# Neutral *B*-meson mixing from three-flavor lattice quantum chromodynamics: Determination of the $SU(3)$-breaking ratio $\xi$

A. Bazavov,[1] C. Bernard,[2] C. M. Bouchard,[3,4,5] C. DeTar,[6] M. Di Pierro,[7] A. X. El-Khadra,[3] R. T. Evans,[3,8]
E. D. Freeland,[3,9] E. Gámiz,[4,10,*] Steven Gottlieb,[11] U. M. Heller,[12] J. E. Hetrick,[13] R. Jain,[3] A. S. Kronfeld,[4] J. Laiho,[14]
L. Levkova,[6] P. B. Mackenzie,[4] E. T. Neil,[4] M. B. Oktay,[6] J. N. Simone,[4] R. Sugar,[15]
D. Toussaint,[16] and R. S. Van de Water[1]

(Fermilab Lattice and MILC Collaborations)

[1]*Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA*
[2]*Department of Physics, Washington University, St. Louis, Missouri 63130, USA*
[3]*Physics Department, University of Illinois, Urbana, Illinois 61801, USA*
[4]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[5]*Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA*
[6]*Physics Department, University of Utah, Salt Lake City, Utah 84112, USA*
[7]*School of Computing, DePaul University, Chicago, Illinois 60604, USA*
[8]*Department of Nuclear Engineering, North Carolina State University, Raleigh, North Carolina 27695, USA*
[9]*Department of Physics, Benedictine University, Lisle, Illinois 60532, USA*
[10]*CAFPE and Departamento de Física Teórica y del Cosmos, Universidad de Granada, Granada, Spain*
[11]*Department of Physics, Indiana University, Bloomington, Indiana 47405, USA*
[12]*American Physical Society, Ridge, New York 11961, USA*
[13]*Physics Department, University of the Pacific, Stockton, California 95211, USA*
[14]*SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, Scotland, United Kingdom*
[15]*Department of Physics, University of California, Santa Barbara, California 93106, USA*
[16]*Department of Physics, University of Arizona, Tucson, Arizona 85721, USA*

We study $SU(3)$-breaking effects in the neutral $B_d$-$\bar{B}_d$ and $B_s$-$\bar{B}_s$ systems with unquenched $N_f = 2 + 1$ lattice quantum chromodynamics (QCD). We calculate the relevant matrix elements on the MILC collaboration's gauge configurations with asqtad-improved staggered sea quarks. For the valence light-quarks ($u$, $d$, and $s$) we use the asqtad action, while for $b$ quarks we use the Fermilab action. We obtain $\xi = f_{B_s}\sqrt{B_{B_s}}/f_{B_d}\sqrt{B_{B_d}} = 1.268 \pm 0.063$. We also present results for the ratio of bag parameters $B_{B_s}/B_{B_d}$ and the ratio of Cabibbo-Kobayashi-Maskawa matrix elements $|V_{td}|/|V_{ts}|$. Although we focus on the calculation of $\xi$, the strategy and techniques described here will be employed in future extended studies of the $B$ mixing parameters $\Delta M_{d,s}$ and $\Delta\Gamma_{d,s}$ in the standard model and beyond.

## I. INTRODUCTION

The observation of new particles at high-energy colliders is not the only way for new physics to be discovered. It can also be unveiled through the observation of deviations from the standard model (SM) via high-precision measurements of low-energy observables in high-luminosity experiments. This requires matching precision in the theoretical SM predictions for these observables. In principle, such a comparison could reveal the exchange of virtual, new heavy particles involving scales much higher than those that can be achieved in direct production at high-energy colliders.

Heavy-flavor physics and, in particular, neutral-meson mixing are potentially very sensitive to these virtual effects. Neutral-meson mixing occurs at loop level in the SM, see Fig. 1, and it is further suppressed by small Cabibbo-Kobayashi-Maskawa (CKM) matrix elements, so the effect of new particles in the internal loops could be noticeable in the parameters describing the mixing. Indeed, there are several measurements for which there is a $2 - 3\sigma$ difference from the SM prediction. These include $\sin(2\beta)$ [1], the like-sign dimuon charge asymmetry [2], and unitarity triangle (UT) fits [3–7]. It has been argued that these differences may be due to physics beyond the standard model (BSM) affecting the neutral *B*-meson mixing processes [3,4].

In the $B_s^0$ system, the relative phase between the decay amplitudes with and without mixing, $\beta_s$, could also show BSM effects, as pointed out in Ref. [8] and later hinted at in a Tevatron measurement [9]. Although new measurements at the CDF [10] and D0 detector [11] are in better
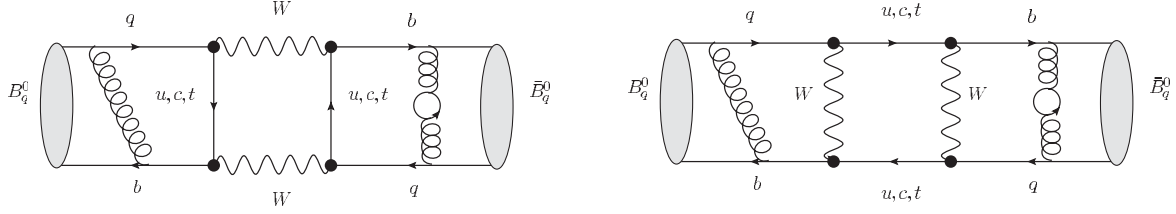
*megamiz@ugr.es

FIG. 1.   Box diagrams contributing to $B^0 - \bar{B}^0$ mixing in the SM. Gluon exchanges shown in the plot are just representative of the QCD corrections.

agreement with the SM, reducing the difference from $\sim 3\sigma$ to $\sim 1\sigma$, there is still room for a large deviation of $\beta_s$ from SM values.

The main parameters describing mixing in the $B_s^0$ and the $B_d^0$ systems are the mass differences, $\Delta M_{s(d)}$, and the decay width differences, $\Delta\Gamma_{s(d)}$, between the heavy and light $B_{s(d)}^0$ mass eigenstates, and the *CP* violating phases $\phi_{s(d)}$. The phases $\phi_{s(d)}$ are defined as the argument of the ratio of the dispersive and absorptive off-diagonal elements of the time evolution matrix which describes the mixing [12]. The existence of new, heavy particles in loops could affect the value of the mass differences, given by the dispersive part of the time evolution matrix. The mass differences $\Delta M_s$ [13–15] and $\Delta M_d$ [16] have been measured with an accuracy better than 1%. Improving the theoretical control on these quantities is thus crucial in order to fully exploit the potential of *CP* violating observables to search for nonstandard physics. In addition, the theoretical calculation of BSM contributions to mixing and the experimental measurement of $B^0$ mixing parameters can help in constraining BSM parameters and understanding new physics [6]. Several recent studies have addressed that task [3–7,17–24], finding that one of the main limitations to further constraining the parameter space in BSM theories is the error associated with the theoretical calculation of the nonperturbative inputs.

The most interesting quantity to analyze in $B^0$ mixing phenomena is the $SU(3)$-breaking ratio $\xi$, which measures the difference between the mixing parameters in the $B_s^0$ and the $B_d^0$ systems, and enters the relation between the ratio of mass differences and CKM matrix elements as

$$\left| \frac{V_{td}}{V_{ts}} \right| = \xi \sqrt{\frac{\Delta M_d M_{B_s}}{\Delta M_s M_{B_d}}}. \tag{1.1}$$

Its value, together with the experimental measurement of the mass differences $\Delta M_{s,d}$, determines the ratio of CKM matrix elements $|V_{td}/V_{ts}|$, which constrains one side of the unitarity triangle [25,26]. Thus, $\xi$ is one of the key ingredients in UT analyses [3,5–7].

In the SM, mixing is due to box diagrams with the exchange of two $W$-bosons, like those in Fig. 1. These box diagrams can be rewritten in terms of an effective Hamiltonian with four-fermion operators describing processes with $\Delta B = 2$. In BSM theories, mixing processes

can receive contributions from additional diagrams due to the exchange of new, heavy particles. These can also be parametrized in terms of four-fermion effective operators built with SM degrees of freedom. The most general effective Hamiltonian describing processes with $\Delta B = 2$ was given in Refs. [27,28], and can also be found in Ref. [29]. There are a total of five independent operators (plus parity conjugates) in the Hamiltonian, but only three of them contribute to mixing in the SM

$$\mathcal{H}_{\text{eff,SM}}^{\Delta B=2} = \sum_{i=1}^{3} C_i \mathcal{O}_i, \tag{1.2}$$

with

$$\begin{aligned} \mathcal{O}_1^q &= (\bar{q}^i \gamma^\nu L b^i)(\bar{q}^j \gamma^\nu L b^j), \\ \mathcal{O}_2^q &= (\bar{q}^i L b^i)(\bar{q}^j L b^j), \\ \mathcal{O}_3^q &= (\bar{q}^i L b^j)(\bar{q}^j L b^i), \end{aligned} \tag{1.3}$$

where $i$ and $j$ are color indices, and $L$ and $R$ are the Dirac projection operators $\frac{1}{2}(1 - \gamma_5)$ and $\frac{1}{2}(1 + \gamma_5)$, respectively. The fields $q$ denote strange or down fields for $B_s^0$ and $B_d^0$ mixing, respectively, and $b$ represents the bottom field.

The matrix element of the first operator in Eq. (1.3), $\mathcal{O}_1^q$, provides the mass difference in the SM:

$$\Delta M_q^{\text{SM}} = \frac{G_F^2 M_W^2}{6\pi^2} |V_{tq}^* V_{tb}|^2 \eta_2^B S_0(x_t) M_{B_q} f_{B_q}^2 \hat{B}_{B_q}, \tag{1.4}$$

where $S_0(x_t)$ is the Inami-Lim function [30], which depends on the top quark mass through $x_t = m_t^2/M_W^2$, and the quantity $\eta_2^B$ is a perturbative quantum chromodynamics correction factor. The products $f_{B_q}^2 \hat{B}_{B_q}$ parametrize the hadronic matrix elements in the effective theory by

$$\langle \bar{B}_q^0 | \mathcal{O}_1^q | B_q^0 \rangle(\mu) = \frac{2}{3} M_{B_q}^2 f_{B_q}^2 B_{B_q}(\mu). \tag{1.5}$$

The factors $f_{B_q}$ are the $B_q^0$ decay constants. The renormalization group invariant bag parameters $\hat{B}_{B_q}$ in Eq. (1.4) are related to the scheme and scale dependent bag parameters in Eq. (1.5) at next-to-leading order (NLO) by

$$\hat{B}_{B_q} = [\alpha_s(\mu)]^{-6/23} \left[ 1 + \frac{\alpha_s(\mu)}{4\pi} J_5 \right] B_{B_q}(\mu), \tag{1.6}$$

where $J_5$ is known in both $\overline{\text{MS}}$-NDR (naive dimensional regularization) and $\overline{\text{MS}}$-HV ('t Hooft-Veltman) schemes [31]. Bag parameters have traditionally been used to measure the deviation of the four-fermion operator matrix elements from their vacuum insertion values, $B_B = 1$.

The $SU(3)$-breaking parameter $\xi$ can be written in terms of decay constants and bag parameters

$$\xi = \frac{f_{B_s}\sqrt{B_{B_s}}}{f_{B_d}\sqrt{B_{B_d}}}. \tag{1.7}$$

Many of the uncertainties that affect the theoretical calculation of the decay constants and bag parameters cancel totally or partially in this ratio, leaving the chiral extrapolation as the dominant error. Hence, $\xi$ and the combination of CKM matrix elements related to it, can be determined with a significantly smaller error than the individual matrix elements.

The hadronic matrix elements in Eq. (1.5) encode the nonperturbative physics of the problem and are best calculated using lattice QCD. Our current knowledge of them, limits the accuracy with which the CKM matrix elements appearing in Eq. (1.4) can be determined from the experimental measurements of $\Delta M_{s(d)}$. In particular, the uncertainty associated with the calculation of $\xi$ is one of the main limiting factors in UT analyses, so improvement in the knowledge of $\xi$ is crucial to disentangle the origin of the $2 - 3\sigma$ tension.

There are two $2 + 1$ unquenched lattice calculations of the ratio $\xi$ in the literature. One is by the HPQCD collaboration [32], which quotes the value $\xi = 1.258(33)$. The other is an exploratory study by the RBC and UKQCD collaborations [33] on a single lattice spacing and using the static limit for the bottom quark; their result is $\xi = 1.13(12)$. In this paper, we report a lattice calculation of $\xi$ at the few-percent level.

Preliminary results related to the work here were presented in Refs. [34–37]. In Ref. [34], the simulation and correlator fitting methods were described using data for one lattice spacing, while Refs. [35,36] focused on the discussion of statistical and fitting errors, and the chiral extrapolation method. In Ref. [37], we studied the matching method and the heavy-quark discretization errors.

The primary difference between this work and the HPQCD calculation in Ref. [32] is the treatment of the valence *b* quarks. The HPQCD collaboration uses lattice NRQCD [38] while we employ the clover action [39] with the Fermilab interpretation [40]. An advantage of the Fermilab method is that it can also be efficiently used to simulate charm quarks, so the analysis performed in this work can be easily extended to the study of the short-distance contributions to $D^0$-$\bar{D}^0$ mixing. Although in the case of neutral $D$ mixing the long-distance contributions are believed to be dominant, a calculation of the short-distance contributions can, nevertheless, provide valuable constraints on extensions of the SM [41].

In order to achieve the few-percent level of precision required by phenomenology, we use lattice QCD simulations with realistic sea quarks. In particular, we employ a subset of the magnetic induced laser cooling (MILC) configurations with $2 + 1$ flavors of asqtad sea quarks [42–44]. In the valence sector, we use the same staggered asqtad action to simulate the light quarks. The configurations we use in this analysis were generated using the fourth-root procedure for eliminating extra degrees of freedom originating from fermion doubling. Despite the nonlocal violations of unitarity of the rooted theory at non-zero lattice spacing [45,46], there are strong theoretical arguments [47–50], as well as other analytical and numerical evidence [51–54], that the local, unitary theory of QCD is recovered in the continuum limit. This gives us confidence that the rooting procedure yields valid results. We also explicitly tested the rooting procedure as well as improvements in our heavy action by calculating the spin-dependent hyperfine splittings for $B_s$ and $D_s$ mesons in Ref. [55].

Our collaboration has already successfully used the asqtad MILC ensembles in similar calculations of other quantities involving $B$ mesons, as part of a broad program of calculating matrix elements: for example, the extraction of the CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$ from the calculation of, respectively, the semileptonic form factors describing the processes $B \to \pi l\nu$ [56] and $B \to D^* l\nu$ [57,58]; or, more recently, the calculation of the $f_B$ and $f_{B_s}$ decay constants [59] and the form factor ratios between the semileptonic decays $\bar{B} \to D^+ l^- \bar{\nu}$ and $\bar{B}_s \to D_s^+ l^- \bar{\nu}$ [60].

