B_{s}→Kℓν form factors from lattice QCD

C. M. Bouchard, G. Peter Lepage, Christopher Monahan, Heechang Na, and Junko Shigemitsu

$B_s \to K\ell\nu$ Form Factors from Lattice QCD

C.M. Bouchard,[1, *] G. Peter Lepage,[2] Christopher Monahan,[3] Heechang Na,[4] and Junko Shigemitsu[1]

(HPQCD Collaboration)

[1] *Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[2] *Laboratory of Elementary Particle Physics, Cornell University, Ithaca, NY 14853, USA*
[3] *Physics Department, College of William and Mary, Williamsburg, Virginia 23187, USA*
[4] *Department Physics and Astronomy, University of Utah, Salt Lake City, Utah 84112, USA*

We report the first lattice QCD calculation of the form factors for the standard model tree-level decay $B_s \to K\ell\nu$. In combination with future measurement, this calculation will provide an alternative exclusive semileptonic determination of $|V_{ub}|$. We compare our results with previous model calculations, make predictions for differential decay rates and branching fractions, and predict the ratio of differential branching fractions between $B_s \to K\tau\nu$ and $B_s \to K\mu\nu$. We also present standard model predictions for differential decay rate forward-backward asymmetries, polarization fractions, and calculate potentially useful ratios of $B_s \to K$ form factors with those of the fictitious $B_s \to \eta_s$ decay. Our lattice simulations utilize NRQCD b and HISQ light quarks on a subset of the MILC $2+1$ asqtad gauge configurations, including two lattice spacings and a range of light quark masses.

I. INTRODUCTION

The decay $B_s \to K\ell\nu$ occurs at tree-level in the standard model via the flavor-changing charged-current $b \to u$ transition, making it an alternative to $B \to \pi\ell\nu$ in the determination of $|V_{ub}|$ from exclusive semileptonic decays. The difference in these processes, a spectator strange quark in $B_s \to K\ell\nu$ *vs.* a spectator down quark in $B \to \pi\ell\nu$, is beneficial for lattice QCD simulations, because it improves the ratio of signal to noise. Though this process has not yet been observed, its measurement is planned at LHCb and is possible during an $\Upsilon(5S)$ run at BelleII. This provides a prediction opportunity for lattice QCD.

In addition to the calculation of form factors for $B_s \to K$, we also calculate their ratios with form factors for the fictitious $B_s \to \eta_s$ decay. Such ratios are essentially free of our largest systematic error, perturbative matching. In combination with a future calculation of $B_s \to \eta_s$ using a HISQ b quark, these ratios would yield a non-perturbative evaluation of the matching factor for the $b \to u$ current with NRQCD b quark. This matching factor would be applicable to $B_s \to K\ell\nu$ and $B \to \pi\ell\nu$ simulations using NRQCD b quarks.

To include correlations among the data for both decays, correlation function fits must include vast amounts of correlated data. To make such fits feasible, we have developed a new technique, called chaining, discussed in Appendix A. In addition, the use of marginalization techniques developed in Ref. [1] significantly reduces the time required for the fits.

The chiral, continuum, and kinematic extrapolations are performed simultaneously using the modified z-expansion [2, 3] with the chiral logarithmic corrections fixed by the results of hard pion chiral perturbation theory (HPChPT) [4, 5]. The factorization of chiral corrections and kinematics, as found at one-loop order by HPChPT, suggests the modified z-expansion is a natural choice for carrying out this simultaneous extrapolation. We refer to the combination of HPChPT chiral logarithmic corrections and the modified z-expansion as the HPChPT z-expansion.

II. FORM FACTORS AND MATRIX ELEMENTS

The vector hadronic matrix element is parameterized by the scalar and vector form factors $f_{0,+}$

$$
\langle K|V^\mu|B_s \rangle = f_+ \left(p^\mu_{B_s} + p^\mu_K - \frac{M^2_{B_s} - M^2_K}{q^2} q^\mu \right) + f_0 \frac{M^2_{B_s} - M^2_K}{q^2} q^\mu, \tag{1}
$$

where $V^\mu = \bar{u}\gamma^\mu b$ and $q^\mu = p^\mu_{B_s} - p^\mu_K$. At intermediate stages of the calculation we recast $f_{0,+}$ in terms of the more convenient form factors $f_{\parallel,\perp}$

$$
\langle K|V^\mu|B_s \rangle = \sqrt{2M_{B_s}} \left(\frac{p^\mu_{B_s}}{M_{B_s}} f_\parallel + p^\mu_\perp f_\perp \right), \tag{2}
$$

where $p^\mu_\perp = p^\mu_K - p^\mu_{B_s}(p_K \cdot p_{B_s})/M^2_{B_s}$. In the B_s meson rest frame, the form factors $f_{\parallel,\perp}$ are simply related to the temporal and spatial components of the hadronic vector matrix elements,

$$
\langle K|V^0|B_s \rangle = \sqrt{2M_{B_s}} f_\parallel, \tag{3}
$$

$$
\langle K|V^k|B_s \rangle = \sqrt{2M_{B_s}} p^k_K f_\perp. \tag{4}
$$

*bouchard.18@osu.edu

The scalar and vector form factors are related to $f_{\parallel,\perp}$ by

$$f_0 \;=\; \frac{\sqrt{2M_{B_s}}}{M_{B_s}^2 - M_K^2}\left[(M_{B_s} - E_K)f_\parallel + \mathbf{p}_K^2 f_\perp\right], \quad (5)$$

$$f_+ \;=\; \frac{1}{\sqrt{2M_{B_s}}}\left[f_\parallel + (M_{B_s} - E_K)f_\perp\right], \quad (6)$$

where $\mathbf{p}_K$ is the kaon three-momentum. This discussion generalizes in a straightforward way for the $B_s \to \eta_s$ matrix elements.

III. SIMULATION

Ensemble averages are performed with the MILC $2+1$ asqtad gauge configurations [6] listed in Table I. Valence quarks in our simulation are nonrelativistic QCD (NRQCD) [7] b quarks, tuned in Ref. [8], and highly improved staggered (HISQ) [9] light and s quarks, the propagators for which were generated in Refs. [2, 3]. Valence quark masses for each ensemble used in the simulations are collected in Table I and correspond to pion masses ranging from, approximately, $260\,\mathrm{MeV}$ to $500\,\mathrm{MeV}$.

