{B7})$$

$$L \in SU(4n)_L, \quad R \in SU(4n)_R. \quad (\text{B8})$$

The standard  $S\chi$ PT power-counting scheme is<sup>26</sup>

$$p^2 \approx m \approx a^2, \quad (\text{B9})$$

so the lowest-order Lagrangian is of  $\mathcal{O}(p^2, m, a^2)$ :

$$\mathcal{L}_\chi = \frac{f^2}{8} \text{Tr}(\partial_\mu \Sigma \partial_\mu \Sigma^\dagger) - \frac{1}{4} \mu f^2 \text{Tr}(\mathcal{M} \Sigma + \mathcal{M} \Sigma^\dagger) + \frac{2m_0^2}{3} (U_I + D_I + S_I)^2 + a^2 \mathcal{V}, \quad (\text{B10})$$

where  $\text{Tr}$  is the full  $4n \times 4n$  trace in both flavor and taste space,  $\mu$  is a dimensionful constant of  $\mathcal{O}(\Lambda_{\text{QCD}})$ , and  $\mathcal{V}$  is the taste-symmetry breaking potential.

The taste-breaking potential leads to important properties of staggered mesons. For discussion of these features it is useful to separate  $\mathcal{V}$  into single and double trace operators:

$$\mathcal{V} = \mathcal{U} + \mathcal{U}', \quad (\text{B11})$$

$$\begin{aligned} -\mathcal{U} &= C_1 \text{Tr}(\xi_5^{(n)} \Sigma \xi_5^{(n)} \Sigma^\dagger) \\ &+ C_3 \frac{1}{2} \sum_\nu [\text{Tr}(\xi_\nu^{(n)} \Sigma \xi_\nu^{(n)} \Sigma) + \text{H.c.}] \\ &+ C_4 \frac{1}{2} \sum_\nu [\text{Tr}(\xi_{\nu 5}^{(n)} \Sigma \xi_{\nu 5}^{(n)} \Sigma) + \text{H.c.}] \\ &+ C_6 \sum_{\mu < \nu} \text{Tr}(\xi_{\mu\nu}^{(n)} \Sigma \xi_{\nu\mu}^{(n)} \Sigma^\dagger), \end{aligned} \quad (\text{B12})$$

<sup>26</sup>Note that standard  $S\chi$ PT does not include  $\alpha$  in its power counting because there are no external operators.

$$\begin{aligned}
-\mathcal{U}' &= C_{2V} \frac{1}{4} \sum_{\nu} [\text{Tr}(\xi_{\nu}^{(n)} \Sigma) \text{Tr}(\xi_{\nu}^{(n)} \Sigma) + \text{H.c.}] \\
&+ C_{2A} \frac{1}{4} \sum_{\nu} [\text{Tr}(\xi_{\nu 5}^{(n)} \Sigma) \text{Tr}(\xi_{\nu 5}^{(n)} \Sigma) + \text{H.c.}] \\
&+ C_{5V} \frac{1}{2} \sum_{\nu} [\text{Tr}(\xi_{\nu}^{(n)} \Sigma) \text{Tr}(\xi_{\nu}^{(n)} \Sigma^{\dagger})] \\
&+ C_{5A} \frac{1}{2} \sum_{\nu} [\text{Tr}(\xi_{\nu 5}^{(n)} \Sigma) \text{Tr}(\xi_{\nu 5}^{(n)} \Sigma^{\dagger})], \quad (\text{B13})
\end{aligned}$$

where  $\xi_T^{(n)}$  is a  $4n \times 4n$  matrix with an ordinary  $4 \times 4$  taste matrix,  $\xi_T$ , repeated along the diagonal:

$$\xi_T^{(n)} = \begin{pmatrix} \xi_T & 0 & 0 & \cdots \\ 0 & \xi_T & 0 & \cdots \\ 0 & 0 & \xi_T & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (\text{B14})$$

The single trace operators in  $\mathcal{U}$  split the tree-level PGB masses into degenerate groups:

$$(m_{\pi}^2)_{\text{LO}} = \mu(m_i + m_j) + a^2 \Delta_F, \quad (\text{B15})$$

where  $\Delta_F$  is different for each of the five  $SO(4)$ -taste irreps  $P$ ,  $A$ ,  $T$ ,  $V$ , and  $I$ .<sup>27</sup> The double trace operators in  $\mathcal{U}'$  generate hairpin diagrams (quark disconnected contractions) for flavor-neutral, taste-vector and axial-vector PGBs. Both  $\mathcal{U}$  and  $\mathcal{U}'$  produce interaction vertices among even numbers of PGBs. In addition, the  $m_0^2$  term, present because only the *overall* singlet can be integrated out of the theory, generates flavor-neutral, taste-singlet hairpins.

It is easy to generalize the standard staggered chiral Lagrangian to quenched and partially quenched theories. For a PQ theory with  $N = n_{\text{val}} + n_{\text{sea}}$  quarks and  $M = n_{\text{val}}$  ghosts, the chiral symmetry group is  $SU(4N|4M)_L \times SU(4N|4M)_R$  and breaks to  $SU(4N|4M)_V$ . The net result is that there are more PGBs, and all of the traces become supertraces ( $\text{Tr} \rightarrow \text{Str}$ ) in the chiral Lagrangian.

### APPENDIX C: DETERMINATION OF CHIRAL $B_K$ OPERATORS

Here we determine all of the mesonic operators which contribute to  $\mathcal{M}_{\text{lat}}$ , and consequently to  $B_K$ , at LO and NLO in our power counting. There are quite a few such operators because the continuum  $SU(3)_{\text{flavor}}$  is enlarged to the graded, staggered symmetry  $SU(4N_{\text{val}}|4N_{\text{sea}}|4N_{\text{val}})$  on the lattice. Fortunately, the fact that  $\mathcal{O}_K$  is invariant under the lattice  $U(1)_A$  symmetry protects the number of operators that contribute to  $B_K$  from becoming unwieldy.

We first enumerate the operators which contribute to  $B_K$  at tree level. These include the standard  $B_K$  chiral operator, and are relatively easy to determine, even in the staggered

theory. Next we enumerate the operators which contribute to  $B_K$  at 1 loop. These can come either from operator mixing on the lattice or from insertions (at the quark level) of the staggered action with the  $B_K$  operator. We then determine the analytic contributions to  $B_K$  at NLO, relying heavily on constraints from  $U(1)_A$  symmetry and from  $CP$  plus  $s \leftrightarrow d$  interchange. Finally we discuss the parametric dependence of the various operator coefficients on the lattice spacing,  $a$ , and the strong coupling constant,  $\alpha$ .

#### 1. Operators that contribute at lowest order

We must map the staggered quark-level operator,  $\mathcal{O}_K^{\text{staggered}}$ , onto operators in the chiral effective theory. This is done most simply by first splitting the operator into spin vector and spin axial-vector parts, mapping *these* operators onto mesonic operators, and then taking appropriate linear combinations to get chiral operators which correspond to the staggered  $B_K$  operator.

We first separate  $\mathcal{O}_K^{\text{staggered}}$  into two pieces<sup>28</sup>:

$$\mathcal{O}_V \equiv [\overline{s1}(\gamma_{\mu} \otimes \xi_5)d1][\overline{s2}(\gamma_{\mu} \otimes \xi_5)d2], \quad (\text{C1})$$

$$\mathcal{O}_A \equiv [\overline{s1}(\gamma_{\mu} \gamma_5 \otimes \xi_5)d1][\overline{s2}(\gamma_{\mu} \gamma_5 \otimes \xi_5)d2]. \quad (\text{C2})$$

The  $B_K$  matrix element is proportional to their sum:

$$\mathcal{M}_K^{\text{staggered}} \propto \langle \overline{K}_2^0 | (\mathcal{O}_V + \mathcal{O}_A) | K_1^0 \rangle. \quad (\text{C3})$$

As in Ref. [26], we label these four-fermion operators by the spin and taste in each bilinear, such that  $\mathcal{O}_V \propto [V \times P]$  and  $\mathcal{O}_A \propto [A \times P]$ . Note the convention that  $A \equiv \gamma_{\mu 5} \otimes \gamma_{5\mu}$ , where the pair of spin (or taste) matrices are in separate bilinears, such that

$$\mathcal{O}_K^{\text{staggered}} \propto [V \times P] - [A \times P] \quad (\text{C4})$$

in this condensed notation.