This paper is organized as follows. In Sec. II, we describe the actions and parameters used in our numerical simulations, as well as the construction of the mixing operators and correlation functions. Section III presents the renormalization method using one-loop mean field improved lattice perturbation theory. We include a discussion of the errors associated with the matching and numerical values of the matching coefficients used. Next, in Sec. IV, we give the details of the procedure for the correlator fits. Section V is devoted to the chiral-continuum extrapolation, which is performed within the framework of rooted staggered chiral perturbation theory [61–65]. We describe and discuss the choice of the functional form used in the extrapolation, the different fitting methods tested, and the choice of parameters and parametrization. In Sec. VI, we list and estimate the different systematic errors. Finally, Sec. VII compiles our final results for the parameter $\xi$ as well as for $|V_{td}|/|V_{ts}|$, and the ratio of bag parameters $B_{B_s}/B_{B_d}$. We also discuss planned future improvements in the calculation of $B^0$ mixing parameters by our collaboration. In Appendix A, we provide the explicit formulas for the chiral fit functions used in the chiral fits described in Sec. V. In Appendix B, we compile the functions needed to estimate the heavy-quark discretization errors in our calculation. Finally, Appendix C

TABLE I. Parameters of the ensembles analyzed in this work. The first two rows show the approximate lattice spacing and the volume. $am_l$ and $am_h$ are the light and strange sea quark masses, respectively. $N_{confs}$ is the number of configurations analyzed from each ensemble, and $am_q$ are the light valence quark masses. The $r_1/a$ values are obtained by fitting the calculated $r_1/a$ to a smooth function [72], as explained in Ref. [66].

| $\approx a$ (fm) | $(\frac{L}{a})^3 \times \frac{T}{a}$ | $am_l/am_h$ | $N_{confs}$ | $am_q$ | $r_1/a$ |
|---|---|---|---|---|---|
| 0.12 | $24^3 \times 64$ | 0.005/0.05 | 529 | 0.005, 0.007, 0.01, 0.02, 0.03, 0.0415 | 2.64 |
| 0.12 | $20^3 \times 64$ | 0.007/0.05 | 833 | 0.005, 0.007, 0.01, 0.02, 0.03, 0.0415 | 2.63 |
| 0.12 | $20^3 \times 64$ | 0.01/0.05 | 592 | 0.005, 0.007, 0.01, 0.02, 0.03, 0.0415 | 2.62 |
| 0.12 | $20^3 \times 64$ | 0.02/0.05 | 460 | 0.005, 0.007, 0.01, 0.02, 0.03, 0.0415 | 2.65 |
| 0.09 | $28^3 \times 96$ | 0.0062/0.031 | 557 | 0.0031, 0.0044, 0.062, 0.0124, 0.0272, 0.031 | 3.70 |
| 0.09 | $28^3 \times 96$ | 0.0124/0.031 | 534 | 0.0031, 0.0042, 0.062, 0.0124, 0.0272, 0.031 | 3.72 |

discusses our choices for prior central values and widths for the correlator fits.

## II. NUMERICAL SIMULATIONS

### A. Parameters of the simulations

The $n_f = 2 + 1$ MILC ensembles [66] used in our calculation include the effect of three sea-quark flavors: two degenerate light quarks corresponding to the up and down quarks (although with larger masses than the physical ones), and one heavier quark corresponding to the strange quark. These dynamical quarks are simulated using the asqtad improved staggered action with errors starting at $O(\alpha_s a^2)$ [67]. The gluon action is a Symanzik improved and tadpole improved action, with $O(\alpha_s a^2)$ errors coming from the gluon loops removed [68,69]. The couplings needed to remove the $O(\alpha_s a^2)$ errors coming from quark loops [70] were available only after the generation of configurations was well advanced, so these effects are not accounted for in the MILC ensembles. Thus, the dominant errors in the gauge action are also of $O(a^4, \alpha_s a^2)$.

The valence light-quark propagators are generated using the asqtad action and converted to naive quark propagators using the relation [71]

$$S_{\text{naive}}(x, y) = \Omega(x)\Omega^\dagger(y)S_{\text{staggered}}(x, y), \quad (2.1)$$

where $\Omega(x) = \gamma_0^{x_0}\gamma_1^{x_1}\gamma_2^{x_2}\gamma_3^{x_3}$.

For the heavy bottom quarks we use the Sheikholeslami-Wohlert action [39] with the Fermilab interpretation for heavy-quark systems [40]. This interpretation retains the full mass dependence of the theory within the parameters of the lattice action. A tree-level matching to QCD is then performed via heavy quark effective theory (HQET), after which it can be shown that the errors in the action begin at $O(\alpha_s \Lambda_{\text{QCD}} a, \Lambda_{\text{QCD}}^2 a^2)$ times bounded functions of $m_b a$, the $b$-quark mass in lattice units.

We perform our analysis at two different values of the lattice spacing, $a \approx 0.12, 0.09$ fm, and for a variety of sea-quark masses. The values used are shown in Table I. The mass of the heavy $b$ quark is fixed to its physical value by computing the spin-averaged $B_s$ kinetic mass [55]. This determines the $b$ quark's hopping parameters, $\kappa_b = 0.0860$

for the $a \approx 0.12$ fm lattice and $\kappa_b = 0.0923$ for the $a \approx 0.09$ fm lattice [55], and thus the bare $b$ quark mass. We simulate the $B$ mesons with the six different values of light-valence quark mass listed in Table I, the smallest of which is around $m_s/8$, in order to facilitate the extrapolation/ interpolation to the physical down/strange quark masses.

### B. Correlators: The open-meson propagator

As described in the introduction, the study of the $SU(3)$-breaking ratio $\xi$ requires the calculation of the hadronic matrix elements[1] $\langle \mathcal{O}_1^q \rangle$ and $\langle \mathcal{O}_2^q \rangle$, the latter of which mixes with $\langle \mathcal{O}_1^q \rangle$ under renormalization, for both $q = d$, $s$. The matrix elements are obtained from three-point correlation functions with zero spatial momentum

$$C_{\mathcal{O}_i^q}(t_x, t_y) = \sum_{x,y} \langle \bar{B}_q^0(t_y, y)\mathcal{O}_i^q(0)B_q^0(t_x, x)^\dagger \rangle, \quad (2.2)$$

where $i = 1$ or 2, and the $B$-meson creation operator $B_q^0(t, x)^\dagger = \sum_{x'} \bar{b}(t, x')S(x, x')\gamma_5 q(t, x)$, with $q(t, x)$ the naive light quark field, whose propagator is constructed from the staggered propagator *via* Eq. (2.1), and with $S(x, x')$ a smearing function. Our choice of smearing function is discussed in Sec. . The structure of the functions in Eq. (2.2) is depicted in Fig. 2. The four-fermion operators $\mathcal{O}_i^q$ are placed at the origin while $B$-mesons are positioned at $x$ and $y$. This layout allows us to perform the three-point function fits over both $t_x$ and $t_y$, maximizing the information included in the fits. In order to extract the relevant matrix elements from Eq. (2.2), we need to determine the overlap of the $B$-meson creation operator with the ground state. Therefore, we also need the pseudoscalar two-point correlator with zero spatial momentum

$$C_{\text{PS}}(t) = \sum_x \langle B_q^0(t, x)B_q^0(0, 0)^\dagger \rangle. \quad (2.3)$$

The calculation of both three-point and two-point correlators can be organized into convenient structures. Starting with a general correlator with Dirac structure $\Gamma_1 \times \Gamma_2$,

---

[1]To simplify the notation, we define $\langle \mathcal{O}_i^q \rangle \equiv \langle \bar{B}_q^0 | \mathcal{O}_i^q | B_q^0 \rangle$ for $i = 1, 2$.
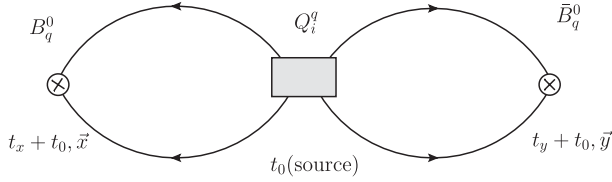
FIG. 2. Structure of the three-point correlators. A $B_q^0$ is created at rest at $t_x + t_0 < t_0$. At time $t_0$, it oscillates into a $\bar{B}_q^0$ via the operator $\mathcal{O}_i^q$, which is subsequently annihilated at $t_y + t_0 > t_0$.

which accommodates a full set of $\Delta B = 2$ four-quark operators, including those in Eq. (1.3),

$$C_3(t_x, t_y) = \sum_{x,y} \langle \bar{B}_q^0(t_y, \boldsymbol{y}) \bar{q}(0)\Gamma_1 b(0) \bar{q}(0)\Gamma_2 b(0) B_q^0(t_x, \boldsymbol{x})^\dagger \rangle,$$

(2.4)

and performing the four possible Wick contractions, we obtain

$$
\begin{aligned}
C_3(t_x, t_y) \\
= \sum_{x,y} & \{ \mathrm{tr}[\gamma_5 L_q(x, 0)\Gamma_1 H_b(0, x)]\mathrm{tr}[\gamma_5 L_q(y, 0)\Gamma_2 H_b(0, y)] \\
& + \mathrm{tr}[\gamma_5 L_q(y, 0)\Gamma_1 H_b(0, y)]\mathrm{tr}[\gamma_5 L_q(x, 0)\Gamma_2 H_b(0, x)] \\
& - \mathrm{tr}[\gamma_5 L_q(x, 0)\Gamma_1 H_b(0, y)\gamma_5 L_q(y, 0)\Gamma_2 H_b(0, x)] \\
& - \mathrm{tr}[\gamma_5 L_q(x, 0)\Gamma_2 H_b(0, y)\gamma_5 L_q(y, 0)\Gamma_1 H_b(0, x)] \}
\end{aligned}
$$

(2.5)

$$
\begin{aligned}
= \sum_{x,y} & \{ \mathrm{tr}[L_q(x, 0)\Gamma_1 \gamma_5 H_b^\dagger(x, 0)]\mathrm{tr}[L_q(y, 0)\Gamma_2 \gamma_5 H_b^\dagger(y, 0)] \\
& + \mathrm{tr}[L_q(y, 0)\Gamma_1 \gamma_5 H_b^\dagger(y, 0)]\mathrm{tr}[L_q(x, 0)\Gamma_2 \gamma_5 H_b^\dagger(x, 0)]
\end{aligned}
$$

(2.6)

$$
\begin{aligned}
& - \mathrm{tr}[L_q(x, 0)\Gamma_1 \gamma_5 H_b^\dagger(y, 0)L_q(y, 0)\Gamma_2 \gamma_5 H_b^\dagger(x, 0)] \\
& - \mathrm{tr}[L_q(x, 0)\Gamma_2 \gamma_5 H_b^\dagger(y, 0)L_q(y, 0)\Gamma_1 \gamma_5 H_b^\dagger(x, 0)] \},
\end{aligned}
$$

(2.7)

where $L_q$ is the (naive) light-quark propagator, and $H_b$ is the heavy-quark propagator. The traces in Eq. (2.5) run over spin and color indices.

These correlators can be rewritten as

$$
\begin{aligned}
C_3(t_x, t_y) = & \Gamma_1^{\beta\alpha} E_{aa}^{\alpha\beta}(t_x)\Gamma_2^{\tau\sigma} E_{cc}^{\sigma\tau}(t_y) \\
& + \Gamma_1^{\beta\alpha} E_{aa}^{\alpha\beta}(t_y)\Gamma_2^{\tau\sigma} E_{cc}^{\sigma\tau}(t_x) \\
& - \Gamma_1^{\beta\alpha} E_{ac}^{\alpha\sigma}(t_x)\Gamma_2^{\sigma\tau} E_{ca}^{\tau\beta}(t_y) \\
& - \Gamma_1^{\beta\alpha} E_{ac}^{\alpha\sigma}(t_y)\Gamma_2^{\sigma\tau} E_{ca}^{\tau\beta}(t_x),
\end{aligned}
$$

(2.8)

where summation over repeated indices is implied, and we have introduced the basic objects

$$E_{ac}^{\alpha\beta}(t) = \gamma_5^{\alpha\sigma} H_{b,da}^{*\tau\sigma}(t, 0)L_{q,dc}^{\tau\beta}(t, 0),$$

(2.9)

with Dirac indices labeled as $\alpha$, $\beta$, $\sigma$, $\tau$ and color indices labeled as $a$, $c$, $d$. We call the combination of propagators $E_{ad}^{\alpha\beta}(t_x)$ defined in Eq. (2.9) "open-meson propagator". Once the open-meson propagators have been computed and saved, all correlation functions needed for $B$-meson mixing, including BSM operators, can be immediately constructed by contracting them with the appropriate Dirac structures. As shown in Eq. (2.8), the three-point correlators are obtained by combining two open-meson propagators, while for the two-point correlators we only need one open-meson propagator.

### C. Doubler modes' effect on the correlation functions

The remnant doubling degeneracy of staggered fermions leads to contributions of scalar states, in addition to pseudoscalar states, in correlation functions with external pseudoscalar particles. The scalar contamination yields oscillating terms in the correlation functions [71]. In this section, we extend the analysis of Ref. [71] for two-point correlation functions, to the three-point functions introduced in Sec. II B. We conclude that the effect of the doubler modes on the three-point functions can be removed at leading order in the lattice spacing through appropriate fits of the Euclidean time-dependence.

The *doubling* symmetry of the original naive action under the transformation

$$\psi(x) \rightarrow e^{ix\cdot\pi_g}M_g\psi(x), \qquad \bar{\psi}(x) \rightarrow e^{ix\cdot\pi_g}\bar{\psi}(x)M_g^\dagger,$$

(2.10)

where

$$M_g = \prod_{\mu\in g} i\gamma_5\gamma_\mu,$$

(2.11)

$$G = \{g : g = (\mu_1, \mu_2, \ldots), \mu_1 < \mu_2 < \ldots\},$$

(2.12)

$$(\pi_g)_\mu = \begin{cases} \frac{\pi}{a} & \text{if } \mu \in g \\ 0 & \text{otherwise} \end{cases}$$

(2.13)

generates sixteen equivalent species of quarks, referred to as *tastes*, that can be reduced to four by staggering the quark field [73]. Each element of $G$ is a list of up to four indices, e.g., (2), (0,3), and (0,1,2,3) are elements of $G$, as is the empty set $\varnothing$. Different $g$'s label different doubler modes, or tastes.

Consider the general three-point function in momentum space,

$C_{\Gamma_1 \times \Gamma_2}(t_x, t_y)$

$\equiv \sum_{x,y} \langle \bar{b}(x) \gamma_5 q(x) [\bar{q}(0) \Gamma_1 b(0) \bar{q}(0) \Gamma_2 b(0)] \bar{b}(y) \gamma_5 q(y) \rangle$

$= \int_{-\pi/a}^{\pi/a} \dfrac{d^3 \boldsymbol{p}}{(2\pi)^3} \dfrac{d^3 \boldsymbol{k}}{(2\pi)^3} \langle \tilde{\bar{b}}(\boldsymbol{p}, t_x) \gamma_5 \tilde{q}(\boldsymbol{p}, t_x)$

$\times [\bar{q}(0) \Gamma_1 b(0) \bar{q}(0) \Gamma_2 b(0)] \tilde{\bar{b}}(\boldsymbol{k}, t_y) \gamma_5 \tilde{q}(\boldsymbol{k}, t_y) \rangle, \quad (2.14)$

where $\Gamma_1 \times \Gamma_2$ denotes the Dirac structure of the four-fermion operators in Eq. (1.3). For simplicity of notation, we omit the smearing function from the $B$-meson operator and write it as $\bar{b}(x) \gamma_5 q(x)$. It would be straightforward (but not particularly instructive) to generalize the expressions of Eqs. (2.14), (2.15), (2.16), (2.17), (2.18), (2.19), and (2.20) to include the smearing function.