Heavy-light B_s meson bilinears $\Phi_{B_s}^\alpha$ are built from NRQCD b and HISQ s quarks (for details see Ref. [8]) and light-light kaon (and similarly for the η_s) bilinears Φ_K are built from HISQ light and s quarks (for details see Ref. [2]). From these bilinears we build two and three point correlation function data

$$C_{B_s}^{\alpha\beta}(t_0,t) \;=\; \frac{1}{L^3}\sum_{\mathbf{x},\mathbf{y}}\langle\Phi_{B_s}^\beta(t,\mathbf{y})\,\Phi_{B_s}^{\alpha\dagger}(t_0,\mathbf{x})\rangle, \quad (7)$$

$$C_{K,\mathbf{p}}(t_0,t) \;=\; \frac{1}{L^3}\sum_{\mathbf{x},\mathbf{y}}e^{i\,\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\langle\Phi_K(t,\mathbf{y})\,\Phi_K^\dagger(t_0,\mathbf{x})\rangle, \quad (8)$$

$$\begin{aligned}C_{J,\mathbf{p}}^\alpha(t_0,t,T) \;=\;& \frac{1}{L^3}\sum_{\mathbf{x},\mathbf{y},\mathbf{z}}e^{i\,\mathbf{p}\cdot(\mathbf{z}-\mathbf{x})}\times\\ &\langle\Phi_K(t_0+T,\mathbf{x})\,J(t,\mathbf{z})\,\Phi_{B_s}^{\alpha\dagger}(t_0,\mathbf{y})\rangle,\end{aligned} \quad (9)$$

where indices α,β specify b quark smearing. We generate data for both a local and Gaussian smeared b quark, with smearing function ϕ introduced via the replacement $\sum_{\mathbf{y}} \to \sum_{\mathbf{y},\mathbf{y}'}\phi(\mathbf{y}'-\mathbf{y})$ in Eqs. (7, 9). Three point and daughter meson two point correlation function data are generated at four daughter meson momenta, corresponding to $\mathbf{p}L \in 2\pi\{(000),(100),(110),(111)\}$. In three point data, these momenta are inserted at $\mathbf{x}$ in Fig. 1. The sum over $\mathbf{x}$ in Eqs. (8, 9) is performed using random wall sources with U(1) phases ξ, *i.e.* $\sum_{\mathbf{x}} \to \sum_{\mathbf{x},\mathbf{x}'}\xi(\mathbf{x})\xi(\mathbf{x}')$. In the three point correlator a B_s meson source is inserted at timeslice t_0, selected at random on each configuration to reduce autocorrelations. The current J is inserted at timeslices t such that $t_0 \leq t \leq t_0+T$ and the daughter meson is annihilated at timeslice t_0+T. Prior to performing the fits, all data are shifted to a
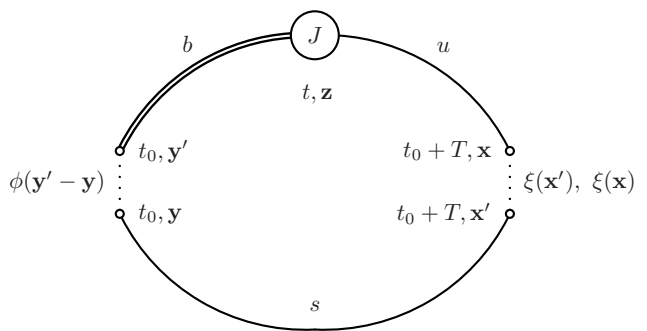

FIG. 1: Setup for three point correlator data generation.

common $t_0 = 0$. This three point correlator setup is depicted in Fig. 1. Additional details regarding the two and three point correlation function generation can be found in Ref. [10].

The flavor-changing current J is an effective lattice vector current $\mathcal{V}_\mu$ corrected through $\mathcal{O}(\alpha_s, \Lambda_{\mathrm{QCD}}/m_b, \alpha_s/(am_b))$. The lattice currents that contribute through this order are

$$\mathcal{V}_\mu^{(0)} \;=\; \overline{\Psi}_u\,\gamma_\mu\,\Psi_b, \quad (10)$$

$$\mathcal{V}_\mu^{(1)} \;=\; -\frac{1}{2am_b}\overline{\Psi}_u\,\gamma_\mu\,\boldsymbol{\gamma}\cdot\boldsymbol{\nabla}\,\Psi_b. \quad (11)$$

Matrix elements of the continuum vector current $\langle V_\mu\rangle$ are matched to those of the lattice vector current according to

$$\langle V_\mu\rangle = (1 + \alpha_s\rho_0^{(V_\mu)})\langle\mathcal{V}_\mu^{(0)}\rangle + \langle\mathcal{V}_\mu^{(1),\mathrm{sub}}\rangle, \quad (12)$$

where

$$\langle\mathcal{V}_\mu^{(1),\mathrm{sub}}\rangle \equiv \langle\mathcal{V}_\mu^{(1)}\rangle - \alpha_s\zeta_{10}^{V_\mu}\langle\mathcal{V}_\mu^{(0)}\rangle. \quad (13)$$

The matching calculation is done to one loop using massless HISQ lattice perturbation theory [11]. In implementing the matching, we omit $\mathcal{O}(\alpha_s\Lambda_{\mathrm{QCD}}/m_b)$ contributions. Ref. [12], which used asqtad valence quarks, found contributions of this order to be negligible. In Ref. [10], which used HISQ valence quarks, these contributions to the temporal component of the vector current were studied and again were found to be negligible. We also omit $\mathcal{O}(\Lambda_{\mathrm{QCD}}/m_b)^2$ relativistic matching corrections. These, and higher order, omitted contributions to the matching result in our leading systematic error. An estimate of this error, and its incorporation in our fit results, is discussed in the following section.