We now use spurion analysis as in Refs. [10,26] to determine the resulting chiral operators. The chiral structure of four-fermion operators with spins  $V$  or  $A$  is

$$\mathcal{O}_F = \pm \sum_{\mu} (\overline{Q}_R(\gamma_{\mu} \otimes F_R)Q_R \pm \overline{Q}_L(\gamma_{\mu} \otimes F_L)Q_L)^2, \quad (\text{C5})$$

where the upper and lower signs correspond to  $V$  and  $A$ , respectively, and  $F_{L,R}$  are Hermitian taste matrices. When we promote the taste matrices to spurion fields,  $F_{L,R}$  must transform in the following way so that  $\mathcal{O}_F$  is invariant under chiral transformations:

$$F_L \rightarrow L F_L L^{\dagger}, \quad F_R \rightarrow R F_R R^{\dagger}. \quad (\text{C6})$$

We will set  $F_L$  and  $F_R$  to the appropriate matrices at the end.

<sup>27</sup>We assume in this expression that the quark flavors  $i$  and  $j$  are different.

<sup>28</sup>From now on we suppress color indices because both color contractions lead to the same chiral operators.

TABLE I. The two linearly independent  $\mathcal{O}(p^2)$  operators corresponding to  $[(V - A) \times F]$ . The  $\pm$  signs correspond to spins  $V$  and  $A$ , respectively. Derivatives act only on the object immediately to the right.

Operator
$[\pm \text{Str}(\Sigma \partial_\mu \Sigma^\dagger F_L \Sigma \partial_\mu \Sigma^\dagger F_L) \pm \text{Str}(\Sigma^\dagger \partial_\mu \Sigma F_R \Sigma^\dagger \partial_\mu \Sigma F_R)]$
$[\pm \text{Str}(\Sigma \partial_\mu \Sigma^\dagger F_L) \text{Str}(\Sigma \partial_\mu \Sigma^\dagger F_L) \pm \text{Str}(\Sigma^\dagger \partial_\mu \Sigma F_R) \text{Str}(\Sigma^\dagger \partial_\mu \Sigma F_R)]$

We first note that chiral operators which result from two left-handed (or two right-handed) taste spurions have a *relative minus sign* when they come from a spin  $V$  versus a spin  $A$  four-fermion operator, whereas those that result from one left-handed and one right-handed taste spurion do not. Thus only operators built out of two spurions with the same handedness remain after taking the linear spin combination  $(V - A)$  which corresponds to  $\mathcal{O}_K^{\text{staggered}}$ . There are no such operators without derivatives, and two such linearly independent ones with two derivatives. These are shown in Table I.

The expression in Eq. (C5) can be made to have the same flavor and taste structure as  $\mathcal{O}_K^{\text{staggered}}$  by setting the two spurions, which we now call  $F_1$  and  $F_2$ , equal to sparse matrices with a single  $\xi_5$  in either the  $s1\bar{d}1$  or  $s2\bar{d}2$  location. With this replacement the first operator in Table I becomes

$$\begin{aligned} \mathcal{O}_\chi^N &= \sum_\mu (\text{Str}(\Sigma \partial_\mu \Sigma^\dagger F_1 \Sigma \partial_\mu \Sigma^\dagger F_2) \\ &\quad + \text{Str}(\Sigma^\dagger \partial_\mu \Sigma F_1 \Sigma^\dagger \partial_\mu \Sigma F_2)) \end{aligned} \quad (\text{C7})$$

and the second becomes

$$\begin{aligned} \mathcal{O}_\chi^K &= \sum_\mu (\text{Str}(\Sigma \partial_\mu \Sigma^\dagger F_1) \text{Str}(\Sigma \partial_\mu \Sigma^\dagger F_2) \\ &\quad + \text{Str}(\Sigma^\dagger \partial_\mu \Sigma F_1) \text{Str}(\Sigma^\dagger \partial_\mu \Sigma F_2)). \end{aligned} \quad (\text{C8})$$

Since  $\mathcal{O}_\chi^K$  has the same left-left current structure as the standard continuum full QCD operator [18], it should contribute to  $B_K$  at tree level, and it does. The matrix  $F_1$  annihilates the  $K_1^0$  (an  $\bar{s}1d1$  meson) while  $F_2$  produces the  $\bar{K}_2^0$  (a  $\bar{d}2s2$  meson).

Because  $B_K$  in continuum  $\chi$ PT only depends on a single parameter [18], the coefficients of  $\mathcal{O}_\chi^K$  and  $\mathcal{O}_\chi^N$  cannot be independent. We can determine their precise relationship using the fact that both operators in the chiral effective theory arise from the same quark-level operator,  $\mathcal{O}_K^{\text{staggered}}$ :

$$\mathcal{O}_K^{\text{staggered}} \xrightarrow{S\chi\text{PT}} C_\chi^K \mathcal{O}_\chi^K + C_\chi^N \mathcal{O}_\chi^N, \quad (\text{C9})$$

but they have nonvanishing tree-level matrix elements between different types of kaons. In particular, while  $\mathcal{O}_\chi^K$  only has a tree-level matrix element between  $\bar{K}_1^0$  and  $K_2^0$  (by construction),  $\mathcal{O}_\chi^N$  only has one between mixed kaons,

i.e.  $s1\bar{d}2$  and  $\bar{s}2d1$ :

$$\langle \bar{K}_{1P}^0 | C_\chi^K \mathcal{O}_\chi^K + C_\chi^N \mathcal{O}_\chi^N | K_{2P}^0 \rangle = \frac{32 C_\chi^K}{f^2} m_K^2, \quad (\text{C10})$$

$$\langle \bar{K}_{1P}^{0\text{mix}} | C_\chi^K \mathcal{O}_\chi^K + C_\chi^N \mathcal{O}_\chi^N | K_{2P}^{0\text{mix}} \rangle = \frac{8 m_K^2 C_\chi^N}{f^2}. \quad (\text{C11})$$

By comparing the same two matrix elements at the quark level we can extract the ratio  $C_\chi^K : C_\chi^N$ . As written in Eq. (6),  $\mathcal{O}_K^{\text{staggered}}$  can contract with  $\bar{K}_1^0$  and  $K_2^0$ , but not the mixed kaons. However, it can be Fierz transformed into an operator with the appropriate mixed-flavor structure, at the cost of introducing new tastes. Using the Fierz-transformation rules in Appendix A of Ref. [27], we find that

$$[(V - A) \times P] \xrightarrow{\text{Fierz}} \frac{1}{4} [(V - A) \times (S + P + T - V - A)]. \quad (\text{C12})$$

Thus the ratio of the two matrix elements is

$$\frac{\langle \bar{K}_{1P}^0 | \mathcal{O}_K^{\text{staggered}} | K_{2P}^0 \rangle}{\langle \bar{K}_{1P}^{0\text{mix}} | \mathcal{O}_K^{\text{staggered}} | K_{2P}^{0\text{mix}} \rangle} = 4, \quad (\text{C13})$$

and  $C_\chi^K = C_\chi^N$ . The Fierz transformation also reveals why the seemingly ‘‘extra’’ operator,  $\mathcal{O}_\chi^N$ , is necessary in the chiral effective theory—it accounts for the fact that  $\mathcal{O}_K^{\text{staggered}}$  has a nonvanishing tree-level matrix element between flavor-mixed kaons of *all* tastes.