For now the bracketed four-quark operator is left in position space and $\tilde{b}$, $\tilde{q}$ are the spatial momentum-space bottom and strange/down fermion fields. Because of the

doubling symmetry, we can integrate over the central half of the Brillouin zone and sum over the spatial doublers

$C_{\Gamma_1 \times \Gamma_2}(t_x, t_y) = \sum_{g_s, g_s'} \int_{-\pi/2a}^{\pi/2a} \dfrac{d^3 \boldsymbol{p}}{(2\pi)^3} \dfrac{d^3 \boldsymbol{k}}{(2\pi)^3}$

$\times \langle \tilde{\bar{b}}(\boldsymbol{p} + \boldsymbol{\pi}_{g_s}, t_x) \gamma_5 \tilde{q}(\boldsymbol{p} + \boldsymbol{\pi}_{g_s}, t_x)$

$\times [\bar{q}(0) \Gamma_1 b(0) \bar{q}(0) \Gamma_2 b(0)]$

$\times \tilde{\bar{b}}(\boldsymbol{k} + \boldsymbol{\pi}_{g_s'}, t_y) \gamma_5 \tilde{q}(\boldsymbol{k} + \boldsymbol{\pi}_{g_s'}, t_y) \rangle, \quad (2.15)$

where $g_s$ denotes a particular spatial doubler mode. Due to the high momentum that is imparted to the heavy quark when $g_s \neq \varnothing$, such states are far off shell and have a negligible effect on the correlation function. The taste of the temporal modes can now be considered by Fourier transforming the light quarks' temporal component, and then again restricting the Brillouin zone and summing over the doublers

$C_{\Gamma_1 \times \Gamma_2}(t_x, t_y) = \int_{-\pi/2a}^{\pi/2a} \dfrac{d^3 \boldsymbol{p}}{(2\pi)^3} \dfrac{d^3 \boldsymbol{k}}{(2\pi)^3} \int_{-\pi/2a}^{\pi/2a} \dfrac{dp_0}{(2\pi)} \dfrac{dk_0}{(2\pi)} e^{ip_0 t_x + ik_0 t_y} \langle \tilde{\bar{b}}(\boldsymbol{p}, t_x) \gamma_5 [\tilde{q}'(\boldsymbol{p}, p_0) + (-1)^{t_x} \tilde{q}'(\boldsymbol{p}, p_0 + \pi/a)]$

$\times [\bar{q}(0) \Gamma_1 b(0) \bar{q}(0) \Gamma_2 b(0)] \tilde{\bar{b}}(\boldsymbol{k}, t_y) \gamma_5 [\tilde{q}'(\boldsymbol{k}, k_0) + (-1)^{t_y} \tilde{q}'(\boldsymbol{k}, k_0 + \pi/a)] \rangle. \quad (2.16)$

With the momentum space spinors $\tilde{f}'^g$ defined as

$$\tilde{f}'^g(p) = \prod_{\mu \in g} i\gamma_5 \gamma_\mu \tilde{q}'(p + \pi_g), \quad (2.17)$$

so that

$$\tilde{q}'(\boldsymbol{p}, p_0) = \tilde{f}'(\boldsymbol{p}, p_0), \qquad \tilde{q}'(\boldsymbol{p}, p_0 + \pi/a) = i\gamma_5 \gamma_0 \tilde{f}'^0(\boldsymbol{p}, p_0), \quad (2.18)$$

the three-point function can be written as

$C_{\Gamma_1 \times \Gamma_2}(t_x, t_y) = \int_{-\pi/2a}^{\pi/2a} \dfrac{d^3 \boldsymbol{p}}{(2\pi)^3} \dfrac{d^3 \boldsymbol{k}}{(2\pi)^3} \langle \tilde{\bar{b}}(\boldsymbol{p}, t_x) \gamma_5 [\tilde{f}(\boldsymbol{p}, t_x) + (-1)^{(t_x)} i\gamma_5 \gamma_0 \tilde{f}^0(\boldsymbol{p}, t_x)][\bar{q}(0) \Gamma_1 b(0) \bar{q}(0) \Gamma_2 b(0)]$

$\times \tilde{\bar{b}}(\boldsymbol{k}, t_y) \gamma_5 [\tilde{f}(\boldsymbol{k}, t_y) + (-1)^{(t_y)} i\gamma_5 \gamma_0 \tilde{f}^0(\boldsymbol{k}, t_y)] \rangle, \quad (2.19)$

where the superscript 0 indicates a temporal taste and no superscript is the null taste at the center of the Brillouin zone.

After Fourier transforming, the bracketed four-quark operator has no restrictions on the tastes that contribute to it. However, it must be contracted with the external quark fields to form the propagators. Because the asqtad action is used, contractions between tastes of different types are suppressed to $O(a^2 \alpha_s^2)$. The three-point function then takes the form

$C_{\Gamma_1 \times \Gamma_2}(t_x, t_y) = \int_{-\pi/2a}^{\pi/2a} \dfrac{d^3 \boldsymbol{p}}{(2\pi)^3} \dfrac{d^3 \boldsymbol{k}}{(2\pi)^3} \langle \bar{b}(\boldsymbol{p}, t_x) \gamma_5 [f(\boldsymbol{p}, t_x) + (-1)^{t_x} i\gamma_5 \gamma_0 f^0(\boldsymbol{p}, t_x)]$

$\times [(\bar{f}(0) + i\gamma_5 \gamma_0 \bar{f}^0(0)) \Gamma_1 b(0) (\bar{f}(0) + i\gamma_5 \gamma_0 \bar{f}^0(0)) \Gamma_2 b(0)] \bar{b}(\boldsymbol{k}, t_y) \gamma_5 [f(\boldsymbol{k}, t_y) + (-1)^{t_y} i\gamma_5 \gamma_0 f^0(\boldsymbol{k}, t_y)] \rangle,$

$(2.20)$

where higher order terms coming from contractions between quarks of different taste give terms of $O(a^2 \alpha_s^2)$ that are not considered here. The effects of such terms are comparable to NLO terms in staggered chiral perturbation theory and need to be considered at that order. They give rise to the "wrong spin" terms discussed below. According to Eq. (2.20), the leading-order correlation functions have

contributions from both the pseudoscalar and the scalar states. The latter ones are known as oscillating states, since the sign of their contribution oscillates with time.

The fit ansatz for our correlators must model both regular and oscillating contributions, so that we can remove the latter and extract the physical matrix elements. This is done using the form

$$C_{\mathcal{O}_i^q}(t_x, t_y) = \sum_{\alpha,\beta=0}^{N_{\text{states}}-1} Z_\alpha Z_\beta \frac{O_{\alpha\beta}^i}{\sqrt{2E_\alpha 2E_\beta}}$$
$$\times (-1)^{(t_x+1)\alpha+(t_y+1)\beta} e^{-E_\alpha t_x - E_\beta t_y}, \quad (2.21)$$

where the sum is over a finite number of states $N_{\text{states}}$. The time $t_x$ in Eq. (2.21) and in the discussion on fitting in Sec. IV is the number of time slices between the initial state and the operator, and thus it is a positive number, unlike the time $t_x$ defined in Fig. 2. The oscillations in Euclidean time given by the factor $(-1)^{(t_x+1)\alpha+(t_y+1)\beta}$ reflect the contribution from the scalar states in Eq. (2.20). The matrix elements of interest are given by the three-point amplitude of the ground state $\alpha = \beta = 0$, $O_{00}^i$. Analogously, we incorporate regular and oscillating contributions to the description of the two-point correlators by using the following functional form in the fits

$$C_{\text{PS}}(t) = \sum_{\alpha=0}^{N_{\text{states}}-1} |Z_\alpha|^2 (-1)^{(t+1)\alpha} (e^{-E_\alpha t} + e^{-E_\alpha(T-t)}),$$
$$(2.22)$$

where $T$ is the temporal size of the lattice. Three- and two-point functions are fit simultaneously in our analysis, as described in Sec. IV.

### D. Improving the heavy-light four-quark operator

In addition to the discretization errors in the heavy-quark action, the mixing operator also has discretization errors due to the difference in the small-momentum behavior of lattice and continuum heavy quarks. In this section, we describe how the lowest order of operator discretization errors are removed in our calculation. We first show that the errors start at $O(ap)$ and then discuss how the error at this order can be removed by a "rotation" of the heavy-quark field.

To begin, consider the small $ap$ expansion of the spinor for the Wilson-like fermion

$$u^{\text{lat}}(\chi, \boldsymbol{p}) = \frac{\gamma_0 \text{sign}\chi \sinh Ea - i\gamma_j \sin(p_j a) + L}{\sqrt{2L(L + \sinh Ea)}} u(\chi, 0)$$
$$= e^{-m_1 a/2}\left[1 - \frac{i\boldsymbol{\gamma} \cdot \boldsymbol{p} a}{2\sinh m_1 a} + O((ap)^2)\right] u(\chi, 0),$$
$$(2.23)$$

where $\chi$ labels spin and particle vs antiparticle, $\hat{p} = (2/a)\sin(pa/2)$, and for the clover action $L = 1 + m_0 a + \frac{1}{2}\hat{p}^2 a^2 - \cos p_0 a$. The continuum spinor has the expansion

$$u^{\text{cont}}(\chi, \boldsymbol{p}) = \left[1 - \frac{i\boldsymbol{\gamma} \cdot \boldsymbol{p}}{2m} + O((ap)^2)\right] u(\chi, 0). \quad (2.24)$$

The mismatch between the small-momentum terms can be easily removed by "rotating" the lattice heavy quark as was done to heavy-light bilinear operators in Ref. [40]. The light lattice spinors for the staggered formulation have the same small-momentum behavior as in the continuum up to $O((\boldsymbol{p}a)^2, (m_q a)^2)$ and need not be matched.

The analysis of Ref. [40] can be generalized from bilinears to four-fermion operators with Dirac structure $\Gamma_1 \times \Gamma_2$. We can take one of the terms in the contraction of the lattice operator

$$\langle q(p_q'), b(p_b')|\bar{q}\Gamma_1 b\bar{q}\Gamma_2 b|q(p_q), b(p_b)\rangle_{\text{lat}} = N_q(p_q')N_q(p_q)N_b(p_b')N_b(p_b)\bar{u}(p_q')\Gamma_1 u_h^{lat}(p_b')\bar{u}(p_q)\Gamma_2 u_h^{lat}(p_b)$$
$$+ \text{(additional contractions)}, \quad (2.25)$$

where $N_q(p)$ and $N_b(p)$ are normalization factors for the $q$ and $b$ one-particle states. Following the Fermilab interpretation in Ref. [40], we demand that lattice and continuum amplitudes match through $O(ap)$,

$$Z\left(\bar{u}(\chi, p_q')\Gamma_1 e^{-am_1^b/2}\left[1 - \frac{i\boldsymbol{\gamma} \cdot \boldsymbol{p}_b' a}{2\sinh am_1^b}\right]u(\chi, 0)\bar{u}(\chi, p_q)\Gamma_2 e^{-am_1^b/2}\left[1 - \frac{i\boldsymbol{\gamma} \cdot \boldsymbol{p}_b a}{2\sinh am_1^b}\right]u(\chi, 0)\right) + aZD_n Q_n$$
$$= \bar{u}(\chi, p_q')\Gamma_1\left[1 - \frac{i\boldsymbol{\gamma} \cdot \boldsymbol{p}_b'}{2m_b}\right]u(\chi, 0) \times \bar{u}(\chi, p_q)\Gamma_2\left[1 - \frac{i\boldsymbol{\gamma} \cdot \boldsymbol{p}_b}{2m_b}\right]u(\chi, 0) + O((\boldsymbol{p}a)^2), \quad (2.26)$$

where $Q_n$ are dimension-seven lattice operators and $D_n$ their corresponding coefficients. These operators and coefficients are straightforward to identify ($m_b$ must be identified with the Fermilab kinetic mass, $M_2$ [74]). There are two dimension-seven operators contributing to the matching, $Q_1 = \bar{q}\Gamma_1 \boldsymbol{\gamma} \cdot \boldsymbol{D} b\bar{q}\Gamma_2 b$ and $Q_2 = \bar{q}\Gamma_1 b\bar{q}\Gamma_2 \boldsymbol{\gamma} \cdot \boldsymbol{D} b$, which will remove the $\boldsymbol{p}a$ discrepancy for appropriate values of the Wilson coefficients and the normalization constant. From the relation above, we find

$$D_1 = D_2 = \left[\frac{1}{\sinh am_1^b} - \frac{1}{2am_2^b}\right] \quad (2.27)$$

and

$$Z = e^{am_1^b}. \quad (2.28)$$

Here, $m_1$ and $m_2$ are again the Fermilab rest and kinetic masses defined in Ref. [40].

However, through $O(ap)$, adding these operators has the same effect as inducing a "rotation" of the heavy field

$$b^r(x) = [1 + ad_1 \boldsymbol{\gamma} \cdot \boldsymbol{D}]b(x), \qquad (2.29)$$

where

$$d_1 = D_1 = D_2. \qquad (2.30)$$

This removes $O(\Lambda_{\text{QCD}}a)$ discretization errors in the operator. The leading errors are then $O((\Lambda_{\text{QCD}}a)^2)$ and $O(\alpha_s\Lambda_{\text{QCD}}a)$. In the way we have set up the calculation, the open-meson propagators, Eq. (2.9), include the rotation.

## III. MATCHING OF THE LATTICE MATRIX ELEMENTS

In order to cancel the scheme and scale dependence of the Wilson coefficients in the effective Hamiltonian, we must relate the bare hadronic matrix elements of the lattice operators in Eq. (1.3) to a continuum scheme. We perform that renormalization and matching perturbatively at one-loop. In the lattice part of this renormalization calculation we use mean field improved lattice perturbation theory [75] to improve the convergence of the theory by resumming the tadpole contributions.

Already at one-loop, even in the continuum, the operators in Eq. (1.3) mix with each other under renormalization. To extract the renormalized value of $\langle \mathcal{O}_1^q \rangle$, we use the following matching relation

$$\begin{aligned}
\langle \mathcal{O}_1^q \rangle^{\text{renor}}(\mu) = \mathcal{C}\{&[1 + \alpha_s \cdot \zeta_{11}(\mu, m_b, am_b)]\langle \mathcal{O}_1^q \rangle^{\text{lat}} \\
&+ \alpha_s \cdot \zeta_{12}(\mu, m_b, am_b)\langle \mathcal{O}_2^q \rangle^{\text{lat}}\} \\
&+ O(\alpha_s^2, \alpha_s\Lambda_{\text{QCD}}a), \qquad (3.1)
\end{aligned}$$

where the renormalization coefficients $\zeta_{ij}$ are the difference between the renormalizations in the continuum and on the lattice, $\zeta_{ij} = Z_{ij}^{\text{cont}} - Z_{ij}^{\text{lat}}$. The continuum renormalization scale at which we perform the matching is $\mu$. The lattice spacing is $a$ and $\mathcal{C}$ is a factor which absorbs the lattice field normalization conventions. The values of $Z_{ij}^{\text{cont}}$ are listed in Ref. [76] and a detailed description of the calculation of the lattice renormalization coefficients will be given in Ref. [77]. Table II lists the tadpole-improved renormalization coefficients relevant for the lattice data analyzed in this paper. For each lattice spacing and $b$ quark mass we show the infrared (IR) finite part of the $Z_{ij}^{\text{lat}}$'s as

TABLE II. Values of the finite part of the lattice one-loop renormalization coefficients $Z_{ij}^{\text{lat}}$, the difference of the continuum and lattice one-loop coefficients $\zeta_{ij}$ needed in Eq. (3.1) for the 0.12 fm and 0.09 fm lattices, and the coupling $\alpha_s$ used in the matching relation. The continuum ($\overline{\text{MS}}$-NDR) scale used in the matching is $\mu = m_b$.

| $\approx a$ (fm) | $am_b$ | $Z_{11}^{\text{lat,finite}}$ | $Z_{12}^{\text{lat,finite}}$ | $\zeta_{11}^{\overline{\text{MS}}\text{-NDR}}$ | $\zeta_{12}^{\overline{\text{MS}}\text{-NDR}}$ | $\alpha_s$ |
|---|---|---|---|---|---|---|
| 0.12 | 2.1881 | −0.726 | −0.325 | 0.1998 | −0.312 | 0.32 |
| 0.09 | 1.7728 | −0.945 | −0.369 | 0.3041 | −0.268 | 0.26 |

well as the corresponding $\zeta_{ij}$ (in the $\overline{\text{MS}}$-NDR continuum scheme). The $\zeta_{ij}$ are IR finite since the IR divergent contributions to $Z_{ij}^{\text{lat}}$ and $Z_{ij}^{\text{cont}}$ cancel in the difference. All the coefficients $Z_{ij}^{\text{lat}}$ in Table II are between 0.3 and 1, which indicates a sensible behavior of the lattice perturbation series.