IV. CORRELATION FUNCTION FITS

Two and three point correlation function fit Ansätze, and the selection of priors, closely follows the methods of

ens	$L^3 \times N_t$	r_1/a	$au_0 m_{\text{sea}}$	u_0	N_{conf}	N_{tsrc}	am_u	am_s	am_b	$aE_{b\bar{b}}^{\text{sim}}$
C1	$24^3 \times 64$	2.647(3)	0.005/0.05	0.8678	1200	2	0.0070	0.0489	2.650	0.28356(15)
C2	$20^3 \times 64$	2.618(3)	0.01/0.05	0.8677	1200	2	0.0123	0.0492	2.688	0.28323(18)
C3	$20^3 \times 64$	2.644(3)	0.02/0.05	0.8688	600	2	0.0246	0.0491	2.650	0.27897(20)
F1	$28^3 \times 96$	3.699(3)	0.0062/0.031	0.8782	1200	4	0.00674	0.0337	1.832	0.25653(14)
F2	$28^3 \times 96$	3.712(4)	0.0124/0.031	0.8788	600	4	0.01350	0.0336	1.826	0.25558(28)

TABLE I: Left to right: labels for the ensembles used in this analysis; lattice volume; inverse lattice spacing in r_1-units; light/strange sea-quark masses; tadpole improvement factor $u_0 = \langle \text{plaquette} \rangle^{1/4}$; number of configurations; number of time sources; valence u-quark mass; valence s-quark mass; b-quark mass; and the spin-averaged $b\bar{b}$ ground state energies used to relate our B_s meson simulation energies to their physical values.

Ref. [10]. Two point B_s data are fit to

$$C_{B_s}^{\alpha\beta}(t) = \sum_{n=0}^{N-1} b^{\alpha(n)} b^{\beta(n)\dagger} e^{-E_{B_s}^{\text{sim(n)}}t}$$
$$+ \sum_{m=0}^{\tilde{N}-1} \tilde{b}^{\alpha(m)} \tilde{b}^{\beta(m)\dagger} (-1)^t e^{-\tilde{E}_{B_s}^{\text{sim(m)}}t}, \quad (14)$$

where tildes denote oscillating state contributions and $E_{B_s}^{\text{sim}}$ is the simulated B_s energy. The physical ground state B_s mass is related to the simulation ground state energy by

$$E_{B_s}^{(0)} = E_{B_s}^{\text{sim}(0)} + \frac{1}{2}(M_{b\bar{b}}^{\text{expt}} - E_{b\bar{b}}^{\text{sim}}) \quad (15)$$

where $M_{b\bar{b}}^{\text{expt}} = 9.450(4)$ GeV [13] is adjusted from experiment to remove electromagnetic, η_b annihilation, and charmed sea effects not present in our simulations, and $E_{b\bar{b}}^{\text{sim}}$ is the spin-averaged energy of $b\bar{b}$ states calculated on the ensembles used in the simulation and listed in Table I. The b quark smearing is indicated by indices α, β. Kaon and η_s two point correlator data are fit to an expression of the form[1]

$$C_{\mathbf{p}}(t) = \sum_{n=0}^{N-1} |d_{\mathbf{p}}^{(n)}|^2 \left(e^{-E^{(n)}t} + e^{-E^{(n)}(N_t-t)} \right)$$
$$+ \sum_{m=0}^{\tilde{N}-1} |\tilde{d}_{\mathbf{p}}^{(m)}|^2 (-1)^t \left(e^{-\tilde{E}^{(m)}t} + e^{-\tilde{E}^{(m)}(N_t-t)} \right). \quad (16)$$

Results of two point fits satisfy the dispersion relation and are stable with respect to variations in $(N, \tilde{N})$ and the range of timeslices included in the fits, as demonstrated for kaon two point data in Ref. [10].

_____

[1] The zero momentum η_s has no oscillating state contributions due to mass degeneracy of its valence quarks.

Three point correlation function data are described by

$$C_{J,\mathbf{p}}^{\alpha}(t,T) = \sum_{n,m=0}^{N-1} d_{\mathbf{p}}^{(n)} A_{J,\mathbf{p}}^{(n,m)} b^{\alpha(m)\dagger}$$
$$\times\ e^{-E^{(n)}(T-t)} e^{-E_{B_s}^{\text{sim(m)}}t}$$
$$+ \sum_{n=0}^{N-1} \sum_{m=0}^{\tilde{N}-1} d_{\mathbf{p}}^{(n)} B_{J,\mathbf{p}}^{(n,m)} \tilde{b}^{\alpha(m)\dagger} (-1)^t$$
$$\times\ e^{-E^{(n)}(T-t)} e^{-\tilde{E}_{B_s}^{\text{sim(m)}}t}$$
$$+ \sum_{n=0}^{\tilde{N}-1} \sum_{m=0}^{N-1} \tilde{d}_{\mathbf{p}}^{(n)} C_{J,\mathbf{p}}^{(n,m)} b^{\alpha(m)\dagger} (-1)^{T-t}$$
$$\times\ e^{-\tilde{E}^{(n)}(T-t)} e^{-E_{B_s}^{\text{sim(m)}}t}$$
$$+ \sum_{n,m=0}^{\tilde{N}-1} \tilde{d}_{\mathbf{p}}^{(n)} D_{J,\mathbf{p}}^{(n,m)} \tilde{b}^{\alpha(m)\dagger} (-1)^T$$
$$\times\ e^{-\tilde{E}^{(n)}(T-t)} e^{-\tilde{E}_{B_s}^{\text{sim(m)}}t}, \quad (17)$$

where the three point amplitudes A, B, C, and D are proportional to the hadronic matrix elements. The ground state hadronic matrix element is obtained from $A^{(0,0)}$

$$\frac{4}{\sqrt{2}} A_{J,\mathbf{p}}^{(0,0)} = \frac{a^3 \langle K_{\mathbf{p}}^{(0)} | J | B_s^{(0)} \rangle}{\sqrt{2a^3 E_K^{(0)}} \sqrt{2a^3 E_{B_s}^{(0)}}}, \quad (18)$$

where the factor of $4/\sqrt{2}$ accounts for numerical factors introduced in the simulation and associated with taste averaging and HISQ inversion. In the correlator fits we include data for several temporal separations T between the mother and daughter mesons. On the coarse ensembles we include data for $T = 13, 14, 15$ while for the fine ensembles we include $T = 23, 24$ data.