## 2. Operators that contribute at NLO

$B_K$  receives NLO contributions from a host of additional operators because of lattice discretization effects. In order to determine all such operators we must consider the three different ways in which they can arise at the quark level: through perturbative matching, mixing with higher-dimension operators, and insertions of operators from the staggered action.

### a. 1-loop matching coefficients

Lattice operators can mix with each other as long as they are in the same representation of the symmetry group that maps a hypercube onto itself. Consequently, lattice operators which correspond to continuum operators with a particular spin and taste *mix* with other operators that correspond to a *different continuum spin and taste*. Specifically, although  $\mathcal{O}_K^{\text{staggered}}$  in the continuum is strictly a  $[(V - A) \times P]$  four-fermion operator, once it is put on the lattice it mixes with  $[(V + A) \times P]$  and other spin-taste structures. Like all taste-breaking effects, this operator mixing is due to hard gluon exchange, so the amount of mixing can be calculated perturbatively in  $\alpha$ . Such 1-loop ‘‘matching coefficients’’ for staggered four-fermion operators have been calculated for both naive fermions and using

TABLE II. Mesonic operators corresponding to the four-fermion operators listed in Eq. (C15). Repeated indices are summed with the constraint that  $\mu \neq \nu$  in the taste tensor matrices.

Four-fermion operator	Chiral operator
$[V \times S]$ and $[A \times S]$	$\rightarrow \text{Str}(\xi_I \Sigma \xi_I \Sigma^\dagger)$
$[V \times T]$ and $[A \times T]$	$\rightarrow \text{Str}(\xi_{\mu\nu} \Sigma \xi_{\nu\mu} \Sigma^\dagger)$
$[S \times V]$ and $[P \times V]$	$\rightarrow \text{Str}(\xi_\mu \Sigma) \text{Str}(\xi_\mu \Sigma^\dagger)$ $\text{Str}(\xi_\mu \Sigma) \text{Str}(\xi_\mu \Sigma) + \text{Str}(\xi_\mu \Sigma^\dagger) \text{Str}(\xi_\mu \Sigma^\dagger)$ $\text{Str}(\xi_\mu \Sigma \xi_\mu \Sigma) + \text{Str}(\xi_\mu \Sigma^\dagger \xi_\mu \Sigma^\dagger)$
$[S \times A]$ and $[P \times A]$	$\rightarrow \text{Str}(\xi_{\mu 5} \Sigma) \text{Str}(\xi_{5\mu} \Sigma^\dagger)$ $\text{Str}(\xi_{\mu 5} \Sigma) \text{Str}(\xi_{5\mu} \Sigma) + \text{Str}(\xi_{\mu 5} \Sigma^\dagger) \text{Str}(\xi_{5\mu} \Sigma^\dagger)$ $\text{Str}(\xi_{\mu 5} \Sigma \xi_{5\mu} \Sigma) + \text{Str}(\xi_{\mu 5} \Sigma^\dagger \xi_{5\mu} \Sigma^\dagger)$

hypercubic fat links [14]. Some of this operator mixing is accounted for and removed in present lattice calculations of  $B_K$ . It is now standard to use 1-loop matching coefficients to remove the  $\mathcal{O}(\alpha/4\pi)$  mixing between the desired  $[(V - A) \times P]$  lattice operator and the unwelcome  $[(V + A) \times P]$  lattice operator. This still leaves mixing at  $\mathcal{O}(\alpha^2)$  with  $[(V + A) \times P]$  and at  $\mathcal{O}(\alpha/4\pi)$  with wrong-taste operators. We discuss them in turn.

Lattice operators with spin structure  $(V + A)$  contribute to  $B_K$ , albeit suppressed by  $\mathcal{O}(\alpha^2)$ . Recall from the previous section that such operators arise from the combination of one left-handed spurion and one right-handed spurion, where  $F_L$  and  $F_R$  transform as in Eq. (C6). Here we need only consider operators without derivatives. This is because any two-derivative operators are of  $\mathcal{O}(\alpha^2 p^2)$  and can only contribute to  $B_K$  at NLO through analytic terms, which we determine separately in Appendix C 3. There is only one operator without derivatives,

$$\mathcal{O}_\chi^{1P} = \text{Str}(F_1 \Sigma F_2 \Sigma^\dagger), \quad (\text{C14})$$

and it cannot contribute to  $B_K$  until 1 loop because of its flavor structure. Note that the 1-loop contribution of  $\mathcal{O}_\chi^{1P}$  is of NLO in our power counting because its coefficient is of  $\mathcal{O}(\alpha^2)$  and the loop momentum further suppresses it by  $\mathcal{O}(p^2)$ .

Lattice calculations to date have not taken into account the  $\mathcal{O}(\alpha/4\pi)$  mixing between  $[(V - A) \times P]$  and lattice operators of different tastes.  $B_K$  therefore receives contributions from wrong-taste lattice operators, but suppressed by  $\mathcal{O}(\alpha/4\pi)$ . Using the matching coefficients from Ref. [14] we find that the desired lattice operator mixes with the following eight four-fermion operators at 1 loop:

$$\begin{aligned} &[V \times S], \quad [A \times S], \quad [V \times T], \quad [A \times T], \quad [S \times V], \\ &[S \times A], \quad [P \times V], \quad [P \times A], \quad [T \times V], \quad [T \times A], \end{aligned} \quad (\text{C15})$$

as well as taste off-diagonal operators, i.e. those in which the two bilinears have different tastes, and operators which violate rotational and  $SO(4)$ -taste symmetry. We map all of these quark-level operators onto chiral operators using a straightforward generalization of the spurion analysis in

Ref. [26] to flavor off-diagonal operators. Table II contains the results. We do not show the chiral operators generated by  $[T \times V, A]$  because they are simply a subset of those from  $[S, P \times V, A]$ . Nor do we include taste off-diagonal operators, as they do not contribute to any PGB quantities of interest. Furthermore, we neglect operators which break rotational and/or  $SO(4)$ -taste symmetry, as doing so would require at least four repeated Lorentz indices in the chiral operator. This would have to come from the presence of four derivatives, four additional taste matrices, or two derivatives and two taste matrices, and therefore be of a higher order than we consider here.

Because these operators arise from mixing with  $\mathcal{O}_K^{\text{staggered}}$ , they have the *same flavor structure* as the original continuum operator. Thus  $\xi_{\mu\nu} \rightarrow F_{1T}$  and  $\xi_{\nu\mu} \rightarrow F_{2T}$ , where  $F_{1T}$  ( $F_{2T}$ ) is a sparse matrix with a single  $\xi_{\mu\nu}$  ( $\xi_{\nu\mu}$ ) in the  $s1\bar{d}1$  ( $s2\bar{d}2$ ) location. We define  $F_{1I}, F_{1J}, F_{1V}, F_{2V}, F_{1A}$ , and  $F_{2A}$  analogously, such that the following eight operators contribute to  $B_K$ <sup>29</sup>:

$$\mathcal{O}_\chi^{1I} = \sum_{\mu \neq \nu} \text{Str}(F_{1I} \Sigma F_{2I} \Sigma^\dagger), \quad (\text{C16})$$

$$\mathcal{O}_\chi^{1T} = \sum_{\mu \neq \nu} \text{Str}(F_{1T} \Sigma F_{2T} \Sigma^\dagger), \quad (\text{C17})$$

$$\mathcal{O}_\chi^{1V} = \sum_{\mu} (\text{Str}(F_{1V} \Sigma F_{2V} \Sigma) + \text{Str}(F_{1V} \Sigma^\dagger F_{2V} \Sigma^\dagger)), \quad (\text{C18})$$

$$\begin{aligned} \mathcal{O}_\chi^{2V} &= \sum_{\mu} (\text{Str}(F_{1V} \Sigma) \text{Str}(F_{2V} \Sigma) \\ &+ \text{Str}(F_{1V} \Sigma^\dagger) \text{Str}(F_{2V} \Sigma^\dagger)), \end{aligned} \quad (\text{C19})$$

$$\mathcal{O}_\chi^{3V} = \sum_{\mu} \text{Str}(F_{1V} \Sigma) \text{Str}(F_{2V} \Sigma^\dagger), \quad (\text{C20})$$