In order to apply the matching relation Eq. (3.1), we need to choose both a scale $\mu$ and a value for the strong coupling constant $\alpha_s$. For the scale, we use the bottom quark mass and, in that way, eliminate higher-order logarithmic contributions that come in powers of $\log(\mu/m_b)$. For the strong coupling constant, we use the renormalized coupling in the $V$-scheme [69] evaluated at a scale $q^*$, as in Ref. [78]. The scale $q^*$ should be the size of a typical gluon loop momentum in this process and can be calculated using the methods outlined in Refs. [69,79]. Here, we use $q^* = 2/a$ which is close to the calculated value for heavy-light currents using the same actions we are employing [69,80]. This is justified since the contributions coming from the current renormalization are larger than the intrinsic four-quark contributions [77]. The values of $\alpha_V$ are determined from the static-quark potential in a manner similar to that described in Ref. [78] and are also given in Table II.

## IV. FITTING METHOD AND STATISTICAL ERRORS

The correlation functions are calculated at four different time sources $t$, and then averaged over time sources. For the $a \approx 0.12$ fm ensembles, $t_0 = 0, 16, 32, 48$, and for the $a \approx 0.09$ fm ensembles, $t_0 = 0, 24, 48, 72$. The statistical errors in the data and fits decrease with each additional time source by approximately what is expected, suggesting that the correlators from different time sources are weakly correlated and statistical power is gained by averaging.

In order to extract the renormalized matrix elements, we tried two methods for the correlator fits. In the first method, we fit the bare correlators and combine the results afterwards with the matching coefficients in Sec. III to get the renormalized matrix elements. In the second method, we first apply the matching coefficients to the correlators for each configuration and then perform the fits to obtain the renormalized matrix elements. The central values are nearly identical with both methods, but the errors are slightly better and the fits more stable with the latter, so for the rest of this article we discuss only the results obtained with the second method.

### A. Description of the fitting method and stability tests

The heavy quark in the two-point and three-point correlation functions is always rotated at the source as explained in Sec. II D. For three-point functions, we smear the heavy quarks at the sink using a function based on the quark-onium 1S wave function [81,82]. For two-point functions, at the sink we either rotate local heavy quarks, or we smear

them with a 1S wave function. Smearing greatly improves the overlap with the ground state. The additional rotation at the sink is to ensure that the local-local meson correlator is positive-definite. The naive light-quark propagator is always local at source and sink.

The two-point and three-point correlators used to determine the matrix element $\langle \mathcal{O}_1^q \rangle$ on a particular ensemble and for a particular choice of valence-mass $m_q$ are fit simultaneously using the Bayesian fitting approach described in Refs. [83,84]. For the matrix elements on the coarse ensembles, we find the smallest errors and greatest stability using three correlators (two two-point correlation functions and one three-point correlation function):

  (i) $C_{\text{PS}}$ in Eq. (2.22) with local source and local sink.
  (ii) $C_{\text{PS}}$ with local source and 1S smeared sink.
  (iii) $C_{\mathcal{O}_1^q}$ in Eq. (2.21) with local source and 1S smeared sink.

For the fine ensembles, the best results are obtained with only one two-point function and one three-point function:

  (i) $C_{\text{PS}}$ with 1S smeared source and 1S smeared sink.
  (ii) $C_{\mathcal{O}_1^q}$ with local source and 1S smeared sink.

The prior central values function as the initial starting guesses for our fits. Hence, we choose ground-state values guided by our data to help the fits converge. The prior central values for the ground-state masses are obtained from effective mass plots. For the overlap factors $Z_0^d$ and $Z_0^{1S}$, where superscripts $d$ and $1S$ denote factors corresponding to local or 1S smeared sources/sinks, we examine the amplitude of the *B*-meson propagator with the exponential of the ground state removed. We do the same for $O_{00}$, where the $Z_0^{1S}$ amplitudes are accounted for. The prior widths are taken to be large compared with the statistical error of the parameters as reported by the fitter to avoid influencing our fit results by our choice of priors. For the higher states' overlap factors, the prior width is chosen based on the expectation that the overlaps should not be larger than the corresponding ground state ones. The energy differences have prior central values and widths that allow them to vary from $\Delta E_{i+1,i} \equiv a(E_{i+1} - E_i) \approx$ 0.14–0.37, where experimental values [16] have been used as a guide. We checked that the prior widths for all fitting parameters are large enough so they do not influence the central value of relevant quantities extracted from the fits.

The same priors are used for all ensembles, except for the masses of the regular and oscillating ground states, $E_0$ and $E_0'$, respectively. These parameters are strongly determined by the data, and very different at each lattice spacing, therefore the prior choice must also be lattice-spacing dependent. Appendix C contains a list of the prior central values and widths we use in the calculation.

Statistical errors are estimated with the bootstrap method. Specifically, for each ensemble and valence mass, 500 bootstrap ensembles are constructed from the
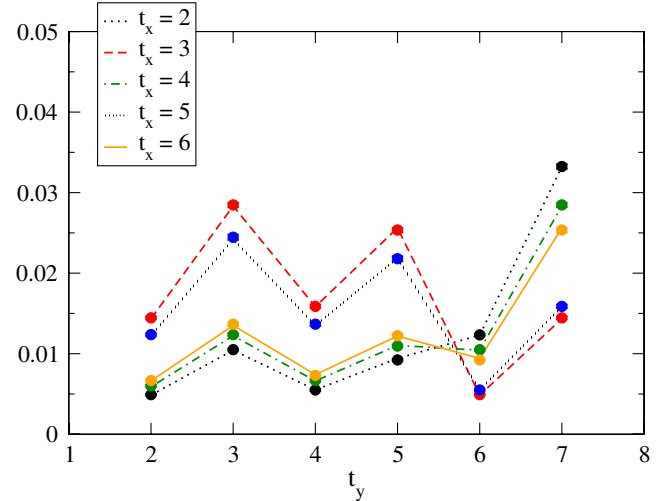


FIG. 3 (color online). Comparison of a fit with correlator data for the correlation function $C_{\mathcal{O}_1^q}(t_x, t_y)$ at fixed $t_x = 2$–6 (labeled by line type) and as a function of $t_y$, for the $a \approx 0.09$ fm ensemble with quark masses 0.0124/0.031 and valence quark mass $am_q = 0.031$. The lines connect the fit function results for integer values of $t_x$, $t_y$ coming from the same single fit and evaluated at specific $t_x$, and the dots are the average over simulation data. Statistical errors on the simulation data are smaller than the plot symbols. The fit results describe very well the oscillation in time shown by the data.

original ensemble by sampling with replacement. A fit is then performed to each ensemble. We find that, as long as the bootstrap ensembles are larger than ∼100, the estimated error is independent of bootstrap ensemble size. For fitting methodology checks and plotting purposes in Figs. 3–6, statistical errors in the parameters are estimated by the average 68% bootstrap error, which is defined as half of the distance between the two points at which 16% of the distribution has a higher (lower) value.

Autocorrelations necessarily exist between correlation functions calculated on different configurations within an ensemble and can be minimized by binning the data. The autocorrelations are observable only in a few ensembles and valence masses. In many ensembles and mass combinations, the noise is large enough that the autocorrelations are not observable. We choose a conservative bin size of 4.

The number of states included in the sum in Eq. (2.21) and the time ranges we use in the fits are shown in Table III. The minimum time slice is fixed to be the same for all the correlators in the fit, three-point as well as two-point. However, the maximum time is fixed separately for the two- and three-point functions. Following Ref. [83], the number of states are determined by first performing the fit using $1 + 1$ states (1 regular state + 1 oscillatory state) starting at large time slices, where the higher-energy states no longer contribute significantly and a good $\chi^2$ per degree of freedom (d.o.f.) is obtained ($\approx 1$). The fit is then performed using one lower time slice as the starting time
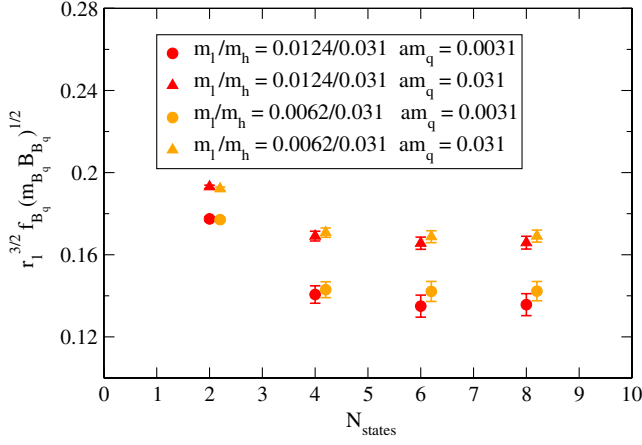
FIG. 4 (color online). $r_1^{3/2} f_{B_q} \sqrt{M_{B_q} B_{B_q}}$ for the two $a \approx$ 0.09 fm ensembles for $am_q = 0.031$, 0.0031 as a function of the number of states $N_{\text{states}}$. The fit results for $f_{B_q} \sqrt{M_{B_q} B_{B_q}}$ reach a plateau for $N_{\text{states}} \geq 4$.
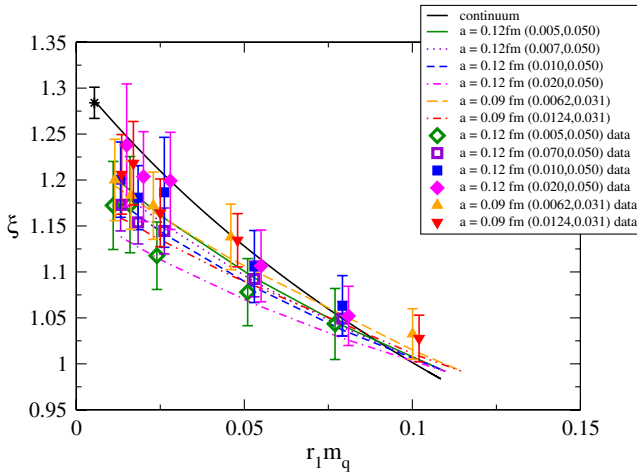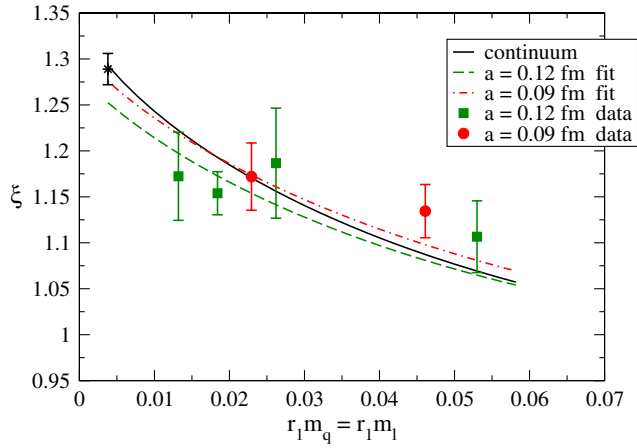


FIG. 5 (color online). Fits results using the NLO rHMS$\chi$PT, first line in Eq. (5.6). The (black) star is the physical value of $\xi$ in both panels. The top panel shows only the full QCD data, while the bottom one shows all the data included in the fits. The (green) squares and (red) circles and lines in the upper panel represent the 0.12 fm and 0.09 fm data and fit results, respectively. In the bottom panel, each color (symbol) labels a different ensemble.
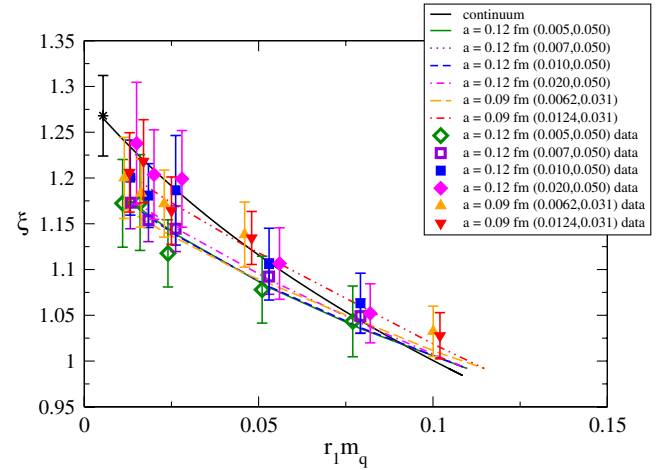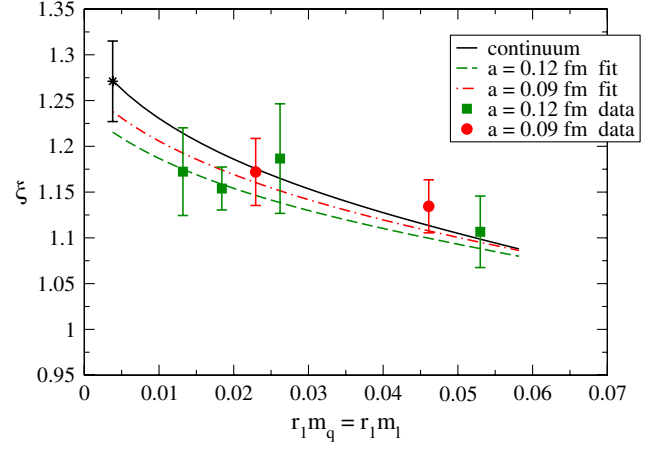


FIG. 6 (color online). Fits results using the NNLO rHMS$\chi$PT in Eq. (5.6). The (black) star is the physical value of $\xi$ in both panels. The top panel shows only the full QCD data, while the bottom one shows all the data included in the fits. The (green) squares and (red) circles and lines in the upper panel represent the 0.12 fm and 0.09 fm data and fit results, respectively. In the bottom panel, each color (symbol) labels a different ensemble.

$t_{\text{min}}$, and this is repeated, reducing $t_{\text{min}}$ until the $\chi^2$/d.o.f. is no longer reasonable, $\gtrsim 1.5$. Then an additional pair of states is added to the model function, and the process iterated. Once the time slice $t = 2$ can be included, that number of states is used in our central-value fits.[2]

For the three-point function, we fit using $N_{\text{states}} = N_{\text{states}}^{\text{regular}} + N_{\text{states}}^{\text{oscillating}} = 2 + 2 = 4$ and time slices $t_x, t_y \in [2, 10]$ for all ensembles. The two-point functions are fit for $t \in [2, 20]$ using $3 + 3$ states. The output of these fits successfully describe the oscillations in the correlation functions as can be seen in Fig. 3, which shows the typical behavior in one of the ensembles analyzed.

---

[2]Time slices $t = 0$ and $t = 1$ contain unconstrained contamination from higher-energy states. At $t = 0$ all states contribute because they have the same exponential weighting, and $t = 1$ is contaminated by higher-energy states because the degrees of freedom of staggered fermions spread over two time slices.

TABLE III. Number of states and time ranges used for each correlator in the fits for both the $a \approx 0.12$ fm and $a \approx 0.09$ fm ensembles. For the number of states, the first value indicates the number of regular states and the second one, the number of oscillating states. The labels between parentheses in the first column indicate the type of source/sink in that correlator.

| Correlator | Number of states | Time range |
|---|---|---|
| $C_{\mathrm{PS}}$ (local/local) | $3 + 3$ | 2–20 |
| $C_{\mathrm{PS}}$ (local/1S) | $3 + 3$ | 2–20 |
| $C_{Q_q^i}$ (local/1S) | $2 + 2$ | $(2 - 10) \times (2 - 10)$ |

In order to check that this number of states is sufficient, we add more states and examine the stability of the fits. Stability plots over numbers of states for the $a \approx 0.09$ fm ensembles, which illustrate the typical behavior of our fits, are shown in Fig. 4. The stability of central values and errors is very good for $N_{\mathrm{states}} \geq 4$ in all cases.