On each ensemble we perform a simultaneous fit to two and three point correlation function data for the $B_s \to K$ and $B_s \to \eta_s$ decays, at all simulated momenta, including both spatial and temporal currents, and for the temporal separations listed above. This ensures correlations among these data are accounted for in the analysis. However, fits to such large data sets produce unwieldy data covariance matrices and are typically not convergent, or
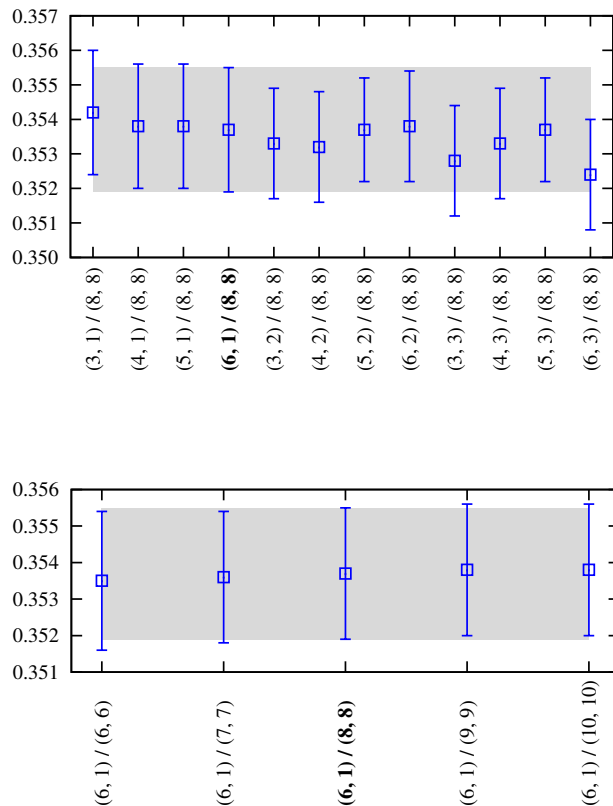
FIG. 2: Chained and marginalized fit results for the ground state amplitude $A_{V_t,(1,0,0)}^{(0,0)}$ of the $B_s \to K$ decay on ensemble F2. Fit results are shown as a function of the number of (*top*) states explicitly included in the fit and (*bottom*) total states accounted for in the fit. Final results are taken from $(6,1)/(8,8)$ fits, represented by gray bands.

require a prohibitively large number of iterations. This can be partially addressed by thinning the data, *e.g.* by the use of SVD cuts, but this reduces the accuracy of the fits.

To address this problem we introduce a technique, which we refer to as chaining, to simplify fits to very large data sets. Consider a data set consisting of N correlators, $data = (correlator_1, correlator_2, \ldots, correlator_N)$. Before the fit, all fit parameters are assigned priors. Chaining first fits $correlator_1$ then uses the best fit mean values and covariances to replace the corresponding priors in subsequent fits. The updated set of priors is then used in the fit to $correlator_2$. In this and all subsequent fits, correlations are accounted for between the data being fit and those priors which are best fit results from previous fits — this is an important step as it prevents "double counting" data. After this 2nd fit, the priors are again updated according to the best fit mean values and covariances. This process is repeated for all correlators. The collection of best fit mean values and covariances following the fit to $correlator_N$ are the final fit results. Chaining is described in greater detail in Appendix A.

ens	$f_0^{B_s K}(000)$	$f_0^{B_s K}(100)$	$f_0^{B_s K}(110)$	$f_0^{B_s K}(111)$
C1	0.8244(23)	0.7081(27)	0.6383(30)	0.5938(41)
C2	0.8427(25)	0.6927(35)	0.6036(49)	0.536(12)
C3	0.8313(29)	0.6953(33)	0.6309(30)	0.5844(46)
F1	0.8322(25)	0.6844(35)	0.5994(43)	0.5551(56)
F2	0.8316(27)	0.6915(38)	0.6119(43)	0.5563(61)
ens		$f_+^{B_s K}(100)$	$f_+^{B_s K}(110)$	$f_+^{B_s K}(111)$
C1		2.087(16)	1.657(14)	1.378(13)
C2		1.880(12)	1.412(16)	1.142(33)
C3		1.773(11)	1.4212(84)	1.184(10)
F1		1.878(13)	1.385(12)	1.158(13)
F2		1.834(14)	1.396(10)	1.163(14)

TABLE II: Fit results for the scalar and vector $B_s \to K$ form factors on each ensemble and for each simulated momentum.

We combine the use of Bayesian [14], marginalized [1], and chained fitting techniques. Our final fit results use marginalization with a total of $(N, \tilde{N}) = (8, 8)$ states accounted for, of which $(6, 1)$ are explicitly fit. We refer to such fits with the shorthand notation, $(6, 1)/(8, 8)$. States accounted for but not explicitly fit are marginalized in that their contributions are subtracted from the data prior to the fit. This technique reduces significantly the time required to perform the fits. In Fig. 2 we show the stability of the fits under variations in the numbers of states explicitly included and the total number of states accounted for in the fit. The $B_s \to K$ form factor results from the correlation function fits are tabulated in Table II and additional details are given in Appendix B.

The form factors obtained from these fits preserve correlations resulting from shared gauge field configurations and quark propagators used in data generation. The preservation of correlations is demonstrated in the top panel of Fig. 3 where, *e.g.*, significant correlations among the $B_s \to K$ form factor fit results are seen at common momenta and nonzero correlations among form factors for the two decays is suggested. The bottom panel of Fig. 3 shows the distribution over all ensembles of correlations among form factors for the two decays. Accounting for these correlations is useful in our determination of the ratio of form factors for the two decays. Fit results for $B_s \to \eta_s$, and the resulting form factor ratios, are presented in Appendix D.