<sup>29</sup>We note as an interesting aside that  $\mathcal{O}_\chi^{1I}$  does not appear in the staggered chiral Lagrangian because, if  $\xi_I$  is flavor diagonal, it is just a trivial constant.

$$\mathcal{O}_X^{1A} = \sum_{\mu} (\text{Str}(F_{1A}\Sigma F_{2A}\Sigma) + \text{Str}(F_{1A}\Sigma^{\dagger} F_{2A}\Sigma^{\dagger})), \quad (\text{C21})$$

$$\mathcal{O}_X^{2A} = \sum_{\mu} (\text{Str}(F_{1A}\Sigma)\text{Str}(F_{2A}\Sigma) + \text{Str}(F_{1A}\Sigma^{\dagger})\text{Str}(F_{2A}\Sigma^{\dagger})), \quad (\text{C22})$$

$$\mathcal{O}_X^{3A} = \sum_{\mu} \text{Str}(F_{1A}\Sigma)\text{Str}(F_{2A}\Sigma^{\dagger}). \quad (\text{C23})$$

Moreover, since Eq. (C15) turns out to be a complete list of  $U(1)_A$  and  $SO(4)$ -rotation invariant four-fermion operators, these are also the only wrong-taste operators that can contribute to  $B_K$  at  $\mathcal{O}(\alpha^2)$ . Thus all of the operators in Eq. (C23) have two independent coefficients, one of  $\mathcal{O}(\alpha/4\pi)$  and the other of  $\mathcal{O}(\alpha^2)$ . Note that, as one would expect, such wrong-taste operators can only contribute to the  $B_K$  matrix element through loops. Such contributions are of NLO in our power counting because of the factor of  $\alpha/4\pi$  or  $\alpha^2$  in the coefficients and the further suppression of  $\mathcal{O}(p^2)$  by the loop momentum. We therefore do not need to consider wrong-taste operators with derivatives.

### b. Dimension 8 quark-level operators

The four-fermion operator  $\mathcal{O}_K^{\text{staggered}}$  is dimension 6. As discussed in the previous subsection, it mixes with other dimension 6 operators on the lattice, but such mixing is due to gluon exchange and suppressed by  $\alpha/4\pi$ . Therefore the naively discretized version of the continuum operator,  $\mathcal{O}_K^{\text{staggered,lat}}$ , generates the dominant contribution to the lattice  $\overline{K}_2^0 - K_1^0$  matrix element,  $\mathcal{M}_{\text{lat}}$ . However, many higher-dimension quark-level operators respect the *same lattice symmetries* as  $\mathcal{O}_K^{\text{staggered}}$  and thus can also contribute to  $B_K$  on the lattice. Dimension 7 and 8 quark-level operators are explicitly suppressed in the action relative to dimension 6 operators by factors of  $a$  and  $a^2$ , respectively, so they can map onto chiral operators which generate 1-loop contributions to  $B_K$  at NLO. In fact, because  $a^2 \sim \alpha^2 \sim \alpha/4\pi$  in our power counting, they are just as or more important than the dimension 6 operators which contribute through mixing. Fortunately we need not go higher than dimension 8 at the quark level because 1-loop contributions to  $B_K$  from dimension 9 operators are at least  $\mathcal{O}(a^3 p^2)$  and thus of higher than NLO.

At the quark level, dimension 7 operators contain four fermions plus either a derivative or a factor of the quark mass matrix. However, all dimension 7 quark-level operators turn out to be *taste off diagonal* [28], and therefore cannot contribute to any PGB processes of interest. Thus we move on to dimension 8 quark-level operators, which contain four fermions plus two derivatives, two factors of the quark mass matrix, or one of each. Dimension 8 operators with one derivative and one mass matrix must have a single nontrivial spin or taste matrix (either  $V$  or  $A$ )

in order to contract up the free derivative index. Such operators are either spin off diagonal or taste off diagonal; both are irrelevant. Dimension 8 operators with two mass matrices map onto chiral operators with at least two mass spurions. Because they also carry a factor of  $a^2$  from the action, such operators are of  $\mathcal{O}(a^2 m^2)$  or higher in our power counting and can be neglected. We therefore need only consider dimension 8 four-fermion operators with two derivatives that possess the same lattice symmetries as  $\mathcal{O}_K^{\text{staggered}}$ . Enumerating all such operators would be a tedious exercise in staggered group theory. Fortunately, this exercise is unnecessary if we start with all *chiral* operators which could contribute to  $B_K$  at NLO and work backwards.

All chiral operators which arise from dimension 8 quark-level operators automatically come with a power of  $a^2$ . Thus, in order to generate 1-loop contributions to  $B_K$  at NLO, they cannot contain any derivatives or be suppressed by additional powers of  $a$  or  $\alpha$ . All of the operators listed in Table II satisfy this condition. In fact, it turns out that there are no additional chiral operators which are invariant under  $U(1)_A$  and derivative free. Therefore all of the chiral operators in Eqs. (C14) and (C16)–(C23) should contribute to  $B_K$  both through 1-loop mixing, suppressed by  $\mathcal{O}(\alpha/4\pi)$ , and through dimension 8 operators, suppressed by  $\mathcal{O}(a^2)$ .

Yet there is a subtlety—not all  $a^2$ 's are created equal. This is because, at the quark level, taste-changing four-fermion interactions require gluon exchange, and therefore have hidden powers of  $\alpha$  which do not appear explicitly in the chiral effective Lagrangian, Eq. (B10). We must therefore consider the  $\mathcal{O}(\alpha^2)$  operators in more detail to allow correct extrapolation of lattice data.

In the case of matching the continuum version of  $\mathcal{O}_K^{\text{staggered}}$  onto dimension 8 lattice four-fermion operators, we want to find the dimension 8 operators that match onto  $[(V - A) \times P]$  *without* gluon exchange, because they truly contribute at  $\mathcal{O}(a^2)$ . Any other dimension 8 operators contribute at  $\mathcal{O}(\alpha a^2)$  or higher, and thus are of at least NNLO in our power counting. As discussed in both Refs. [29,30], lattice bilinears of different spins and tastes “mix”<sup>30</sup> with each other at tree level in perturbation theory, but to all orders in  $a$ , simply because of the fact that the spinor components of staggered fermions are spread out over a hypercube.<sup>31</sup> Fortunately, at tree level, one can consider the two bilinears in each four-fermion operator separately. Thus, in order to determine which four-fermion operators mix at  $\mathcal{O}(a^2)$ , one needs only find *which bilinears mix at  $\mathcal{O}(1)$ ,  $\mathcal{O}(a)$ , and  $\mathcal{O}(a^2)$* . The product of two  $\mathcal{O}(a)$  bilinears, or of one  $\mathcal{O}(1)$  and one  $\mathcal{O}(a^2)$

<sup>30</sup>Recall that we are using the terminology mix as shorthand for this process of matching lattice operators onto those in the continuum.

<sup>31</sup>Note that these mixings are independent of both the choice of gauge links and the choice of how to construct gauge-invariant operators.

bilinear, will be a four-fermion operator that mixes at  $\mathcal{O}(a^2)$ .

We do so using Eqs. (26) and (27) in Ref. [30]. We find that the dimension 3 bilinear  $[\gamma_\mu(1 + \gamma_5) \otimes \xi_5]$  mixes with taste  $V$  dimension 4 bilinears at  $\mathcal{O}(a)$ . It also mixes with dimension 3 and 5 taste  $P$  bilinears at  $\mathcal{O}(1)$  and  $\mathcal{O}(a^2)$ , respectively. Because spin off-diagonal four-fermion operators cannot contribute to PGB processes, we are restricted to multiplication of each of these bilinears by itself. For example, we can multiply the bilinear  $[1 \times \xi_\mu]$  times itself to generate the four-fermion operator

$$\sum_\mu \sum_\nu \bar{Q} D_\mu (I \otimes \xi_\mu) Q \bar{Q} D_\nu (I \otimes \xi_\nu) Q. \quad (\text{C24})$$

The above operator is simply a generalization of the dimension 6 four-fermion operator  $[S \times V]$  to dimension 8. The derivatives give it the appropriate dimension, and allow it to have both  $SO(4)$  rotational symmetry breaking ( $\mu = \nu$ ) and conserving ( $\mu \neq \nu$ ) pieces.