## V. CHIRAL PERTURBATION THEORY

The light sea- and valence-quark masses that are used in our lattice simulations have unphysically large values, with our lightest pion mass $\approx 240$ MeV. To obtain information about the quark-mass dependence of the relevant matrix elements, which allows us to extrapolate our results to the physical masses, we perform our calculation at six sea × six valence quark masses, thus including numerous partially quenched data points. In addition, the leading-order taste violations, which arise at $O(a^2 \alpha_s)$, are included in the theory and then removed when the extrapolation is performed using rooted heavy-meson staggered chiral perturbation theory (rHMS$\chi$PT) [61,62,65]. The $\chi$PT expression for $\langle \mathcal{O}_1^q \rangle$, as well as for the matrix elements of all the other operators in the $\Delta B = 2$ effective Hamiltonian, was first described in Ref. [85] for partially quenched Wilson-type quarks in the framework of continuum heavy-meson $\chi$PT (HM$\chi$PT). With staggered fermions, we also must include the effects of taste-violating interactions, using rHMS$\chi$PT [65].

With the four-quark operators, a careful examination of the Fierz properties shows that there are additional operators with both wrong taste and *spin*, i.e., wrong $(\Gamma_1, \Gamma_2)$. As far as we know, this property of local heavy-staggered four-quark operators has not been discussed in the literature before. The needed rHMS$\chi$PT expressions are derived in Ref. [86]. We became aware of these contributions after our analysis was nearly complete, so we have not included them in the chiral fit functions used here. We do, however, estimate the associated systematic error on $\xi$ in our error budget (cf. Sec. VI C). Explicit expressions for the chiral fit functions used in this work are given in Appendix A.

The NLO rHMS$\chi$PT in Eq. (A1) and subsequent equations in the appendix can be schematically written as

$$\langle \bar{B}_q | \mathcal{O}_1^q | B_q \rangle = \frac{2}{3} M_{B_q}^2 f_{B_q}^2 B_q$$
$$= M_{B_q} \alpha [1 + (\mathcal{W}_q + \mathcal{T}_q + \mathcal{Q}_q) + L_v m_q$$
$$+ L_s (2m_l + m_h) + L_a a^2], \qquad (5.1)$$

where $\alpha$, $L_v$, $L_s$, and $L_a$ are low-energy constants (LECs) to be determined from fitting the data, the factor of $M_{B_q}$ comes from the HQET normalization of states, and the masses $m_q$, $m_l$, and $m_h$ are the light valence, light sea, and strange sea-quark masses, respectively. The light sea quarks are treated as degenerate, and the isospin average is used, i.e., $\hat{m} = (m_u + m_d)/2$. For staggered quarks, the taste-nonsinglet pseudoscalar meson masses are split

$$M_{ij,\rho}^2 = \mu(m_i + m_j) + a^2 \Delta_\rho, \qquad (5.2)$$

where $m_i$ and $m_j$ are the quark masses and the sixteen meson masses are labeled by their taste representation, $\rho = P, A, T, V, I$. The parameters $\mu$ and the $\Delta_\rho$'s are determined from lattice calculations for pions and kaons [87]. Their values are collected in Table IV.

The chiral logarithms, $\mathcal{W}_q$, $\mathcal{T}_q$, $\mathcal{Q}_q$, stem from wave function renormalization, tadpole, and sunset diagrams,

TABLE IV. Inputs for the priors of the free parameters and for the fixed parameters in the fits. The NLO low-energy constants $L_v$, $L_s$, and $L_a$ are not constrained in the fits. The parameter $s$ is given by the quantity $1/(8\pi^2(r_1 f_\pi)^2)$. We do not consider errors on the slope $\mu r_1$ or the taste splittings $r_1^2 a^2 \Delta_\rho$ because those have negligible effect on the final results. In the right-hand side table, the two last columns correspond to lattice spacings $a \approx 0.12$ fm and $a \approx 0.09$ fm. See the text for explanations of the choices of parameters.

| Fit parameters (central value ± width) | | |
|---|---|---|
| $a \approx 0.12$ fm | | $(a \approx 0.09$ fm$)$ |
| $\beta$ | | $1 \pm 1$ |
| $g_{B^* B \pi}$ | | $0.51 \pm 0.20$ |
| $r_1^2 a^2 \delta'_V$ | | $0.0 \pm 0.07$ $(0.0 \pm 0.07) \times 0.35$ |
| $r_1^2 a^2 \delta'_A$ | | $-0.28 \pm 0.06$ $(-0.28 \pm 0.06) \times 0.35$ |
| $L_v$ | | Unconstrained |
| $L_s$ | | Unconstrained |
| $L_a$ | | Unconstrained |
| $Q_{1-4}$ | | $0 \pm s^2$ |
| $P_{1-3}$ | | $0 \pm s^2$ |

| Input (fixed) parameters | | |
|---|---|---|
| $a \approx 0.12$ fm | | $a \approx 0.09$ fm |
| $f_\pi r_1$ | | 0.2106 |
| $\mu r_1$ | 6.234 | 6.382 |
| $r_1^2 a^2 \Delta_P$ | 0 | 0 |
| $r_1^2 a^2 \Delta_V$ | 0.439 | 0.152 |
| $r_1^2 a^2 \Delta_T$ | 0.327 | 0.115 |
| $r_1^2 a^2 \Delta_A$ | 0.205 | 0.0706 |
| $r_1^2 a^2 \Delta_I$ | 0.537 | 0.206 |

respectively. The explicit expressions can be found in Appendix A.

In this work, we do not include the effects of the hyperfine splitting $\Delta^*$ or the light-flavor splittings $\delta_k$ defined in the Appendix. The wrong spin terms contribute to the tadpole and sunset diagrams [86].

When extrapolating to the physical point, we set the parameters $\Delta_\rho$ and $\delta'_{A,V}$ (which describe discretization effects) and the lattice spacing to zero, and set the sea quark masses to their physical values, $m_l \to (m_u + m_d)/2$, and $m_h \to m_s$. We then obtain $\langle \bar{B}^0_d | \mathcal{O}^d_1 | B^0_d \rangle$ or $\langle \bar{B}^0_s | \mathcal{O}^s_1 | B^0_s \rangle$ by setting $m_q = m_d$ or $m_s$. Thus, it is an extrapolation to the *u*- and *d*-quark masses and an interpolation to the *s*-quark mass.

An additional consideration is that $SU(3)$ NLO $\chi$PT may not be valid for data with masses as large as the strange quark's. It would be desirable to include next-to-next-to-leading order (NNLO) contributions to test the validity of the NLO expression, but the effort needed to calculate the NNLO logs is prohibitive. It is reasonable instead to test the chiral expansion by including just NNLO analytic contributions. Wherever the quark masses or splittings are large enough for such analytic NNLO terms to be significant, the NNLO logarithms should be slowly varying and well approximated by the analytic terms. We follow this strategy and supplement the NLO rHMS$\chi$PT expressions with NNLO analytic terms in our fits, with prior constraints estimated based on $\chi$PT power counting as explained in the next section.

### A. Parametrization of the chiral expression

Dimensionful quantities are extracted first in units of the lattice spacing and then converted to their physical values using the $r_1$ scale [88,89]. This absolute scale is defined as $r_1^2 F(r_1) = 1.0$, where $F(r_1)$ is the force between static quarks. In our chiral fits, all parameters are first converted to units of $r_1$ by multiplying by the relative scale $r_1/a$. The values for $r_1/a$ on every MILC ensemble used in our calculation are listed in Table I. After the chiral-continuum extrapolation, we convert from $r_1$ units with a physical value of $r_1$. We take the result obtained by combining the 2009 MILC determination of $r_1 f_\pi$ [90] and the PDG value of $f_\pi$ [16]. Following Ref. [59], the error for $r_1$ is determined by averaging the MILC value with the HPQCD value in Ref. [91] and then by inflating the uncertainty to take conservatively into account the possible correlations coming from the use of the same configurations in both determinations. The final value we use is $r_1 = 0.3117(22)$ [59]. The error associated with $r_1$ has a very small effect on the dimensionless quantity $\xi$.

The dominant lattice artifacts to take into account in our rHMS$\chi$PT expressions are expected to be taste-violating contributions of $O(a^2 \alpha_s^2)$, since the $O(a^2 \alpha_s)$ taste-violating effects are absent for asqtad quarks. We parametrize these effects in our fits by defining a quantity $A_a^2$,

which is the ratio of the size of taste violations on lattices with spacing *a* to those on the $a \approx 0.12$ fm lattices. Thus $A_{0.12\,\mathrm{fm}}^2 = 1$ and

$$A_{0.09\,\mathrm{fm}}^2 \equiv \frac{(\alpha_s^2 a^2)_{0.09\,\mathrm{fm}}}{(\alpha_s^2 a^2)_{0.12\,\mathrm{fm}}} \sim 0.35. \qquad (5.3)$$

The NLO rHMS$\chi$PT function in Eq. (5.1) can then be rewritten as

$$
\begin{aligned}
\beta_q &\equiv \sqrt{\frac{3}{2} \langle \bar{B}_q | \mathcal{O}^q_1 | B_q \rangle / M_{B_q}} \\
&= f_{B_q} \sqrt{M_{B_q} B_{B_q}} \\
&= \beta^\chi \left[ 1 + \frac{1}{2}(\mathcal{Q}_q + \mathcal{W}_q + \mathcal{T}_q) + \frac{L_v}{2} M_{qq}^2 \right. \\
&\quad \left. + \frac{L_s}{2}(2M_{ll}^2 + M_{hh}^2) + \frac{L_a}{2} A_a^2 \right],
\end{aligned} \qquad (5.4)
$$

where $\beta^\chi = \sqrt{\alpha}$. The masses $M_{ij}$ are defined in Eq. (5.2), but here we disregard $a^2$ corrections in the masses since they can be absorbed by a redefinition of the low-energy constants at higher order in the chiral expansion. To the NLO expression above we add, inside the square brackets, the allowed NNLO analytic terms, which contribute with seven more unknown LECs

$$
\begin{aligned}
&Q_1 M_{qq}^4 + Q_2 (2M_{ll}^2 + M_{hh}^2)^2 + Q_3 M_{qq}^2 (2M_{ll}^2 + M_{hh}^2) \\
&\quad + Q_4 (2M_{ll}^4 + M_{hh}^4) + P_1 A_a^2 M_{qq}^2 + P_2 A_a^2 (2M_{ll}^2 + M_{hh}^2) \\
&\quad + P_3 A_a^4.
\end{aligned} \qquad (5.5)
$$

In the above expressions, we have suppressed the factors of $r_1$ for simplicity. They can be deduced using dimensional considerations. In these expressions, which are the ones we use as fit functions, we write the analytic terms for convenience as functions of the pseudoscalar masses $M_{ij}$ rather than the quark masses.

The ratio $\xi$ can be extracted by first interpolating $\beta_q$ to $m_q = m_s$ and extrapolating to $m_q = m_d$ separately according to expressions (5.4) and (5.5), and then forming the ratio $\beta_s/\beta_d$. Alternatively, one can consider the ratio of chiral expressions and expand up to NNLO to obtain

$$
\begin{aligned}
\xi' = \frac{\beta_{q'}}{\beta_q} &= \xi \frac{\sqrt{M_{B_{q'}}}}{\sqrt{M_{B_q}}} \\
&= 1 + \frac{1}{2}(\mathcal{Q}_{q'} + \mathcal{W}_{q'} + \mathcal{T}_{q'} - \mathcal{Q}_q - \mathcal{W}_q - \mathcal{T}_q) \\
&\quad + \frac{L_v}{2}(M_{q'q'}^2 - M_{qq}^2) + Q_1(M_{q'q'}^4 - M_{qq}^4) \\
&\quad + Q_3(M_{q'q'}^2 - M_{qq}^2)(2M_{ll}^2 + M_{hh}^2) \\
&\quad + P_1(M_{q'q'}^2 - M_{qq}^2)A_a^2,
\end{aligned} \qquad (5.6)
$$

with $m_{q'}$ fixed to the value closest to $m_s$. In Eq. (5.6) we disregard the NNLO terms coming from squaring the NLO terms in the denominator, since they are not necessary to obtain good fits and they are difficult to disentangle from those already included. We can then interpolate/extrapolate to $m_{q'} = m_s$ and $m_q = m_d$. We call these two strategies for the chiral and continuum extrapolation of $\xi$ the indirect and direct methods, respectively. Many of the fit parameters cancel in the chiral expression for $\xi'$ in Eq. (5.6), improving the reliability and stability of the fits. In addition, discretization errors of $O(\alpha_s \Lambda_{\rm QCD} a, (\Lambda_{\rm QCD} a)^2)$ from the heavy-quark action that are not included in the chiral perturbation theory, partially cancel in the ratio. We thus choose this method as our preferred fitting strategy.

### B. Results from the chiral fits

In order to perform the chiral fits, we first create 200 bootstrap samples of $\beta_q$ for each sea- and valence-quark mass combination from the two- and three-point correlator fits. The bootstrap data is then fit to the chiral expression using Bayesian techniques. The fits are simultaneously performed to all ensembles in Table I.

The input and fit parameters are set as in Table IV. We do not impose any constraint on the NLO low-energy constants. For the NNLO LECs we use prior widths based on a simple power counting argument. The NLO analytic terms should be of magnitude similar to the NLO logs, which are $\sim m_\pi^2/(8\pi^2 f_\pi^2) = s(r_1 m_\pi)^2$, (with $s \equiv 1/(8\pi^2(r_1 f_\pi)^2)$). Hence, the NNLO terms are $\sim (m_\pi^2/(8\pi^2 f_\pi^2))^2 = s^2(r_1 m_\pi)^4$. The taste-violating hairpin parameters, $\delta_V'$ and $\delta_A'$, were also determined from lattice calculations for pions and kaons in Ref. [66]. We constrain the parameters $\delta_V'$ and $\delta_A'$ in our fits using the results of Ref. [66] as prior central values and widths. We also take the effective coupling of the $B^* B\pi$ interaction, $g_{B^* B\pi}$, as a fit parameter in our analysis. The prior central value and width we use for this parameter, shown in Table IV, covers the main ranges of determinations of $g_{B^* B\pi}$ [92–98], as discussed in Ref. [57]. A more recent, precise value of $g_{B^* B\pi}$, obtained with $N_f = 2 + 1$ domain-wall fermions and static $b$ quarks [99], was not yet available when this stage of the analysis was carried out. Nevertheless, the result obtained by the authors in Ref. [99], $0.449 \pm 0.047 \pm 0.019$, falls well within the prior central value and width considered here. For the pion decay constant, we use the PDG value, $f_\pi = (130.41 \pm 0.20)$ MeV [16].

The fit results for $\xi$ using different ansätze for the fitting function and the direct and indirect methods explained in Sec. V A are listed in Table V. The results and errors obtained using the direct and indirect methods agree very well, especially when NNLO terms are included. This constitutes a good check of how well our results are encompassing higher-order terms in the chiral expansion, which are different in these two methods.

TABLE V. Results from the rHMS$\chi$PT fits. Errors are only statistical and obtained from 200 bootstrap samples. For a full discussion of systematic errors, see Sec. VI.

| Ansatz | $\chi^2$/d.o.f. | $\xi$ Direct | $\chi^2$/d.o.f. | $\xi$ Indirect |
|---|---|---|---|---|
| NNLO | 0.45 | $1.268^{+0.035}_{-0.044}$ | 0.23 | $1.255^{+0.034}_{-0.041}$ |
| NLO | 0.78 | $1.284^{+0.018}_{-0.016}$ | 0.49 | $1.262^{+0.008}_{-0.012}$ |

In Figs. 5 and 6, we show the NLO and NNLO fit results for $\xi$ from the direct method as a function of the light valence mass in $r_1$ units, $r_1 m_q$. The top panels in both figures show only the full QCD points, $m_q = m_l$, while the bottom panels show all (partially quenched) data included in the fits (see Table I). The fit curve is the same in both panels of each figure. The black lines show the results of the fit in the continuum limit, after the dominant lattice artifacts are removed using rHMS$\chi$PT, and after interpolating the physical sea and valence strange-quark masses to the physical value, as a function of the valence light-quark mass. The black point is our result for $\xi$ at the physical masses, and includes statistical errors.

From the spread of data in the bottom panels of both figures (same data), one can see that the light sea-quark mass dependence is mild; all different sea-quark masses (squares or triangles at a particular axis value) agree within one statistical $\sigma$. The discretization errors are also small, as can be seen in both the data and the extrapolation lines in the upper panels.