V. CHIRAL, CONTINUUM, AND KINEMATIC EXTRAPOLATION

The results of HPChPT [4, 5] suggest a factorization, to at least one-loop order, of the soft physics of logarithmic chiral corrections and the physics associated with kinematics in the form factors describing semileptonic
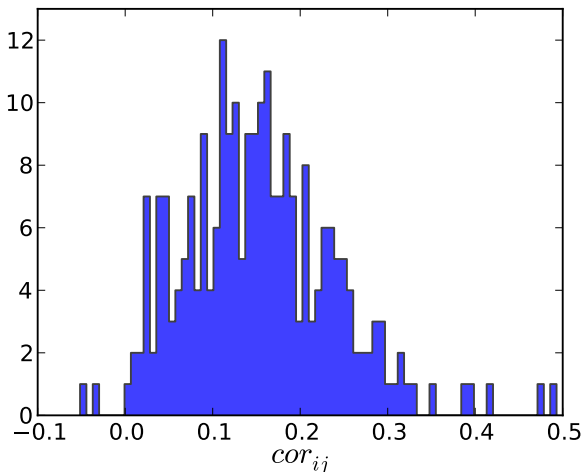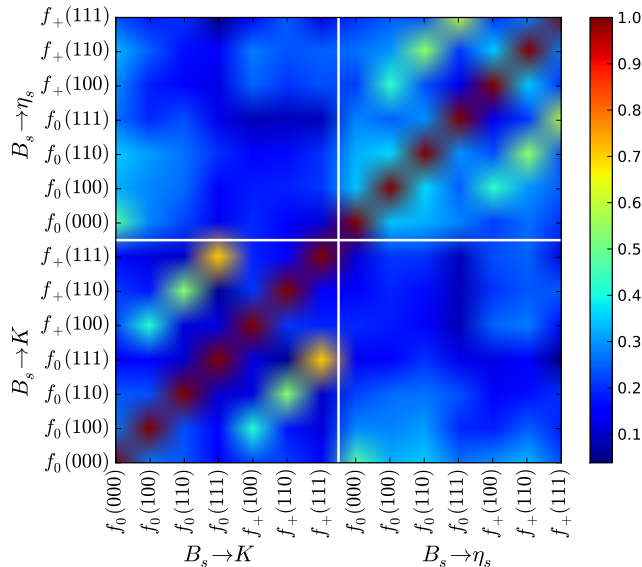
FIG. 3: (*color online*) (*top*) Heat map of the correlation matrix for ensemble C1. (*bottom*) Distribution of correlations among the form factors for $B_s \to K$ and $B_s \to \eta_s$ for all ensembles.

decays of heavy mesons,

$$f_{\parallel,\perp}(E) = (1 + [\text{logs}]) \, \mathcal{K}_{\parallel,\perp}(E). \quad (19)$$

The logarithmic chiral corrections, calculated in Ref. [5] for several $B_{(s)}$ decays, are independent of E. An unspecified function $\mathcal{K}$ characterizes the kinematics.

To obtain results over the full kinematic range one must include lattice simulation data over a range of energies. However, for any relevant physical scale Λ (*e.g.* Λ_{QCD}, $1/r_1$, Λ_{ChPT}, . . .), $E \gtrsim \Lambda$ at nominal lattice momenta and there is no convergent expansion of the unknown function $\mathcal{K}(E)$ in powers of E/Λ. This is an inherent limitation of characterizing the kinematics in

terms of energy. The energy of the daughter meson is a poor variable with which to describe the kinematics.

In contrast, the z-expansion [15–17] provides a convergent, model-independent characterization of the kinematics over the entire kinematically accessible range. Combining a z-expansion on each ensemble[2] with the HPChPT inspired factorization of Eq. (19) allows a simultaneous chiral, continuum, and kinematic extrapolation of lattice data at arbitrary energies. Because the chiral logs are the same for $f_{\parallel}$ and $f_{\perp}$, linear combinations (*i.e.* f_0 and f_+) factorize in the same way and have the same chiral logs. Motivated by these observations, we construct a HPChPT-motivated modified z-expansion, which we call the "HPChPT z-expansion", and fit the lattice data of Tables II and IX, with accompanying covariance matrix, to fit functions of the form

$$\begin{aligned} P_{0,+}(q^2) f_{0,+}(q^2) &= (1 + [\text{logs}]) \\ &\quad \times \sum_{k=0}^{K} a_k^{(0,+)} D_k^{(0,+)} z(q^2)^k, \quad (20) \end{aligned}$$

where [logs] are the continuum HPChPT logs of Ref. [5], and generic analytic chiral and discretization effects are accounted for by D_k. Resonances above q_{max}^2 but below the $B_s K$ production threshold, *i.e.* those in the range $q_{\text{max}}^2 < q^2 < (M_{B_s} + M_K)^2$, are accounted for via the Blaschke factor, $P = 1 - q^2/M_{\text{res}}^2$. Though not observed, we allow for the possibility of a $J^P = 0^+$ state in P_0, with choice of mass guided by Ref. [13]. Our fit results are insensitive to the presence of this state. The factorization suggested by HPChPT may not hold at higher order [18] so we allow chiral analytic terms, which help parameterize effects from omitted higher order chiral logs, to have energy dependence (*i.e.* to vary with k).

We note that Eq. (20) is the modified z-expansion introduced in Refs. [2, 3], with the coefficients of the chiral logarithmic corrections fixed by the results of HPChPT. In the chiral and continuum limits

$$\lim_{\substack{m \to m_{\text{physical}} \\ a \to 0}} (1 + [\text{logs}]) \, a_k D_k = b_k \text{ of Ref. } [17], \quad (21)$$

and Eq. (20) is equivalent to the BCL parametrization [17] of the form factors.

Following Ref. [17] we impose a constraint on $a_K^{(+)}$ from the expected scaling behavior of $f_+(q^2)$ in the neighborhood of q_{max}^2. The resulting fit function for f_+ is

$$\begin{aligned} P_+(q^2) f_+(q^2, a) &= (1 + [\text{logs}]) \\ &\times \sum_{k=0}^{K-1} a_k^{(+)} D_k^{(+)}(a) \big[z(q^2)^k - (-1)^{k-K} \frac{k}{K} z(q^2)^K \big]. \quad (22) \end{aligned}$$

_____

[2] This assumes the general arguments on which the z-expansion is based hold for heavier than physical quark masses and at finite lattice spacing.