It is clear now that the lattice version of  $\mathcal{O}_K^{\text{staggered}}$  really only mixes at  $\mathcal{O}(a^2)$  with dimension 8 quark-level operators that have spin-taste structures  $[S \times V]$ ,  $[P \times V]$ ,  $[T \times V]$ ,  $[(V - A) \times P]$ , and  $[(V + A) \times P]$  along with two derivatives. These lead to the four chiral operators  $\mathcal{O}_\chi^{1V}$ ,  $\mathcal{O}_\chi^{2V}$ ,  $\mathcal{O}_\chi^{3V}$ , and  $\mathcal{O}_\chi^{1P}$ .<sup>32</sup> We therefore conclude that their coefficients are of  $\mathcal{O}(a^2)$ , while the rest of the operators enumerated in this section can be neglected. This is summarized in the second column of Table IV.

### c. Insertions of the staggered action

The lattice operator mixing discussed in the past two subsections is not particularly intuitive because it stems from perturbation theory and complicated lattice symmetries. In contrast, the fact that operators from the staggered action modify the quark-level  $B_K$  operator,  $\mathcal{O}_K^{\text{staggered}}$ , can be shown through a simple diagrammatic argument. Consider the four-fermion  $\mathcal{O}_K^{\text{staggered}}$  and another four-fermion operator from the staggered action. Two of the quarks in  $\mathcal{O}_K^{\text{staggered}}$  can contract with two of the quarks in the other operator to form a 1-loop diagram with four external quarks, as shown in Fig. 3. This produces a local operator in the chiral effective theory because the loop propagates over  $\mathcal{O}(1/\Lambda_{\text{QCD}})$ , rather than over  $\mathcal{O}(1/m_\pi)$ . It corrects  $\mathcal{O}_K^{\text{staggered}}$  at  $\mathcal{O}(a_\alpha^2)$  because all of the dimension 6 four-fermion operators in the staggered action are taste breaking. Such  $\mathcal{O}(a_\alpha^2)$  corrections turn out to be numerically enhanced, however, so they are still of NLO in our power counting. We must therefore determine all chiral operators produced by combining  $\mathcal{O}_K^{\text{staggered}}$  with each of

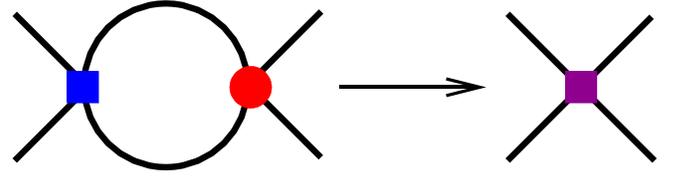


FIG. 3 (color online). Schematic of  $B_K$  vertex modification due to the insertion of an operator from the staggered action. In the first diagram the square represents the original  $B_K$  vertex and the circle represents the inserted strong four-fermion vertex. In the second diagram the square represents the resulting local vertex in the effective theory.

the dimension 6 four-fermion operators in the staggered action.

Reference [10] developed a method for combining four-fermion operators at the chiral level. Essentially, one forms chirally invariant operators which contain two taste spurions from the first four-fermion operator and two from the second.  $\mathcal{O}_K^{\text{staggered}}$  has spin-taste structure  $[(V - A) \times P]$ , so the two taste spurions from  $\mathcal{O}_K^{\text{staggered}}$  must transform as in Eq. (C6). Moreover, we learned in Appendix C 1 that only operators with two left-handed (or two right-handed) taste spurions survive after taking the linear spin combination  $(V - A)$ . In contrast, the four-fermion operator may have any spin— $V$ ,  $A$ ,  $S$ ,  $P$ , or  $T$ —combined with an appropriate taste to maintain overall  $U(1)_A$  invariance. Furthermore, the two taste spurions from the inserted four-fermion operator must be *flavor diagonal* because the action does not change quark flavor.

Operators with spins  $V$  or  $A$  can have tastes  $S$ ,  $P$ , or  $T$ , and their taste spurions transform as Eq. (C6). Generically, there are three possible ways of combining these spurions with those from  $\mathcal{O}_K^{\text{staggered}}$ :

$$(F_L F_L)(F'_L F'_R) + \text{p.c.}, \quad (F_L F_L)(F'_L F'_L) + \text{p.c.}, \quad (\text{C25}) \\ (F_L F_L)(F'_R F'_R) + \text{p.c.},$$

where “p.c.” indicates parity conjugate. Here the unprimed spurions are from  $\mathcal{O}_K^{\text{staggered}}$  and the primed are from the four-fermion operator in the staggered action. Because the two  $F_L$ ’s will ultimately have different off-diagonal flavor structures, they must be separated by at least one  $\Sigma$  field on either side within a supertrace. We use this fact to eliminate a few operators. Consequently, these spurion combinations only lead to two nontrivial generic chiral structures, shown at the top of Table III. To produce operators which contribute to  $B_K$  we let  $F_L \rightarrow F_1 \& F_2$  and  $F'_R \rightarrow \xi_1, \xi_5, \text{ or } \xi_{\mu\nu}$ . It is easy to see that the chiral operators are trivial when the taste spurions from the quark-level operator in the action are just the identity. Thus we are left with four new operators which contribute to  $B_K$  with coefficients of  $\mathcal{O}(a_\alpha^2)$ :

$$\mathcal{O}_\chi^{2P} = \text{Str}(\xi_5 \Sigma F_1 \Sigma^\dagger \xi_5 \Sigma F_2 \Sigma^\dagger) + \text{p.c.}, \quad (\text{C26})$$

<sup>32</sup>They also lead to higher-order chiral operators which break  $SO(4)$  taste symmetry, corresponding to the piece of Eq. (C24) with  $\mu = \nu$ .

TABLE III. Mesonic operators corresponding to insertions of four-fermion operators from the staggered action.  $F_1$  and  $F_2$  are flavor off diagonal and come from the  $B_K$  quark-level operator. The remaining taste matrices are flavor diagonal. Repeated indices are summed with the constraint that  $\mu \neq \nu$  in the taste tensor matrices. The notation ‘‘p.c.’’ indicates the parity conjugate of the previous operator. The difference between upper and lower panels is described in the text.