We obtain fits that match the data well and have good $\chi^2$/d.o.f. with only the inclusion of NLO terms, as shown in Fig. 5. When we add the NNLO terms, the central values for $\xi$ are also within one statistical $\sigma$, although errors are significantly larger. This is to be expected, since the NNLO LECs are poorly known. Related to this is the fact that the $\chi^2$/d.o.f. for the NLO fits are larger than for the NNLO fits. At NNLO, we are including extra degrees of freedom with large prior widths that are poorly determined by the fit, so, in practice, we are dividing the same $\chi^2$ by a larger number of degrees of freedom. In fact, the NNLO fits seem to give a slightly better description of the data, as can be seen in the full QCD plots. The chiral extrapolation for $\xi$ is also milder in the NNLO case. Based on these arguments and, as mentioned above, the fact that direct and indirect fits agree better at NNLO, we choose the direct NNLO fit for our central value and statistical error. The systematic error associated with our choice of fit function is discussed in Sec. VI D.

### VI. ERROR ANALYSIS

In this section, we discuss all sources of systematic uncertainty affecting our calculation of $\xi$. The systematic errors have to be added to the statistical uncertainty listed in Table V, which also encompasses our imperfect knowledge about chiral parameters such as $g_{B^* B\pi}$ and $\delta_{V,A}'$.

### A. Heavy-quark mass uncertainty

The mixing parameters depend on the *b* quark mass used in our simulations through the hopping parameter $\kappa_b$, which is tuned so that the kinetic meson mass, $M_2$, agrees with the experimental result. The dispersion relation for a heavy particle can be written for low-momentum as

$$E(\boldsymbol{p}) = M_1 + \frac{\boldsymbol{p}^2}{2M_2} - \frac{a^3 W_4}{6} \sum_j p_j^4 - \frac{(\boldsymbol{p}^2)^2}{8M_4^3} + O(\boldsymbol{p}^6),$$
(6.1)

where $\kappa_b$ enters into the definitions of $M_1$ and $M_2$ [40]. $W_4$ and the deviation of $M_4$ from $M_2$ capture lattice artifacts. We calculated two-point functions for the pseudoscalar and vector mesons at several momenta and extract the energy, $E(\boldsymbol{p})$, for each particle at each momentum. A fit to the dispersion relation then determines $M_2$ and the spin average of the results is taken. The $\kappa_b$ value is then adjusted until $M_2$ agrees with the spin-averaged $B_s$ meson mass.

In this work, we use the values for $\kappa_b$ on the $a \approx 0.12$ fm and $a \approx 0.09$ fm ensembles tuned this way in Ref. [55]. The error in the determination translates into a systematic error in the mixing matrix elements. However, in the ratio $\xi$ the effect of the uncertainty is minimal, since the corrections go in the same direction in both denominator and numerator, and, thus, largely cancel. In addition, most of the remaining dependence is encoded in the decay constants rather than in the bag parameters, which are very insensitive to the exact values of the quark masses. In Ref. [59], we studied the decay constants with the same choice of actions, parameters, and configurations as here. We expect systematic errors to be very similar in both analyses. We therefore, adopt the error due to the uncertainty in the *b* quark mass obtained in Ref. [59] for the ratio of decay constants $f_{B_s}/f_{B_d}$, namely, 0.4%, as a good estimate of this systematic error for $\xi$.

### B. Higher-order effects in the perturbative matching

The most straightforward and conservative way to estimate the effects of the missing higher order terms in the perturbative matching is to assume two-loop coefficients of order 1 and to multiply the central value by $\alpha_s^2 = \alpha_V^2(2/a)$. This estimate gives an error $\sim 5\%$ for $f_B\sqrt{B_B}$ on the $a \approx 0.12$ fm lattices and $\sim 3.6\%$ on the $a \approx 0.09$ fm lattices, becoming the main source of uncertainty for this quantity [37]. If there was no mixing between $\langle \mathcal{O}_1 \rangle$ and $\langle \mathcal{O}_2 \rangle$ under renormalization, there would be an exact cancellation of the renormalization coefficients for the ratio $\xi = (f_{B_s}\sqrt{B_{B_s}})/(f_{B_d}\sqrt{B_{B_d}})$, as long as the valence light-quarks are taken to be massless in the renormalization calculation. The mixing under renormalization prevents this exact cancellation from happening, but the renormalization corrections in the ratio are still largely suppressed, by a factor of $\langle \mathcal{O}_2^s \rangle / \langle \mathcal{O}_1^s \rangle - \langle \mathcal{O}_2^d \rangle / \langle \mathcal{O}_1^d \rangle$, with respect to those for a single matrix element. We estimate this suppression factor via the

ratio $(m_s - m_d)/\Lambda_{\text{QCD}}$ and multiply the perturbative error for $f_B\sqrt{B_B}$ given above by it. As a result, the perturbative matching uncertainty for $\xi$ from this estimate is 0.2–0.5% (for $\Lambda_{\text{QCD}} = 700$ MeV).

The ratio $\xi$ changes by 0.2% when the one-loop renormalization is omitted entirely, supporting our power counting argument. Another way of estimating $O(\alpha_s^2)$ effects is by varying the scale $q^*$ at which $\alpha_V$ is evaluated. If we change $q^*$ from our central value of $2/a$ to $1/a$ and $3/a$ we find that the extrapolated $\xi$ changes between 0.2 and 0.4%.

Since the initial estimate yields the largest uncertainty, 0.5%, we take this as the error associated with the missing higher order terms in the perturbative renormalization. Hence, this source of uncertainty is subdominant in our determination of $\xi$.

### C. Mixing with wrong spin four-fermion operators

As mentioned in Sec. V, there are contributions at NLO in rHMS$\chi$PT originating from the mixing of $\langle \mathcal{O}_1 \rangle$ with the matrix elements of four-fermion operators of different spin and taste. We have omitted these contributions from our chiral fits, because we discovered these terms after this stage of the analysis was complete. From Figs. 5 and 6, one can see that the effect of the wrong spin mixing is unlikely to be very large, perhaps being mostly absorbed into the LECs.

We cannot include the effects of the wrong spin contributions, because they require the matrix elements of $\mathcal{O}_3$, which we have not computed here. Fortunately, however, we have started a more comprehensive analysis of $B$-$\bar{B}$ mixing on a larger set of higher-statistics ensembles, including $\mathcal{O}_3$. We have added the wrong spin operators to that analysis and find that their inclusion tends to increase the slope of the continuum extrapolated chiral fit function for $\langle \mathcal{O}_1 \rangle$ and, hence, $\xi$. For example, taking priors and widths similar to those in Table IV, we find a 2% increase in $\xi$, while for other reasonable choices of the priors the variation is not larger than 3.2%. We add a 3.2% systematic error to account for the missing terms in our chiral extrapolation functions.

### D. Chiral-extrapolation systematics and light-quark discretization

The errors due to the choice of fit ansatz and light-quark discretization effects cannot be disentangled, because every fit ansatz necessarily treats the discretization errors differently. So any estimate of the systematic uncertainty associated with the choice of ansatz also accounts for the light-quark discretization errors left over after removing the dominant ones using rHMS$\chi$PT.

In Fig. 7, we show the distribution of values for $\xi$ obtained with the NNLO direct fit for the 200 bootstrap samples analyzed. We check that 200 bootstrap samples is enough to obtain a (nearly) Gaussian distribution, as can be seen in the panel. With the goal of testing our choice of
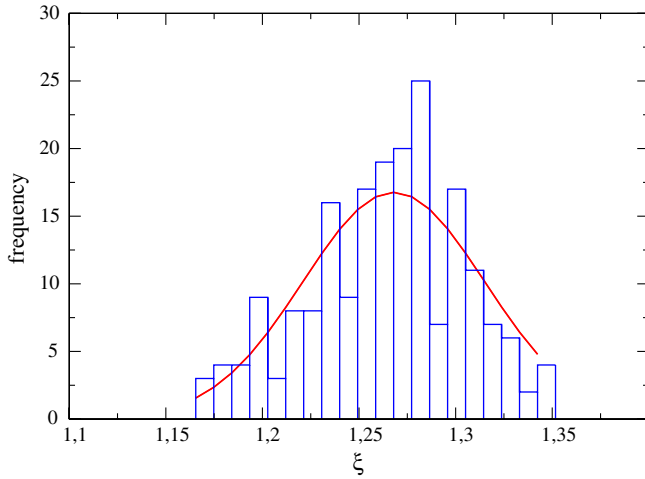
FIG. 7 (color online). Histogram of the distribution of values of $\xi$ obtained from the 200 bootstrap samples. The curve is a Gaussian distribution corresponding to $\xi = 1.268 \pm 0.047$ (the NNLO result with augmented errors as explained in the text).

functional form and the error associated with the truncation of the chiral series, we perform fits with only two of the three NNLO terms, omitting each one in turn. All fits give good values of $\chi^2/\text{d.o.f.}$ and p value. The values of $\xi$ obtained are scattered around the distribution in Fig. 7 but always within 1.5 statistical $\sigma$ of the central value. This consistency, together with the fact that the NNLO LECs are not well determined by the fit, indicates that the statistical error already accounts for the possibility of having one of the unknown constants equal or close to zero. If we inflate and symmetrize our statistical errors to $\pm 0.047$, we cover the spread of results from the different fitting functions tried (including the NLO one). We take this value as our estimate of both statistical and chiral systematic uncertainties.

An alternative way of estimating the uncertainty in the truncation of the chiral series and the fitting function would be taking the difference between the NLO and the NNLO fits results. If we add this difference with the statistical errors in Table V in quadrature, the uncertainty would be slightly smaller than the $\pm 0.047$ we are taking as our estimate of these two sources of error.

In the rest of this section, we list the errors associated with the uncertainty of several input parameters used in the continuum and chiral fits, that typically can be estimated by varying the inputs and redoing the fits.

### 1. Light-quark mass uncertainty

The physical values of the light-quark masses used for the extrapolations and interpolations for $\xi$ are determined by the MILC collaboration [87,100]. They are obtained by making the charged pions and kaons take on their physical values after removal of electromagnetic effects and are listed in Table VI.

TABLE VI. Input for the physical light-quark masses used in the chiral extrapolations. These values were determined by the MILC collaboration [87,100]. Physical values are found from chiral fits that have been extrapolated to the continuum, but masses are still in units of the $a \approx 0.09$ fm lattice spacing.

| Quantity | Physical |
|---|---|
| $am_s \times 10^2$ | 2.72 (8) |
| $a \frac{(m_u + m_d)}{2} \times 10^3$ | 0.997 (35) |
| $am_d \times 10^3$ | 1.40(6) |

The error on $\xi$ due to the light-quark mass uncertainties is obtained by individually varying each quark mass within this uncertainty and repeating the preferred chiral fit and extrapolation. The central values arrived at using each mass variation are compared to the results of the fit which used the central values for the masses, and their differences are added in quadrature. This gives a total systematic error due to the light-quark mass uncertainties of 0.5% for $\xi$.

### 2. Uncertainty in the scale $r_1$

The value of $r_1$ used in this analysis to convert from lattice to physical units, as described in Sec. VA, is $r_1 = 0.3117(22)$ fm. The results discussed in previous sections are obtained by fixing $r_1$ to its central value. In order to estimate the uncertainty due to the error in $r_1$ we change $r_1$ by $\pm 0.0022$ fm and all parameters that depend on the physical $r_1$ are appropriately adjusted. The uncertainty in scale gives a systematic error of 0.2%, which is very small due to the fact that $\xi$ is a dimensionless quantity and the scale only enters in the normalization to lattice units of the chiral corrections $(1/(f_\pi r_1)^2)$ and indirectly via the tuning of the quark masses.

### E. Heavy-quark discretization effects

The discretization errors associated with our choice of heavy-quark action to simulate the bottom quarks can be described in terms of the difference in the lattice and continuum Wilson coefficients of higher dimension operators in the HQET expansion. Those come from two sources: the mismatch between continuum and lattice in the Lagrangian and the mismatch in the four-fermion operators whose matrix elements yield $f_B\sqrt{B_B}$ and, thus, $\xi$. For a particular operator $Q_i$ the error can be written in terms of the usual power counting magnitudes times functions that reflect the particular $m_0 a$ dependence of the action [56,101]

$$\text{error}_i = z_i f_i(m_0 a)(a\Lambda_{\text{QCD}})^{s_i}, \qquad (6.2)$$

where $s_i = \dim Q_i - 4$ for Lagrangian operators $Q_i$ of dimension 4 and 5, and $s_i = \dim Q_i - 6$ for four-fermion operators $Q_i$ of dimension 7 and 8, and the $z_i$ are constants. The functions $f_i(m_0 a)$ can be deduced from Refs. [40,102] and were discussed in detail in Ref. [56] for the form

factors parametrizing $B \rightarrow \pi l \nu$ decays and in Ref. [59] for heavy-light decay constants. A detailed study of these corrections for the matrix elements of all the operators contributing to neutral $B$ mixing in the SM and beyond will be presented elsewhere [77]. Here we only summarize the sources of the different corrections for $\langle \mathcal{O}_1 \rangle$ and $\xi$. The explicit form of the different functions $f_i(m_0 a)$ can be found in Appendix B.

From the Lagrangian, there are $O(a^2)$ errors and $O(\alpha_s a)$ errors which are identical to those in Eqs. (A12) and (A19) in Ref. [59]. They are proportional to the functions $f_E(m_0 a)$ in Eq. (B1) and $f_B(m_0 a)$ in Eq. (B2) of Appendix B, respectively.

From the four-fermion operators, we have $O(a^2)$ errors coming from higher order corrections to the rotation relation in Eq. (2.29). They are generated by the mismatch between lattice and continuum coefficients of the operators $\bar{q}\Gamma \boldsymbol{D}^2 b$, $\bar{q}\Gamma i \cdot \boldsymbol{B} b$ and $\bar{q}\Gamma \boldsymbol{\alpha} \cdot \boldsymbol{E} b$ in the same way as in Eqs. (A13) and (A14) of Ref. [59], but with an extra overall factor of two due to the fact that we have two heavy fields in our leading-order operator, not one. These corrections are proportional to the functions $f_X(m_0 a)$ and $f_Y(m_0 a)$ in Eq. (B3) of Appendix B, respectively.

The last contribution, and the least straightforward, comes from the $O(\alpha_s a)$ corrections to the four-fermion operators. In principle, to subleading order, there are a basis of twelve new local operators in the effective Hamiltonian. However, using symmetry constraints, Fierz transformations, and rewriting some combinations as total derivatives, only five independent operators remain [37,103]. Because we separate the temporal and spatial parts of the operators through this analysis, the total temporal and spatial components of those five operators should be compared with the temporal and spatial parts of the leading operators. As explained in Ref. [77], this produces a difference of $3f_3(m_0 a)$, where $f_3$ is given in Eq. (B4).

In Table VII, we list all the contributions together with the functions $f_i$, and the proportionality constant $z_i$ in Eq. (6.2). We also list the numerical values of the different contributions to the heavy-quark discretization error in percentage for $f_B\sqrt{B_B}$. In order to get the numerical results, we use $\Lambda_{\text{QCD}} = 700$ MeV and $\alpha_s = \alpha_V(2/a)$ listed in Table II. The final heavy-quark discretization error for

$f_B\sqrt{B_B}$ is 1.3% for the $a \approx 0.09$ fm ensembles and 2.1% for the 0.12 fm ones.