We write $f(q^2, a)$, $z(q^2)$, and $D_k(a)$, explicitly exposing the dependence on q^2 and a. This is useful in explaining the implementation of a second kinematic constraint we impose on the form factors. At the kinematic endpoint $q^2 = 0$, the continuum extrapolated form factors f_0 and f_+ are equal, i.e. $f_0(0, 0) = f_+(0, 0)$. We impose this constraint by fixing the coefficient $a_0^{(0)}$,

$$a_0^{(0)} D_0^{(0)}(0) = -\sum_{k=1}^{K} a_k^{(0)} D_k^{(0)}(0) z(0)^k$$

$$+ \sum_{k=0}^{K-1} a_k^{(+)} D_k^{(+)}(0) \big[z(0)^k - (-1)^{k-K} \frac{k}{K} z(0)^K \big]. \quad (23)$$

Imposing this constraint results in the fit function for f_0

$$P_0(q^2) f_0(q^2, a) = (1 + [\text{logs}])$$

$$\times \Bigg\{ \sum_{k=1}^{K} a_k^{(0)} \Big[D_k^{(0)}(a) z(q^2)^k - \frac{D_0^{(0)}(a)}{D_0^{(0)}(0)} D_k^{(0)}(0) z(0)^k \Big]$$

$$+ \frac{D_0^{(0)}(a)}{D_0^{(0)}(0)} \sum_{k=0}^{K-1} a_k^{(+)} D_k^{(+)}(0) \big[z(0)^k - (-1)^{k-K} \frac{k}{K} z(0)^K \big] \Bigg\}.$$

$$(24)$$

In the fit functions for f_0 and f_+, Eqs. (22, 24), D_k and [logs] are given by,

$$D_k = 1 + c_1^{(k)} x_\pi + c_2^{(k)} \big(\tfrac{1}{2} \delta x_\pi + \delta x_K \big)$$

$$+ c_3^{(k)} \delta x_{\eta_s} + d_1^{(k)} (a/r_1)^2 + d_2^{(k)} (a/r_1)^4$$

$$+ e_1^{(k)} (aE_K)^2 + e_2^{(k)} (aE_K)^4, \quad (25)$$

$$[\text{logs}] = -\frac{3}{8} x_\pi (\log x_\pi + \delta_{FV}) - \frac{1 + 6g^2}{4} x_K \log x_K$$

$$- \frac{1 + 12g^2}{24} x_\eta \log x_\eta, \quad (26)$$

with implicit indices in Eq. (25) specifying the scalar or vector form factor. We account for momentum-independent and momentum-dependent discretization effects in D_k. The values of aE_K that enter the fit are the values from the simulation and are, of course, small. Finite volume effects in the simulation are included via a shift δ_{FV} in the pion log [19]. The infinite volume limit is taken by setting this shift to zero. Eq. (26) gives the HPChPT [5] result for the chiral logarithmic correction to $B_s \to K$ form factors. These expressions make use of the dimensionless quantities

$$x_{\pi, K, \eta} = \frac{M_{\pi, K, \eta}^2}{(4\pi f_\pi)^2}, \quad (27)$$

$$\delta x_{\pi, K} = \frac{(M_{\pi, K}^{\text{asqtad}})^2 - (M_{\pi, K}^{\text{HISQ}})^2}{(4\pi f_\pi)^2}, \quad (28)$$

$$\delta x_{\eta_s} = \frac{(M_{\eta_s}^{\text{HISQ}})^2 - (M_{\eta_s}^{\text{physical}})^2}{(4\pi f_\pi)^2}, \quad (29)$$

where $M_\eta^2 = (M_\pi^2 + 2M_{\eta_s}^2)/3$. We determine q^2 and z on each ensemble using correlator fit results for meson masses and simulation momenta. Light and heavy quark discretization effects are accommodated for by making the $d_i^{(k)}$ mild functions of the masses, accomplished by the replacements

$$d_1^{(k)} \rightarrow d_1^{(k)}(1 + l_1^{(k)} x_\pi + l_2^{(k)} x_\pi^2)(1 + h_1^{(k)} \delta x_b + h_2^{(k)} \delta x_b^2),$$

$$d_2^{(k)} \rightarrow d_2^{(k)}(1 + l_3^{(k)} x_\pi + l_4^{(k)} x_\pi^2)(1 + h_3^{(k)} \delta x_b + h_4^{(k)} \delta x_b^2),$$

$$(30)$$

where $\delta x_b = am_b - 2.26$ is chosen so that as am_b varies over the coarse and fine ensembles $-0.4 \lesssim \delta x_b \lesssim 0.4$.

Lastly, we account for uncertainty associated with the perturbative matching of Sec. III. With the matching coefficients calculated in Ref. [11], we find $\mathcal{O}(\alpha_s, \Lambda_{\text{QCD}}/m_b, \alpha_s/(am_b))$ contributions to be $\sim 4\%$ of the total contribution to $\langle V_0 \rangle$. Of this 4% the majority, $\sim 3.5\%$, comes from the one loop $\mathcal{O}(\alpha_s)$ correction and $< 1\%$ from the NRQCD matching via $\langle J_0^{(1), \text{sub}} \rangle$. For $\langle V_k \rangle$ we find contributions at this order to be $\sim 2\%$, with $\sim 1\%$ coming from the $\mathcal{O}(\alpha_s)$ correction and $< 1\%$ from the NRQCD matching. The matching error results from omitted higher order corrections, the size of which we estimate from observed leading order effects, where we conservatively use the larger 4%. Following the arguments outlined in Ref. [10] we estimate the matching error to be the same size as the observed $\mathcal{O}(\alpha_s, \Lambda_{\text{QCD}}/m_b, \alpha_s/(am_b))$ contributions and take the matching error to be 4%. This is equivalent to taking the $\mathcal{O}(\alpha_s^2)$ matching coefficient to be four times larger than the $\mathcal{O}(\alpha_s)$ matching coefficient $\rho_0^{(V_0)}$ (13 times larger than $\rho_0^{(V_k)}$). This uncertainty is associated with the hadronic matrix elements and therefore, by Eqs. (3, 4), with $f_\parallel$ and $f_\perp$. To correctly incorporate it in the results for f_0 and f_+ we convert our fit functions for $f_{0,+}$ into $f_{\parallel, \perp}$, multiply by $(1 + m_{\parallel, \perp})$, where $m_{\parallel, \perp}$ is a coefficient representing the matching error with a prior central value of zero and width 0.04, then convert back to $f_{0,+}$ before performing the fit. Schematically, we modify the fit functions, defined in Eqs. (22, 24), by