General structure	Specific operator
$\text{Str}(F_L \Sigma F'_R \Sigma^\dagger F_L \Sigma F'_R \Sigma^\dagger) + \text{p.c.}$	$\rightarrow \text{Str}(F_1 F_2) + \text{p.c.}$ $\text{Str}(\xi_5 \Sigma F_1 \Sigma^\dagger \xi_5 \Sigma F_2 \Sigma^\dagger) + \text{p.c.}$ $\text{Str}(\xi_{\mu\nu} \Sigma F_1 \Sigma^\dagger \xi_{\nu\mu} \Sigma F_2 \Sigma^\dagger) + \text{p.c.}$
$\text{Str}(F_L \Sigma F'_R \Sigma^\dagger) \text{Str}(F_L \Sigma F'_R \Sigma^\dagger) + \text{p.c.}$	$\rightarrow \text{Str}(F_1) \text{Str}(F_2) + \text{p.c.}$ $\text{Str}(\xi_5 \Sigma F_1 \Sigma^\dagger) \text{Str}(\xi_5 \Sigma F_2 \Sigma^\dagger) + \text{p.c.}$ $\text{Str}(\xi_{\mu\nu} \Sigma F_1 \Sigma^\dagger) \text{Str}(\xi_5 \Sigma \xi_{\nu\mu} \Sigma^\dagger) + \text{p.c.}$
$\text{Str}(F_L \tilde{F}_L \Sigma^\dagger F_L \Sigma \tilde{F}_R) + \text{p.c.}$	$\rightarrow \text{Str}(F_1 \xi_\mu \Sigma^\dagger F_2 \Sigma \xi_\mu) + \text{p.c.}$ $\text{Str}(F_1 \xi_{\mu 5} \Sigma^\dagger F_2 \Sigma \xi_{5\mu}) + \text{p.c.}$
$\text{Str}(F_L \tilde{F}_L \Sigma^\dagger) \text{Str}(F_L \Sigma \tilde{F}_R) + \text{p.c.}$	$\rightarrow \text{Str}(F_1 \xi_\mu \Sigma^\dagger) \text{Str}(F_2 \Sigma \xi_\mu) + \text{p.c.}$ $\text{Str}(F_1 \xi_{\mu 5} \Sigma^\dagger) \text{Str}(F_2 \Sigma \xi_{5\mu}) + \text{p.c.}$
$\text{Str}(F_L \tilde{F}_L \Sigma^\dagger F_L \tilde{F}_L \Sigma^\dagger) + \text{p.c.}$	$\rightarrow \text{Str}(F_1 \xi_\mu \Sigma^\dagger F_2 \xi_\mu \Sigma^\dagger) + \text{p.c.}$ $\text{Str}(F_1 \xi_{\mu 5} \Sigma^\dagger F_2 \xi_{5\mu} \Sigma^\dagger) + \text{p.c.}$
$\text{Str}(F_L \tilde{F}_L \Sigma^\dagger) \text{Str}(F_L \tilde{F}_L \Sigma^\dagger) + \text{p.c.}$	$\rightarrow \text{Str}(F_1 \xi_\mu \Sigma^\dagger) \text{Str}(F_2 \xi_\mu \Sigma^\dagger) + \text{p.c.}$ $\text{Str}(F_1 \xi_{\mu 5} \Sigma^\dagger) \text{Str}(F_2 \xi_{5\mu} \Sigma^\dagger) + \text{p.c.}$

$$\mathcal{O}_\chi^{3P} = \text{Str}(\xi_5 \Sigma F_1 \Sigma^\dagger) \text{Str}(\xi_5 \Sigma F_2 \Sigma^\dagger) + \text{p.c.}, \quad (\text{C27})$$

$$\mathcal{O}_\chi^{2T} = \sum_{\mu \neq \nu} \text{Str}(\xi_{\mu\nu} \Sigma F_1 \Sigma^\dagger \xi_{\nu\mu} \Sigma F_2 \Sigma^\dagger) + \text{p.c.}, \quad (\text{C28})$$

$$\mathcal{O}_\chi^{3T} = \sum_{\mu \neq \nu} \text{Str}(\xi_{\mu\nu} \Sigma F_1 \Sigma^\dagger) \text{Str}(\xi_{\nu\mu} \Sigma F_2 \Sigma^\dagger) + \text{p.c.} \quad (\text{C29})$$

Operators with spins  $S$  or  $P$  can have tastes  $V$  or  $A$ .<sup>33</sup> Their chiral structure is as follows:

$$\mathcal{O}_F^I = (\overline{Q}_L (1 \otimes \tilde{F}_L) Q_R \pm \overline{Q}_R (1 \otimes \tilde{F}_R) Q_L)^2, \quad (\text{C30})$$

where the upper and lower signs correspond to spin  $S$  and  $P$ , respectively, so their taste spurions must transform as

$$\tilde{F}_L \rightarrow L \tilde{F}_L R^\dagger, \quad \tilde{F}_R \rightarrow R \tilde{F}_R L^\dagger. \quad (\text{C31})$$

As before, there are three possible ways of combining these spurions with those from  $\mathcal{O}_K^{\text{staggered}}$ :

$$\begin{aligned} (F_L F_L)(\tilde{F}_L \tilde{F}_R) + \text{p.c.}, & \quad (F_L F_L)(\tilde{F}_L \tilde{F}_L) + \text{p.c.}, \\ (F_L F_L)(\tilde{F}_R \tilde{F}_R) + \text{p.c.} & \quad (\text{C32}) \end{aligned}$$

These spurion combinations generate the four nontrivial chiral structures shown in the lower panel of Table III. In this case, to produce operators which contribute to  $B_K$  we let  $F_L \rightarrow F_1 \& F_2$  and  $\tilde{F}_L, \tilde{F}_R \rightarrow \xi_\mu$  or  $\xi_{\mu 5}$ . Because two of the taste matrices are diagonal, the following (and similar)

<sup>33</sup>We choose to neglect spin  $T$  in this discussion because it ultimately leads to a subset of operators already generated by spins  $S$  and  $P$ .

relations allow further simplification of the resulting operators:

$$\xi_\mu F_1 \xi_\mu = -F_1, \quad \xi_\mu F_1 = F_{1A}, \quad \xi_{\mu 5} F_1 = F_{1V}. \quad (\text{C33})$$

Such simplification reveals that these operators are none other than  $\mathcal{O}_\chi^{1P}$  and  $\mathcal{O}_\chi^{1V} - \mathcal{O}_\chi^{3A}$ . However, here they have arisen through insertions of the staggered action rather than lattice operator mixing.

In principle, we must also consider insertions of the dimension 4 quark mass term, which can modify  $\mathcal{O}_K^{\text{staggered}}$  in a similar manner to the four-fermion operators already considered. The resulting chiral operators would contain the two taste spurions from  $\mathcal{O}_K^{\text{staggered}}$  and a mass spurion, and thus be of  $\mathcal{O}(m)$  in our power counting. However, it turns out that there are no such operators for quite a simple reason. Recall that, since  $F_1$  and  $F_2$  are flavor off diagonal, they must either be separated by a  $\Sigma$  field on both sides or be in different supertraces, each with a  $\Sigma$ . A single mass spurion is insufficient for meeting this criterion. Therefore operators with mass spurions only contribute to  $B_K$  at NLO through analytic terms, which we discuss next.

### 3. Analytic NLO contributions

We finally consider analytic NLO contributions to the  $B_K$  matrix element,  $\mathcal{M}_K$ . Many operators generate contributions with the same quark mass dependence; it is unnecessary to separate them in fits to lattice data. Thus rather than enumerating all possible operators, we can determine

all linearly independent functions of the quark masses, allowing symmetries to do much of the work.

Given our power-counting scheme, generically six types of operators can contribute to  $\mathcal{M}_K$ :

$$\begin{aligned} \mathcal{O}(p^4), \quad \mathcal{O}("a^2"p^2), \quad \mathcal{O}("a^4"), \quad \mathcal{O}(m^2), \quad \mathcal{O}(p^2m), \\ \mathcal{O}("a^2"m), \end{aligned} \quad (\text{C34})$$

where quotation marks indicate that  $a^2$  can be interchanged with either  $\alpha^2$ ,  $\alpha/4\pi$ , or  $a_\alpha^2$ . We can immediately rule out contributions from  $\mathcal{O}(p^4)$  operators because there are four derivatives but only two fields upon which they can act. We can also rule out NLO analytic contributions from  $\mathcal{O}("a^4")$  operators. Without derivatives or mass matrices the only possible matrix element from such operators is  $\propto "a^4"$ . However, this contribution does not vanish in the chiral limit, and therefore violates the  $U(1)_A$  symmetry possessed by  $\mathcal{O}_K^{\text{staggered}}$ , so it cannot occur. Effectively,  $U(1)_A$  forces all tree-level matrix elements between taste- $\xi_5$  kaons to vanish in the chiral limit.<sup>34</sup>