These errors largely cancel in $\xi$. The effect of the cancellation on the error can be estimated by multiplying the errors in $f_B\sqrt{B_B}$ by a factor of $(m_s - m_d)/\Lambda_{\text{QCD}}$ which gives a final heavy-quark discretization error for $\xi$ of 0.2% for the coarse lattice and 0.1% for the fine lattice. This agrees well with the estimate of this type of error for the ratio $f_{B_s}/f_B$ [59], $\sim$0.3%, using a very similar set of data and statistics. The strategy followed in Ref. [59] differs from the one described here. In that paper, terms of the form in Eq. (6.2) were directly added to the chiral and continuum extrapolation fitting functions with a coefficient of order one to be determined by the fit. Ultimately, we would like to employ that strategy also for $B^0$-$\bar{B}^0$ mixing studies. For this work, however, we simply take the larger estimate from the ratio $f_{B_s}/f_B$ as our estimate of the uncertainty in $\xi$ due to heavy-quark discretization errors.

### F. Finite volume corrections

In order to evaluate the finite volume corrections in our calculation, we follow the prescription in Refs. [65,104]. The MILC lattices are large enough in the time direction that it can be treated as infinite to a very good approximation, so we are interested in corrections due to finite spatial volume only. They are estimated by replacing infinite-volume integrals in the chiral expression with finite sums over the spatial momentum.

Including finite volume corrections in the chiral expressions and redoing the fits reveals negligible errors, $<$0.1%.

### G. Tuning of the tadpole parameter $u_0$

The tadpole improvement factor $u_0$ is a parameter of the gauge and asqtad staggered (sea) quark action and is determined from the fourth root of the average plaquette. The tadpole improvement factor also enters into the valence light and heavy quark actions. On the $a \approx 0.09$ fm ensembles, the valence quarks are generated with the same values of $u_0$ as the sea. However, on the $a \approx 0.12$ fm ensembles, the valence quark actions use values of $u_0$ obtained from the average link in Landau gauge instead. The differences between the values of $u_0$ obtained with the two methods is around 3–4%.

The effect on $f_{B_s}/f_B$ of the mismatch between $u_0$ values in the valence and the sea sectors of the $a \approx 0.12$ fm ensembles was estimated to be $<$0.1% in Ref [59]. Since this is much smaller than the errors due to statistics, chiral fits, and continuum and chiral extrapolation, we take this estimate as our error on $\xi$.

### VII. DISCUSSION OF RESULTS AND FUTURE IMPROVEMENTS

The error budget for the $SU(3)$-breaking mixing parameter $\xi$ described in the previous sections is summarized

TABLE VII. Heavy-quark discretization errors given in percentages.

| Contribution | $f_i$ | $z_i$ | Error $f_B\sqrt{B_B}$ (%) (coarse,fine) | Error $\xi$ (%) (coarse,fine) |
|---|---|---|---|---|
| $\mathcal{O}(a^2)$ Lagrangian | $f_E$ | 2 | (0.28,0.16) | |
| $\mathcal{O}(a_s a)$ Lagrangian | $f_B$ | 2 | (0.96,0.58) | |
| $\mathcal{O}(a^2)$ Operator | $f_X$ | 4 | (1.29,0.74) | |
| | $f_Y$ | 2 | (0.23,0.18) | |
| $\mathcal{O}(a_s a)$ Operator | $f_3$ | 3 | (1.32,0.75) | |
| Total error | | | (2.1,1.3) | (0.2,0.1) |

TABLE VIII. Complete error budget and total error for the $B^0$ mixing parameter $\xi$. All errors are given in percentages.

| Source of uncertainty | Error (%) |
|---|---|
| Statistics ⊕ light-quark disc. ⊕ chiral extrapolation | 3.7 |
| Mixing with wrong-spin operators | 3.2 |
| Heavy-quark discretization | 0.3 |
| Scale uncertainty ($r_1$) | 0.2 |
| Light-quark masses | 0.5 |
| One-loop matching | 0.5 |
| Tuning $\kappa_b$ | 0.4 |
| Finite volume | 0.1 |
| Mistuned coarse $u_0$ | 0.1 |
| Total error | 5.0 |

in Table VIII. For the first error in the table we prefer not to attempt to disentangle the statistical, light-quark discretization, and chiral extrapolation errors since, as explained in Sec. VI D, the lack of knowledge about the LECs at NNLO makes a reliable separation impossible. Our final result is

$$\xi = 1.268 \pm 0.063. \qquad (7.1)$$

The total uncertainty is dominated by the combined statistical, light-quark discretization, and chiral extrapolation error, and the uncertainty associated with the wrong spin operators in the chiral-continuum extrapolation.

Combining our result in Eq. (7.1) with the averages of the experimentally measured values of the mass differences $\Delta M_d = (0.507 \pm 0.004)$ ps$^{-1}$ [16] and $\Delta M_s = (17.69 \pm 0.08)$ ps$^{-1}$ [105], and the meson masses $M_{B_s^0} = (5366.0 \pm 0.9)$ MeV and $M_{B_d^0} = (5279.5 \pm 0.5)$ MeV [16], we quote a value for the ratio of the CKM matrix elements

$$\left| \frac{V_{td}}{V_{ts}} \right| = 0.216 \pm 0.011, \qquad (7.2)$$

assuming no new physics in $B_{(s)}^0$-$\bar{B}_{(s)}^0$ mixing. The error includes the uncertainties in the $B$-meson masses and mass differences but it is strongly dominated by the error in $\xi$.

We can also take our result for $\xi$ and combine it with the value of the decay constant ratio $f_{B_s}/f_{B_d} = 1.229 \pm 0.026$ calculated by our collaboration [59] to determine the ratio of bag parameters

$$\frac{B_{B_s}}{B_{B_d}} = \xi^2 \left( \frac{f_{B_d}}{f_{B_s}} \right)^2 = 1.06 \pm 0.11. \qquad (7.3)$$

The two results for $\xi$ and $f_{B_s}/f_{B_d}$ are correlated, but the statistical analyses were done independently so we cannot include the correlations in calculating the uncertainty in the ratio of bag parameters. Therefore, the error shown in the result of Eq. (7.3) is overestimated. However, as part of the future work, we plan to perform a common analysis of matrix elements and decay constants, from which we will

be able to account for correlations in extracting the value of the bag parameters and thus greatly reduce the error in Eq. (7.3). We will do the same for the individual bag parameters corresponding to all the operators in the basis in Eq. (1.3).

Our result for $\xi$ in Eq. (7.1) is in good agreement with the HPQCD value obtained in Ref. [32], $\xi = 1.258(33)$. Note, however, that HPQCD did not estimate the effects of the wrong spin operators that appear in the complete NLO chiral expression, so the full error in their result may be somewhat bigger than what was quoted. The agreement of these two determinations of $\xi$ provides an excellent check of the methodology and systematic error study in both analyses. In addition, it helps to increase the confidence in the robustness of lattice results for a parameter of great importance in phenomenological studies. In this article, we have established and tested the methodology to apply to broader studies of $B^0$ mixing with the same lattice formulations for light and bottom quarks as used here.

Statistical errors could be reduced significantly by expanding the analysis to include the full set of available configurations (approximately 2000) at each of the $a \approx 0.12$ fm and $a \approx 0.09$ fm ensembles. The current runs of our collaboration on the extended ensembles are also implementing sources located at a random spatial and time location to reduce further the statistical errors. We expect a reduction of the statistical errors by about a factor of two.

The other dominant error of our calculation, the omission in the rHMS$\chi$PT analysis of terms generated by wrong spin operators, will be eliminated when a complete analysis is done with the full rHMS$\chi$PT expressions [86]. A result for $\xi$ that properly includes the wrong spin terms requires the calculation of the continuum matrix elements not only of the operator $\mathcal{O}_1$ as we have done in this work, but also of $\mathcal{O}_2$ and $\mathcal{O}_3$, and simultaneous chiral and continuum extrapolations of all three matrix elements.

The discretization errors, related to both heavy and light quarks, will be reduced in a straightforward way by simulations at smaller lattice spacing, i.e., on the $a \approx 0.06$ fm and $a \approx 0.045$ fm MILC lattices. The reduction of both statistical and discretization errors will also yield cleaner and more accurate continuum and chiral extrapolations. Including data at smaller lattice spacings will also reduce the uncertainty associated with the perturbative matching from the reduction of $\alpha_s = \alpha_V(2/a)$. Although not relevant for the reduction of the total error in $\xi$, this will be important in the determination of the matrix elements $\langle \mathcal{O}_i \rangle$ themselves.

Similarly, although the uncertainty associated with heavy-quark discretization effects is a subdominant source of error in the determination of $\xi$, it is one of the main errors in the determination of $\langle \mathcal{O}_i \rangle$ [37]. In order to have a more reliable, data-driven estimation of these effects, in our on-going analyses we plan to employ the strategy used in Ref. [59], in which terms like the ones in Eq. (6.2) are

included in the chiral-continuum extrapolation fitting form with free parameters to be determined from the fit.

Our new analysis, which incorporates the improvements mentioned above, includes the study of the matrix elements of all five operators that contribute to $H_{\text{eff}}^{\Delta B=2}$ [29]. This will allow not only the precise SM determination of $\Delta M_{s,d}$, $\Delta\Gamma_{s,d}$, and $\xi$, but will also provide the nonperturbative inputs needed to put constraints on BSM models using experimental data on $B^0$ mixing and related observables.

## ACKNOWLEDGMENTS

## APPENDIX A: STAGGERED CHIRAL PERTURBATION THEORY FOR $B^0$-$\bar{B}^0$ MIXING

In this appendix we describe the functional form we use in the chiral and continuum extrapolation of the matrix elements $\langle \bar{B}_q^0 | \mathcal{O}_1^q | B_q^0 \rangle$. Further discussion, as well as complete NLO rHMS$\chi$PT expressions for $\langle \bar{B}_q^0 | \mathcal{O}_1^q | B_q^0 \rangle$ with $i = 1, \ldots, 5$ and the corresponding bag parameters can be found in Ref. [86].

At NLO in rHMS$\chi$PT and at first order in the heavy-quark expansion we use

$$\langle \bar{B}_q^0 | \mathcal{O}_1^q | B_q^0 \rangle = \alpha \left( 1 + \frac{\mathcal{W}_{q\bar{b}} + \mathcal{W}_{b\bar{q}}}{2} + \mathcal{T}_q + \mathcal{Q}_q \right) + L_v m_q + L_s (2m_l + m_h) + L_a a^2. \quad \text{(A1)}$$

$\alpha$, $L_v$, $L_s$, and $L_a$ are constants to be determined from the fits to lattice data. The quantities in script for the partially quenched $2 + 1$ ($m_u = m_d \neq m_s$) case are

$$\mathcal{W}_{q\bar{b}} = \mathcal{W}_{b\bar{q}}$$
$$= \frac{ig_{B^*B\pi}^2}{f_\pi^2} \left\{ \frac{1}{16} \sum_{\mathcal{S},\rho} N_\rho \mathcal{H}_{\mathcal{S}q,\rho}^{\Delta^* + \delta_{\mathcal{S}q}} + \frac{1}{3} \left[ R_{X_I}^{[2,2]}(\{M_{X_I}^{(5)}\}; \{\mu_I\}) \frac{\partial \mathcal{H}_{X,I}^{\Delta^*}}{\partial m_{X_I}^2} - \sum_{j \in \{M_I^{(5)}\}} D_{j,X_I}^{[2,2]}(\{M_{X_I}^{(5)}\}; \{\mu_I\}) \mathcal{H}_{j,I}^{\Delta^*} \right] \right.$$
$$\left. + a^2 \delta_V' \left[ R_{X_V}^{[3,2]}(\{M_{X_V}^{(7)}\}; \{\mu_V\}) \frac{\partial \mathcal{H}_{X,V}^{\Delta^*}}{\partial m_{X_V}^2} - \sum_{j \in \{M_V^{(7)}\}} D_{j,X_V}^{[3,2]}(\{M_{X_V}^{(7)}\}; \{\mu_V\}) \mathcal{H}_{j,V}^{\Delta^*} \right] + (V \to A) \right\}, \quad \text{(A2)}$$

$$\mathcal{T}_q = \frac{-i}{f_\pi^2} \left\{ \frac{1}{16} \sum_{\mathcal{S},\rho} N_\rho \mathcal{I}_{q\mathcal{S},\rho} + \frac{1}{16} \sum_\rho N_\rho \mathcal{I}_{X,\rho} + \frac{2}{3} \left[ R_{X_I}^{[2,2]}(\{M_{X_I}^{(5)}\}; \{\mu_I\}) \left( \frac{\partial \mathcal{I}_{X_I}}{\partial m_{X_I}^2} \right) - \sum_{j \in \{M_I^{(5)}\}} D_{j,X_I}^{[2,2]}(\{M_{X_I}^{(5)}\}; \{\mu_I\}) \mathcal{I}_j \right] \right.$$
$$\left. + a^2 \delta_V' \left[ R_{X_V}^{[3,2]}(\{M_{X_V}^{(7)}\}; \{\mu_V\}) \left( \frac{\partial \mathcal{I}_{X_V}}{\partial m_{X_V}^2} \right) - \sum_{j \in \{M_V^{(7)}\}} D_{j,X_V}^{[3,2]}(\{M_{X_V}^{(7)}\}; \{\mu_V\}) \mathcal{I}_j \right] + (V \to A) \right\}, \quad \text{(A3)}$$

$$\mathcal{Q}_q = \frac{-ig_{B^*B\pi}^2}{f_\pi^2} \left\{ \frac{1}{16} \sum_\rho N_\rho \mathcal{H}_{X,\rho}^{\Delta^*} + \frac{1}{3} \left[ R_{X_I}^{[2,2]}(\{M_{X_I}^{(5)}\}; \{\mu_I\}) \left( \frac{\partial \mathcal{H}_{X_I}^{\Delta}}{\partial m_{X_I}^2} \right) - \sum_{j \in \{M_I^{(5)}\}} D_{j,X_I}^{[2,2]}(\{M_{X_I}^{(5)}\}; \{\mu_I\}) \mathcal{H}_j^{\Delta} \right] \right\}. \quad \text{(A4)}$$

In the equations above, the index $\rho$ runs over the taste representation $(P, A, T, V, I)$ with degeneracies $N_\rho$ ($N_\rho = 1, 4, 6, 4, 1$, respectively), and $\mathcal{S}$ runs over the sea flavors $u$, $d$, $s$. The meson $X$ is made of two light valence quarks $q$, and $m_X$ is its mass. The functions $\mathcal{H}$ and $\mathcal{I}$ are the integrals defined in Appendix A of Ref. [85]. The subscripts on those functions label the flavor and taste of the meson masses at which they are evaluated.

The superscript in $\mathcal{H}$ is the second argument for that function as defined in Ref. [85]. In addition to the hyperfine splitting $\Delta^* = M_{B^*} - M_B$, it includes a light flavor splitting $\delta_{Sq}$ whenever the light flavor of the vector meson in the loop is different from the external flavor.

The splitting is $\delta_{Sq} \equiv M_{B^0 S} - M_{B_q^0} = 2\lambda_1 \mu(m_S - m_q)$, where $\lambda_1$ and $\mu$ are low-energy constants. The constant $\lambda_1$ comes from heavy quark effective theory, and $\mu$ is defined in Eq. (5.2).

The residue functions $R_j^{[n,k]}$ and $D_{j,l}^{[n,k]}$ in the expressions above are defined by Ref. [63]

$$R_j^{[n,k]}(\{m\}, \{\mu\}) \equiv \frac{\prod_{a=1}^{k}(\mu_a^2 - m_j^2)}{\prod_{i \neq j}(m_i^2 - m_j^2)},$$
$$D_{j,l}^{[n,k]}(\{m\}, \{\mu\}) \equiv -\frac{d}{dm_l^2} R_j^{[n,k]}(\{m\}, \{\mu\}). \tag{A5}$$

The mass combinations appearing as arguments of these functions in the $2 + 1$ partially quenched theory are

$$\{M_X^{(5)}\} \equiv \{m_\eta, m_X\},$$
$$\{M_X^{(7)}\} \equiv \{m_\eta, m_{\eta'}, m_X\}, \tag{A6}$$
$$\{\mu\} \equiv \{m_L, m_H\},$$

where $m_L$ is the meson mass made from $l\bar{l}$ sea quarks, and $m_H$ is the meson mass made from $h\bar{h}$ sea quarks. The tastes of these mesons are indicated explicitly in the equations above.