$$f_0, f_+ \rightarrow f_\parallel, f_\perp \quad (31)$$

$$f_\parallel, f_\perp \rightarrow (1 + m_\parallel) f_\parallel, (1 + m_\perp) f_\perp \quad (32)$$

$$(1 + m_\parallel) f_\parallel, (1 + m_\perp) f_\perp \rightarrow f_0^{\text{corrected}}, f_+^{\text{corrected}}, \quad (33)$$

then we use $f_{0,+}^{\text{corrected}}$ to fit the results of the correlation function fits of Sec. IV. Conversions between the form factors $f_{0,+}$ and $f_{\parallel, \perp}$ are performed using Eqs. (5, 6).

The results of a simultaneous fit to the data for $f_{0,+}^{B_s K}$ and $f_{0,+}^{B_s \eta_s}$, in which the maximum order of z (specified by K in Eqs. (22, 24)) is 3 and $\chi^2/\text{dof} = 84.0/70$, are shown relative to the data in Fig. 4 for $B_s \to K$. Details of prior choices and fit results are given in Appendix C.

We test the stability of this fit to the following modifications of the fit Ansätze:

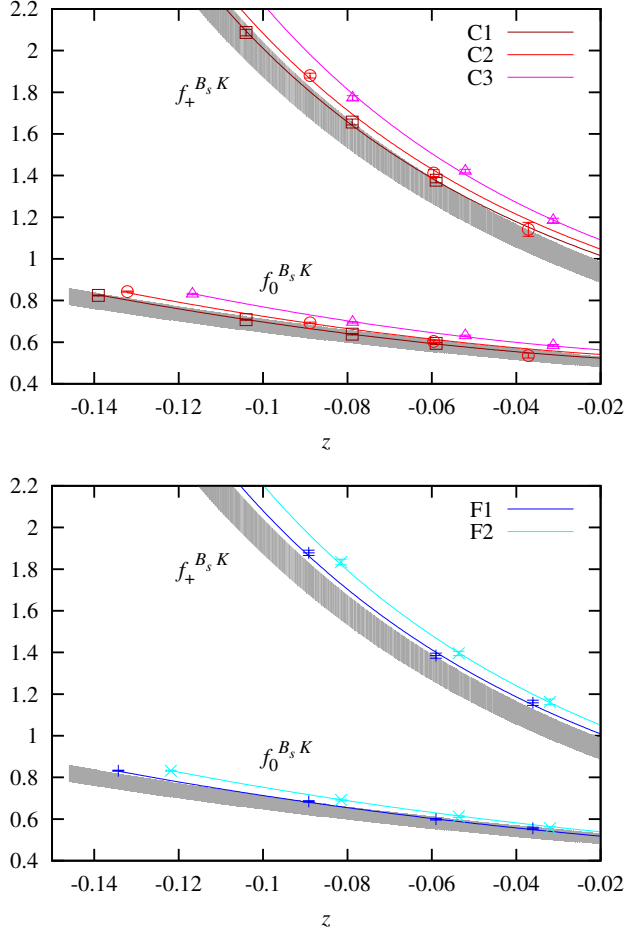1. Truncate the z expansion at $\mathcal{O}(z^2)$.

FIG. 4: (*color online*) $B_s \to K$ form factor results from a simultaneous chiral, continuum, and kinematic extrapolation via the HPChPT z-expansion are shown (*top*) relative to coarse ensemble data (C1, C2, and C3) and (*bottom*) relative to fine ensemble data (F1 and F2).

2. Truncate the z expansion at $\mathcal{O}(z^4)$.

3. Truncate the z expansion at $\mathcal{O}(z^5)$.

4. Drop $\mathcal{O}(aE_K)^4$ momentum-dependent and $\mathcal{O}(a^4)$ momentum-independent discretization terms in Eq. (25).

5. Drop the am_b-dependent discretization terms in Eq. (30).

6. Drop the light-quark mass-dependent discretization terms in Eq. (30).

7. Add the following NNLO chiral analytic terms to

D_k as defined in Eq. (25):

$$
\begin{aligned}
& c_4^{(k)} x_\pi^2 + c_5^{(k)} \left(\tfrac{1}{2}\delta x_\pi + \delta x_K\right)^2 + c_6^{(k)} \delta x_{\eta_s}^2 \\
& + c_7^{(k)} x_\pi \left(\tfrac{1}{2}\delta x_\pi + \delta x_K\right) + c_8^{(k)} x_\pi \delta x_{\eta_s} \quad\quad (34) \\
& + c_9^{(k)} \left(\tfrac{1}{2}\delta x_\pi + \delta x_K\right)\delta x_{\eta_s} + c_{10}^{(k)} x_\pi (a/r_1)^2 \\
& + c_{11}^{(k)} \left(\tfrac{1}{2}\delta x_\pi + \delta x_K\right)(a/r_1)^2 + c_{12}^{(k)} \delta x_{\eta_s}(a/r_1)^2 .
\end{aligned}
$$

8. Drop the sea- and valence-quark mass difference term $\left(\tfrac{1}{2}\delta x_\pi + \delta x_K\right)$ from Eq. (25).

9. Drop the strange quark mistuning term δx_{η_s} from Eq. (25).

10. Drop finite volume effects, *i.e.* set $\delta_{FV} = 0$ in Eq. (26).

The stability of the $B_s \to K$ fit results to these modifications is shown in Fig. 5, where results are shown at the extrapolated $q^2 = 0$ point. This point is furthest from the data region where simulations are performed and therefore is particularly sensitive to changes in the fit function. In Fig. 5 our final fit result, as defined by Eqs. (22, 24) with $K = 3$ and by Eqs. (25–30), is indicated by the dashed line and gray band.