Determination of the NLO analytic contributions to  $\mathcal{M}_K$  from  $\mathcal{O}("a^2"p^2)$  operators is also straightforward. The derivatives *must* act on the two external kaons, bringing down a factor of  $p_K^2$  which becomes  $m_K^2$  when the kaons are on shell. Therefore the first NLO analytic term in the expression for  $\mathcal{M}_K$  is

$$Am_K^2, \quad (\text{C35})$$

where  $A$  is an undetermined coefficient that is, in principle, suppressed by either  $a^2$ ,  $\alpha^2$ ,  $\alpha/4\pi$ , or  $a_\alpha^2$ . For fitting it is important to know whether or not  $A$  does, in fact, depend upon all four expansion parameters. Operators of  $\mathcal{O}(a^2p^2)$  come from mixing with dimension 8 operators. Recall from Appendix C 2 b that  $[(V - A) \times P]$  mixes with taste  $P$  and taste  $V$  operators at  $\mathcal{O}(a^2)$ . The operator  $\mathcal{O}_\chi^K$  itself comes from a taste  $P$  quark-level operator, so its coefficient,  $C_\chi^K$ , receives a correction of  $\mathcal{O}(a^2)$ . One can also easily map the quark-level operator  $[(V + A) \times P]$  onto a two-derivative chiral operator that contributes to  $\mathcal{M}_K$  at tree level:

$$\text{Str}(\Sigma \partial_\mu \Sigma^\dagger F_1) \text{Str}(\Sigma^\dagger \partial_\mu \Sigma F_2). \quad (\text{C36})$$

Although we lump all NLO analytic contributions to  $\mathcal{M}_K$  of the form  $a^2m_K^2$  together, we note that it *would* be important to separate the correction to  $C_\chi^K$  from the rest if one were to do a fit including some NNLO terms. Operators of  $\mathcal{O}(a^2p^2)$  come from perturbative matching with all tastes of four-fermion operators on the lattice, which we discussed in Appendix C 2 a. As we just showed, mixing with taste  $P$  four-fermion operators leads both to corrections to  $C_\chi^K$  and to a new operator. This time they

produce contributions to  $\mathcal{M}_K$  of the form  $\alpha^2m_K^2$ . In contrast, operators of  $\mathcal{O}(p^2\alpha/4\pi)$  only come from perturbative matching with *wrong-taste* dimension 6 four-fermion operators. Because wrong-taste operators cannot contribute to  $\mathcal{M}_K$  at tree level, however, there is no analytic contribution of the form  $m_K^2\alpha/4\pi$ . Finally, operators of  $\mathcal{O}(p^2\alpha^2)$  come from insertions with operators in the action. The insertion of  $\mathcal{O}_K^{\text{staggered}}$  with the taste-singlet dimension 6 four-fermion operator maps onto  $\mathcal{O}_\chi^K$ , but now at higher order. Thus there is a NLO analytic correction to  $C_\chi^K$ , and therefore a contribution to  $\mathcal{M}_K$ , of the form  $a_\alpha^2m_K^2$ . We summarize the parametric dependence of the coefficient  $A$  in Table V.

Determining the analytic NLO contributions to  $\mathcal{M}_K$  from the last three types of operators requires a bit more work. To do this, we utilize a symmetry of  $\mathcal{O}_K^{\text{staggered}}$  which has not been explicitly needed until now: *CPS symmetry* [31]. Recall that, because  $\mathcal{O}_K^{\text{staggered}}$  is a weak operator, it is *not* parity invariant, as is the strong chiral Lagrangian. However, it *is* invariant under CPS symmetry, and this can be used to restrict the chiral operators to which it corresponds. Standard CPS symmetry is  $CP$  plus  $d \leftrightarrow s$  and  $m_d \leftrightarrow m_s$ . In our case, because there are two sets of valence quarks, we must impose a modified CPS symmetry with both  $d1 \leftrightarrow s1$  and  $d2 \leftrightarrow s2$  so as to make the quark-level operator invariant. At the chiral level, this corresponds to the following transformations:

$$C: \Sigma \leftrightarrow \Sigma^T,$$

$$P: \Sigma \leftrightarrow \Sigma^\dagger,$$

$$"S": d1 \leftrightarrow s1, \quad d2 \leftrightarrow s2,$$

where the last line denotes an interchange of indices in both  $\Sigma$  and  $M$ . One can easily show that all of the operators enumerated in the previous subsections are already invariant under this series of transformations. However, CPS symmetry will prove important for restricting the analytic contributions to  $\mathcal{M}_K$ .

First consider analytic contributions to  $\mathcal{M}_K$  from  $\mathcal{O}(m^2)$  operators, which necessarily contain the two spurions from  $\mathcal{O}_K^{\text{staggered}}$  plus two mass spurions. Because of the *flavor* structure of  $F_1$  and  $F_2$ , such operators will only produce nonvanishing tree-level matrix elements if one mass spurion is within each supertrace, e.g.,

$$\text{Str}(F_1 \Sigma M^\dagger) \text{Str}(F_2 \Sigma M^\dagger) + \text{Str}(F_1 M \Sigma^\dagger) \text{Str}(F_2 M \Sigma^\dagger), \quad (\text{C37})$$

$$\text{Str}(F_1 \Sigma M^\dagger) \text{Str}(F_2 M \Sigma^\dagger) + \text{Str}(F_1 M \Sigma^\dagger) \text{Str}(F_2 \Sigma M^\dagger). \quad (\text{C38})$$

Note that the second term in each operator is the CPS conjugate of the first term. When contracted with an ex-

<sup>34</sup>Of course, this is nothing more than a restatement of the fact that taste- $\xi_5$  mesons are lattice Goldstone bosons, but we have nevertheless checked this result explicitly.

ternal  $\overline{K}_2^0$  and a  $K_1^0$ , the first operator generates an analytic term proportional to  $m_d^2 + m_s^2$ , while the second generates one proportional to  $m_d m_s$ . We choose to combine these terms such that the new analytic contributions to  $\mathcal{M}_K$  at NLO are of the form

$$Bm_K^4, \quad C(m_d - m_s)^2. \quad (\text{C39})$$

Note that the second term vanishes quadratically as  $(m_d - m_s) \rightarrow 0$ .

Next consider  $\mathcal{O}(p^2 m)$  operators which consist of the two spurions from  $\mathcal{O}_K^{\text{staggered}}$ , two Lie derivatives, and a mass spurion. To produce a nonvanishing matrix element, each Lie derivative must be in a separate supertrace with either  $F_1$  or  $F_2$ , e.g.,

$$\text{Str}(F_1 \Sigma \partial_\mu \Sigma^\dagger) \text{Str}(F_2 \Sigma \partial_\mu \Sigma^\dagger) \text{Str}(\Sigma M^\dagger + M \Sigma^\dagger), \quad (\text{C40})$$

$$\begin{aligned} &\text{Str}(F_1 \Sigma \partial_\mu \Sigma^\dagger) \text{Str}(F_2 \Sigma \partial_\mu \Sigma^\dagger \Sigma M^\dagger) \\ &+ \text{Str}(F_1 \Sigma \partial_\mu \Sigma^\dagger) \text{Str}(F_2 M \Sigma^\dagger \Sigma \partial_\mu \Sigma^\dagger). \end{aligned} \quad (\text{C41})$$

Once again, these operators are CPS invariant. The first operator generates a term proportional to  $m_K^2 \text{Str}(M) = m_K^2 \text{Tr}(M_{\text{sea}})$ , while the second generates one proportional to  $m_K^2(m_d + m_s) \propto m_K^4$ . Thus there is only one new analytic contribution to  $\mathcal{M}_K$  at NLO from such operators:

$$Dm_K^2 \text{Tr}(M_{\text{sea}}). \quad (\text{C42})$$

Lastly we consider  $\mathcal{O}("a^2"m)$  operators. These contain two spurions from  $\mathcal{O}_K^{\text{staggered}}$ , possibly two flavor-diagonal taste spurions, and a mass spurion. Thus their tree-level matrix elements are proportional to one power of the quark mass, for which there are three possibilities:  $(m_d + m_s)$ ,  $(m_d - m_s)$ , and  $\text{Tr}(M_{\text{sea}})$ . However,  $"a^2"(m_d - m_s)$  is not CPS invariant, and  $"a^2"(m_d + m_s)$  is included in the parameter  $A$ , so we need only consider analytic terms of the form  $"a^2" \text{Tr}(M_{\text{sea}})$ . The factor  $\text{Tr}(M_{\text{sea}})$  always arises from the operator  $\text{Str}(\Sigma M^\dagger + M \Sigma^\dagger)$  when  $\Sigma = 1$ , so the only operators that can produce tree-level contributions of the form  $"a^2" \text{Tr}(M_{\text{sea}})$  are simply the previously enumerated  $\mathcal{O}(a^2)$ ,  $\mathcal{O}(\alpha^2)$ ,  $\mathcal{O}(\alpha/4\pi)$ , or  $\mathcal{O}(a_\alpha^2)$  operators multiplied by  $\text{Str}(\Sigma M^\dagger + M \Sigma^\dagger)$ . However,  $\mathcal{O}("a^2")$  operators cannot generate tree-level contributions to  $\mathcal{M}_K$  because of the  $U(1)_A$  symmetry, so neither do these particular  $\mathcal{O}("a^2"m)$  ones.