Since we are not including the effects of the hyperfine splitting $\Delta^*$ or the light flavor splittings $\delta_k$ in this work, the functions $\mathcal{H}$ and $\mathcal{I}$ appearing in the wave function, tadpole, and sunset contributions simplify to

$$i\mathcal{H}_{k,\Xi}^0 = -3i\mathcal{I}_{k,\Xi} = -\frac{3}{16\pi^2} M_{k,\Xi}^2 \ln\left(\frac{M_{k,\Xi}^2}{\Lambda_\chi^2}\right). \tag{A7}$$

## APPENDIX B: FUNCTIONS PARAMETRIZING HEAVY-QUARK DISCRETIZATION ERRORS

In this appendix we collect the functions $f_i$ needed in Eq. (6.2) to estimate the heavy-quark discretization errors affecting our calculation. For details on the origin of these functions and the effects of higher-dimension operators in the Lagrangian, see Ref. [102]. For further details on the application to the estimation of heavy-quark discretization errors in $B^0$-$\bar{B}^0$ mixing, see Ref. [77].

(i) $O(a^2)$ errors from the Lagrangian:

$$f_E(m_0 a) = \frac{1}{2}\left[\frac{(1 + m_0 a) - 1}{m_0 a(2 + m_0 a)(1 + m_0 a)} - \frac{1}{4(1 + m_0 a)^2}\right]. \tag{B1}$$

(ii) $O(\alpha_s a)$ errors from the Lagrangian:

$$f_B(m_0 a) = \frac{\alpha_s}{2(1 + m_0 a)}. \tag{B2}$$

(iii) $O(a^2)$ errors from the four-fermion operator:

$$f_X(m_0 a) = \frac{1}{2}\left[\frac{1}{2(1 + m_0 a)} - \left(\frac{m_0 a}{2(2 + m_0 a)(1 + m_0 a)}\right)^2\right] \tag{B3}$$
$$f_Y(m_0 a) = \frac{2 + 4m_0 a + (m_0 a)^2}{4(1 + m_0 a)^2(2 + m_0 a)^2}.$$

(iv) $O(\alpha_s a)$ errors from the four-fermion operator:

$$f_3(m_0 a) = \frac{\alpha_s}{2(2 + m_0 a)}. \tag{B4}$$

## APPENDIX C: PRIOR CENTRAL VALUES AND WIDTHS FOR THE CORRELATOR FITS

In Table IX we collect the prior central values and widths used in the correlator fits described in Sec. IV. The amplitude parameters are defined in Eqs. (2.21) and (2.22), and the energy differences are defined as $\Delta E_{i+1,i} \equiv a(E_{i+1} - E_1)$.

TABLE IX. The priors with index 0 refer to the ground state. Superscripts $d$ and $1S$ refer to the local and $1S$ smeared sources, respectively. Higher-energy state priors have indices $i$ and $j$. The prime in $E_i'$ refers to an opposite parity (oscillating) state.

|  | Prior central value | Prior width |
|---|---|---|
| $Z_0^{1S}$ | 2.2 | 0.5 |
| $Z_i^{1S}$ | 0.01 | 0.5 |
| $Z_0^d$ | 0.45 | 0.45 |
| $Z_i^d$ | 0.01 | 1 |
| $O_{00}$ | 0.01 | 0.02 |
| $O_{ij}$ | 0.01 | 0.1 |
| $E_0$ (0.12 fm) | 1.95 | 0.15 |
| $E_0'$ (0.12 fm) | 2.25 | 0.15 |
| $E_0$ (0.09 fm) | 1.65 | 0.15 |
| $E_0'$ (0.09 fm) | 1.85 | 0.15 |
| $\log\Delta E_{i+1,i}$ | $-1.5$ | 0.5 |
| $\log\Delta E_{i+1,i}'$ | $-1.5$ | 0.5 |

[1] E. Lunghi and A. Soni, Phys. Lett. B **697**, 323 (2011).
[2] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **105**, 081801 (2010).
[3] J. Laiho, E. Lunghi, and R. Van De Water, Proc. Sci., FPCP2010 (2010) 040.
[4] A. Lenz *et al.*, arXiv:1203.0238.
[5] E. Lunghi and A. Soni, Phys. Rev. Lett. **104**, 251802 (2010).
[6] A. Lenz *et al.*, Phys. Rev. D **83**, 036004 (2011).
[7] A. J. Bevan *et al.* (UTfit Collaboration), Proc. Sci., ICHEP2010 (2010) 270.
[8] A. Lenz and U. Nierste, J. High Energy Phys. 06 (2007) 072.
[9] M. Bona *et al.* (UTfit Collaboration), PMC Phys. A **3**, 6 (2009).
[10] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **85**, 072002 (2012).
[11] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **85**, 032006 (2012).
[12] K. Anikeev *et al.*, arXiv:hep-ph/0201071.
[13] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **97**, 242003 (2006).
[14] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **97**, 021802 (2006).
[15] R. Aaij *et al.* (LHCb Collaboration), Phys. Lett. B **709**, 177 (2012).
[16] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).
[17] M. Bona *et al.* (UTfit Collaboration), J. High Energy Phys. 03 (2008) 049.
[18] B. A. Dobrescu and G. Z. Krnjaic, Phys. Rev. D **85**, 075020 (2012).
[19] W. Altmannshofer and M. Carena, Phys. Rev. D **85**, 075006 (2012).
[20] A. J. Buras and J. Girrbach, J. High Energy Phys. 03 (2012) 052.
[21] M. Blanke, A. J. Buras, K. Gemmler, and T. Heidsieck, J. High Energy Phys. 03 (2012) 024.
[22] A. J. Buras, M. Nagai, and P. Paradisi, J. High Energy Phys. 05 (2011) 005.
[23] A. J. Buras, K. Gemmler, and G. Isidori, Nucl. Phys. **B843**, 107 (2011).
[24] A. J. Buras, B. Duling, T. Feldmann, T. Heidsieck, C. Promberger, and S. Recksiegel, J. High Energy Phys. 09 (2010) 106.
[25] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).
[26] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, Phys. Rev. D **50**, 3433 (1994).
[27] F. Gabbiani, E. Gabrielli, A. Masiero, and L. Silvestrini, Nucl. Phys. **B477**, 321 (1996).
[28] D. Becirevic, V. Gimenez, G. Martinelli, M. Papinutto, and J. Reyes, J. High Energy Phys. 04 (2002) 025.
[29] C. M. Bouchard *et al.* (Fermilab Lattice and MILC Collaborations), Proc. Sci., LATTICE2011 (2011) 274.
[30] T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981); **65**, 1772(E) (1981).
[31] A. J. Buras, M. Jamin, and P. H. Weisz, Nucl. Phys. **B347**, 491 (1990).
[32] E. Gámiz, C. T. H. Davies, G. P. Lepage, J. Shigemitsu, and M. Wingate (HPQCD Collaboration), Phys. Rev. D **80**, 014503 (2009).

[33] C. Albertus *et al.* (RBC and UKQCD Collaborations), Phys. Rev. D **82**, 014505 (2010).
[34] R. T. Evans, A. X. El-Khadra, and M. Di Pierro (Fermilab Lattice and MILC Collaborations), Proc. Sci., LAT2006 (2006) 081.
[35] R. Evans, E. Gámiz, A. X. El-Khadra, and M. Di Pierro (Fermilab Lattice and MILC Collaborations), Proc. Sci., LAT2007 (2007) 354.
[36] R. T. Evans *et al.* (Fermilab Lattice and MILC Collaborations), Proc. Sci., LAT2008 (2008) 052.
[37] R. T. Evans *et al.* (Fermilab Lattice and MILC Collaborations), Proc. Sci., LAT2009 (2009) 245.
[38] G. P. Lepage, L. Magnea, C. Nakhleh, U. Magnea, and K. Hornbostel, Phys. Rev. D **46**, 4052 (1992).
[39] B. Sheikholeslami and R. Wohlert, Nucl. Phys. **B259**, 572 (1985).
[40] A. X. El-Khadra, A. S. Kronfeld, and P. B. Mackenzie, Phys. Rev. D **55**, 3933 (1997).
[41] E. Golowich, J. Hewett, S. Pakvasa, and A. A. Petrov, Phys. Rev. D **76**, 095009 (2007).
[42] C. W. Bernard, T. Burch, K. Orginos, D. Toussaint, T. A. DeGrand, C. DeTar, S. Datta, S. Gottlieb, U. M. Heller, and R. Sugar (MILC Collaboration), Phys. Rev. D **64**, 054506 (2001).
[43] M. G. Alford, W. Dimm, G. P. Lepage, G. Hockney, and P. B. Mackenzie, Phys. Lett. B **361**, 87 (1995).
[44] M. Luscher and P. Weisz, Phys. Lett. **158B**, 250 (1985).
[45] S. Prelovsek, Phys. Rev. D **73**, 014506 (2006).
[46] C. Bernard, M. Golterman, and Y. Shamir, Phys. Rev. D **73**, 114511 (2006).
[47] C. Bernard, Phys. Rev. D **73**, 114503 (2006).
[48] Y. Shamir, Phys. Rev. D **71**, 034509 (2005).
[49] Y. Shamir, Phys. Rev. D **75**, 054503 (2007).
[50] C. Bernard, M. Golterman, and Y. Shamir, Phys. Rev. D **77**, 074505 (2008).
[51] S. R. Sharpe, Proc. Sci., LAT2006 (2006) 022.
[52] A. S. Kronfeld, Proc. Sci., LAT2007 (2007) 016.
[53] M. Golterman, Proc. Sci., CONFINEMENT8 (2008) 014.
[54] G. C. Donald, C. T. H. Davies, E. Follana, and A. S. Kronfeld, Phys. Rev. D **84**, 054504 (2011).
[55] C. Bernard *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **83**, 034503 (2011).
[56] J. A. Bailey *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **79**, 054507 (2009).
[57] C. Bernard *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **79**, 014506 (2009).
[58] J. A. Bailey *et al.* (Fermilab Lattice and MILC Collaborations), Proc. Sci., LATTICE2010 (2010) 311.
[59] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **85**, 114506 (2012).
[60] J. A. Bailey *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **85**, 114502 (2012).
[61] W. Lee and S. Sharpe, Phys. Rev. D **60**, 114503 (1999).
[62] C. Aubin and C. Bernard, Phys. Rev. D **68**, 034014 (2003).
[63] C. Aubin and C. Bernard, Phys. Rev. D **68**, 074011 (2003).
[64] S. R. Sharpe and R. S. Van de Water, Phys. Rev. D **71**, 114505 (2005).
[65] C. Aubin and C. Bernard, Phys. Rev. D **73**, 014515 (2006).
[66] A. Bazavov *et al.* (MILC Collaborations), Rev. Mod. Phys. **82**, 1349 (2010).
[67] G. P. Lepage, Phys. Rev. D **59**, 074502 (1999).

[68] K. Symanzik, Nucl. Phys. **B226**, 205 (1983).
[69] G. P. Lepage and P. B. Mackenzie, Phys. Rev. D **48**, 2250 (1993).
[70] Z. Hao, G. M. von Hippel, R. R. Horgan, Q. J. Mason, and H. D. Trottier, Phys. Rev. D **76**, 034507 (2007).
[71] M. Wingate, J. Shigemitsu, C. T. H. Davies, G. P. Lepage, and H. D. Trottier, Phys. Rev. D **67**, 054505 (2003).
[72] C. R. Allton, arXiv:hep-lat/9610016v1.
[73] L. Susskind, Phys. Rev. D **16**, 3031 (1977); H. S. Sharatchandra, H. J. Thun, and P. Weisz, Nucl. Phys. **B192**, 205 (1981).
[74] A. S. Kronfeld, Phys. Rev. D **62**, 014505 (2000).
[75] G. P. Lepage and P. B. Mackenzie, Phys. Rev. D **48**, 2250 (1993).
[76] E. Gámiz, J. Shigemitsu, and H. Trottier, Phys. Rev. D **77**, 114505 (2008).
[77] E. Gámiz, A. El-Khadra, and A. Kronfeld (unpublished).
[78] Q. Mason, H. D. Trottier, C. T. H. Davies, K. Foley, A. Gray, G. P. Lepage, M. Nobes, and J. Shigemitsu (HPQCD and UKQCD Collaborations), Phys. Rev. Lett. **95**, 052002 (2005).
[79] S. J. Brodsky, G. P. Lepage, and P. B. Mackenzie, Phys. Rev. D **28**, 228 (1983).
[80] K. Hornbostel, G. P. Lepage, and C. Morningstar, Phys. Rev. D **67**, 034023 (2003).
[81] J. L. Richardson, Phys. Lett. **82B**, 272 (1979).
[82] M. Di Pierro, A. X. El-Khadraa, S. Gottlieb, A. S. Kronfeld, P. B. Mackenzie, D. P. Menscher, M. B. Oktay, and J. N. Simone, Nucl. Phys. B, Proc. Suppl. **119**, 586 (2003).
[83] G. P. Lepage, B. Clark, C. T. H. Davies, K. Hornbostel, P. B. Mackenzie, C. Morningstar, H. Trottier, Nucl. Phys. B, Proc. Suppl. **106**, 12 (2002).
[84] C. Morningstar, Nucl. Phys. B, Proc. Suppl. **109**, 185 (2002).
[85] W. Detmold and C. J. D. Lin, Phys. Rev. D **76**, 014501 (2007).
[86] C. Bernard (unpublished).

[87] C. Aubin, C. Bernard, C. DeTar, J. Osborn, S. Gottlieb, E. B. Gregory, D. Toussaint, U. M. Heller, J. E. Hetrick, and R. Sugar (MILC Collaboration), Phys. Rev. D **70**, 114501 (2004).
[88] C. W. Bernard, Phys. Rev. D **62**, 034503 (2000).
[89] R. Sommer, Nucl. Phys. **B411**, 839 (1994).
[90] A. Bazavov et al. (MILC Collabortion), Proc. Sci., CD09 (2009) 007.
[91] C. T. H. Davies, E. Follana, I. D. Kendall, G. P. Lepage, and C. McNeile (HPQCD Collaboration), Phys. Rev. D **81**, 034506 (2010).
[92] I. W. Stewart, Nucl. Phys. **B529**, 62 (1998).
[93] R. Casalbuoni, A. Deandrea, N. Di Bartolomeo, R. Gatto, F. Feruglio, and G. Nardulli, Phys. Rep. **281**, 145 (1997).
[94] A. Anastassov et al. (CLEO Collaboration), Phys. Rev. D **65**, 032003 (2002).
[95] A. Abada, D. Becirevic, P. Boucaud, G. Herdoiza, J. P. Leroy, A. Le Yaouanc, O. Pene, and J. Rodriguez-Quintero, Nucl. Phys. B, Proc. Suppl. **119**, 641 (2003).
[96] M. C. Arnesen, B. Grinstein, I. Z. Rothstein, and I. W. Stewart, Phys. Rev. Lett. **95**, 071802 (2005).
[97] H. Ohki, H. Matsufuru, and T. Onogi, Phys. Rev. D **77**, 094509 (2008).
[98] J. Bulava et al. (ALPHA Collaboration), Proc. Sci., LATTICE2010 (2010) 303.
[99] W. Detmold, C.-J. D. Lin, and S. Meinel, Phys. Rev. D **85**, 114508 (2012).
[100] C. Bernard, Proc. Sci., LAT2007 (2007) 090.
[101] A. S. Kronfeld, Nucl. Phys. B, Proc. Suppl. **129**, 46 (2004).
[102] M. B. Oktay and A. S. Kronfeld, Phys. Rev. D **78**, 014504 (2008).
[103] W. Kilian and T. Mannel, Phys. Lett. B **301**, 382 (1993).
[104] C. Bernard (MILC Collaboration), Phys. Rev. D **65**, 054031 (2002).
[105] Average of the CDF and LHCb results by the Heavy Flavor Averaging Group http://www.slac.stanford.edu/xorg/hfag/.