#### 4. Parametric dependence of operator coefficients

Many of the chiral operators enumerated in the previous sections come from multiple quark-level operators. Thus they can have *multiple coefficients*, each multiplied by a different expansion parameter. While this does not affect

TABLE IV. Parametric dependence of NLO  $B_K$  operator coefficients. Each chiral operator can come from more than one quark-level operator, and therefore enter the NLO expression for  $B_K$  multiplied by more than one expansion parameter.

$B_K$ operator	Source—coefficient dependence		
	1-loop mixing?	Dimension 8 op.?	Action insertion?
$\mathcal{O}_\chi^{1P}$	Yes— $\alpha^2$	Yes— $a^2$	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{1T}$	Yes— $\alpha/4\pi, \alpha^2$	No	No
$\mathcal{O}_\chi^{1V}$	Yes— $\alpha/4\pi, \alpha^2$	Yes— $a^2$	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{2V}$	Yes— $\alpha/4\pi, \alpha^2$	Yes— $a^2$	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{3V}$	Yes— $\alpha/4\pi, \alpha^2$	Yes— $a^2$	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{1A}$	Yes— $\alpha/4\pi, \alpha^2$	No	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{2A}$	Yes— $\alpha/4\pi, \alpha^2$	No	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{3A}$	Yes— $\alpha/4\pi, \alpha^2$	No	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{1I}$	Yes— $\alpha/4\pi, \alpha^2$	No	No
$\mathcal{O}_\chi^{2P}$	No	No	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{3P}$	No	No	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{2T}$	No	No	Yes— $a_\alpha^2$
$\mathcal{O}_\chi^{3T}$	No	No	Yes— $a_\alpha^2$

TABLE V. Parametric dependence of NLO analytic contributions to  $B_K$ . An entry of “1” indicates that the coefficient is a constant at this order.

$B_K$ analytic term	Coefficient dependence
$Am_K^2$	$a^2, \alpha^2, a_\alpha^2$
$Bm_K^4$	1
$C(m_d - m_s)^2$	1
$Dm_K^2 \text{Tr}(M_{\text{sea}})$	1

the form of  $B_K$  at NLO in  $S\chi\text{PT}$ , it certainly affects extrapolation of lattice data using multiple lattice spacings. Thus, for ease of use, we summarize the dependence of the NLO operator coefficients on the various expansion parameters in Table IV. All operators which come from 1-loop operator mixing have coefficients of  $\mathcal{O}(\alpha^2)$ ,  $\mathcal{O}(\alpha/4\pi)$ , or both, as described in Appendix C 2 a. Operators which arise through mixing with dimension 8 four-fermion operators come in at  $\mathcal{O}(a^2)$ . Finally, operators that arise from insertions of four-fermion operators are of  $\mathcal{O}(a^2 a^2)$ . Thus the various  $B_K$  operators have anywhere between 1 and 4 independent undetermined coefficients. For example,  $\mathcal{O}_\chi^{1V}$  has four coefficients, multiplied by  $\alpha/4\pi, \alpha^2, a^2$ , and  $a_\alpha^2$ , respectively, whereas  $\mathcal{O}_\chi^{2P}$  has a single coefficient multiplied by  $a_\alpha^2$ .

Finally, for completeness, we include the dependence of the various analytic terms on  $a$  and  $\alpha$  in Table V.

- [1] J. Charles *et al.* (CKMfitter Group), <http://ckmfitter.in2p3.fr>.
- [2] M. Bona *et al.* (UTfit Collaboration), <http://utfit.roma1.infn.it>.
- [3] S. Eidelman *et al.* (Particle Data Group), Phys. Lett. B **592**, 1 (2004).
- [4] E. Gamiz, S. Collins, C. T. H. Davies, J. Shigemitsu, and M. Wingate (HPQCD Collaboration), Nucl. Phys. B, Proc. Suppl. **140**, 353 (2005).
- [5] W. Lee (private communication).
- [6] C. Dawson (RBC Collaboration), Nucl. Phys. B, Proc. Suppl. **140**, 356 (2005).
- [7] C. Aubin *et al.* (MILC Collaboration), Phys. Rev. D **70**, 114501 (2004).
- [8] C. Aubin and C. Bernard, Phys. Rev. D **68**, 034014 (2003).
- [9] C. Aubin and C. Bernard, Phys. Rev. D **68**, 074011 (2003).
- [10] S. R. Sharpe and R. S. Van de Water, Phys. Rev. D **71**, 114505 (2005).
- [11] C. Aubin and C. Bernard, Nucl. Phys. B, Proc. Suppl. **140**, 491 (2005).
- [12] G. Kilcup, R. Gupta, and S. R. Sharpe, Phys. Rev. D **57**, 1654 (1998).
- [13] D. Versteegen, Nucl. Phys. **B249**, 685 (1985).
- [14] W. Lee and S. Sharpe, Phys. Rev. D **68**, 054510 (2003).
- [15] Q. Mason, P. Lepage, P. Mackenzie, H. Trottier, J. Hein, C. Davies, and E. Follana (HPQCD Collaboration), Nucl. Phys. B, Proc. Suppl. **119**, 446 (2003).
- [16] C. Bernard (private communication).
- [17] S. R. Sharpe and R. S. Van de Water, Phys. Rev. D **69**, 054027 (2004).
- [18] J. Bijnens, H. Sonoda, and M. B. Wise, Phys. Rev. Lett. **53**, 2367 (1984).
- [19] S. R. Sharpe, Phys. Rev. D **46**, 3146 (1992).
- [20] S. R. Sharpe and N. Shoresh, Phys. Rev. D **62**, 094503 (2000).
- [21] X. D. Ji, hep-lat/9506034.
- [22] R. Gupta, T. Bhattacharya, and S. R. Sharpe, Phys. Rev. D **55**, 4036 (1997).
- [23] M. Crisafulli *et al.*, Phys. Lett. B **369**, 325 (1996).
- [24] M. F. L. Golterman and K. C. L. Leung, Phys. Rev. D **57**, 5703 (1998).
- [25] D. Becirevic and G. Villadoro, Phys. Rev. D **69**, 054010 (2004).
- [26] W. Lee and S. Sharpe, Phys. Rev. D **60**, 114503 (1999).
- [27] S. Sharpe and A. Patel, Nucl. Phys. **B417**, 307 (1994).
- [28] S. R. Sharpe, Nucl. Phys. B, Proc. Suppl. **34**, 403 (1994).
- [29] D. Daniel and S. Sheard, Nucl. Phys. **B302**, 471 (1988).
- [30] A. Patel and S. Sharpe, Nucl. Phys. **B395**, 701 (1993).
- [31] C. W. Bernard, T. Draper, A. Soni, H. D. Politzer, and M. B. Wise, Phys. Rev. D **32**, 2343 (1985).