Modifications 1, 2, and 3 vary the order of the truncation in z and demonstrate that by $\mathcal{O}(z^3)$ fit results have stabilized and errors have saturated. We therefore conclude that the error of the $\mathcal{O}(z^3)$ fit adequately accounts for the systematic error due to truncating the z expansion.

Momentum-dependent and momentum-independent discretization effects proportional to a^4 are removed in modification 4. This results in a modest increase in χ^2 and a negligible shift in the fit result. This suggests our final fit, which includes the a^4 effects, adequately accounts for all discretization effects observed in the data.

In modifications 5 and 6 we remove heavy- and light-quark mass-dependent discretization effects with essentially no impact on the fit. That our results are independent of light-quark mass dependent discretization effects suggests that staggered taste violating effects are accommodated for by a generic a^2 dependence.

Modification 7 tests the truncation of chiral analytic terms after NLO by adding the NNLO terms listed in Eq. (34). This results in a slight decrease in χ^2 but has no noticeable effect on the fit central value or error. From this we conclude that errors associated with omitted higher order chiral terms are negligible.

Differences in sea and valence quark masses, due in part to our use of HISQ valence- and asqtad sea-quarks, are neglected in modification 8. This results in a small increase in χ^2 and negligible change in the fit results. We account for these small mass differences in our final fit, though this test suggests they are unimportant in the fit.

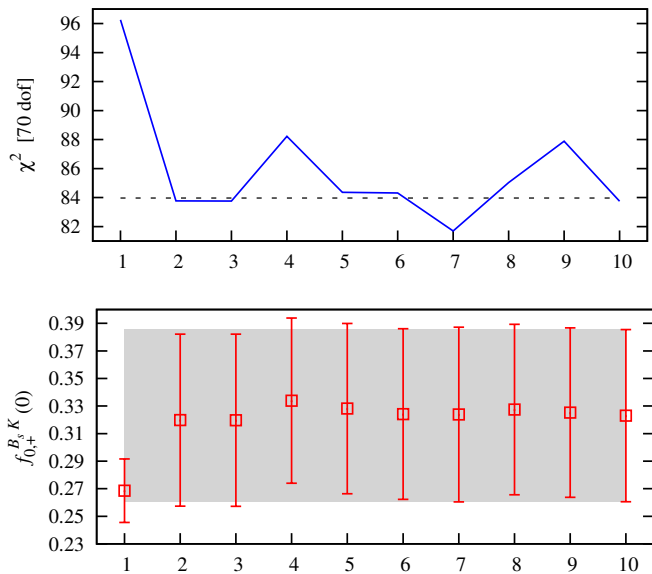Effects due to strange quark mass mistuning on the ensembles are omitted in modification 9, resulting in a

FIG. 5: The stability of the HPChPT z-expansion is demonstrated by studying the fit results under various modifications, discussed in the text. The top panel shows χ^2 with 70 degrees of freedom (dof) for each test fit and the bottom panel shows form factors extrapolated to $q^2 = 0$.
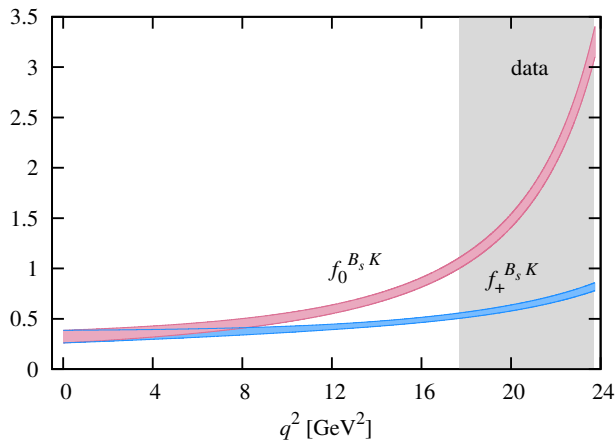

FIG. 6: $B_s \to K$ form factor results from a simultaneous chiral, continuum, and kinematic extrapolation via the HPChPT z-expansion. The q^2 region for which lattice simulation data exist is indicated by the shaded region.

modest increase in χ^2 and no change in the fit central value and error. We include these effects in our final fit.

Modification 10 results in nearly identical fit results, suggesting that finite volume effects are negligible in our data. We include these effects in our final fit results.

VI. FORM FACTOR RESULTS

In this section we present final results, with a complete error budget, for the $B_s \to K$ form factors. We provide the needed information to reconstruct the form factors and compare our results with previous model calculations.

Fig. 6 shows the results of the chiral, continuum, and kinematic extrapolation of Sec. V, plotted over the entire kinematic range of q^2. The form factors, extrapolated to $q^2 = 0$, have the value $f_{0,+}^{B_s K}(0) = 0.323(63)$.

A. Fit errors for the HPChPT z-expansion

The inputs in our chiral, continuum, and kinematic extrapolation fits are data (the correlator fit results for f_0 and f_+ in Tables II and IX with the accompanying covariance matrix) and priors. The total hessian error of the fit can be described in terms of contributions from these inputs, as described in detail in Appendix A. We group priors in a meaningful, though not unique, way and discuss the error associated with the chiral, continuum, and kinematic extrapolation based on these groupings. As the priors are, by construction, uncorrelated with one another, we can group them together in any way we find meaningful. The resulting error groupings are uncorrelated and add in quadrature to the total error. In Fig. 7 we plot the following relative error components as functions of q^2:

- **experiment:** This is the error in the fit due to uncertainty of experimentally determined, and other, input parameters. It is the sum in quadrature of the errors due to priors for the "Group I" fit parameters listed in Table VII. This error is independent of q^2 and subdominant.

- **kinematic:** This error component is due to the priors for the coefficients $a_k^{(0,+)}$ in Eqs. (22, 24). A comparison of the fit results from modifications 1, 2, and 3 in Fig. 5 shows that by $\mathcal{O}(z^3)$ the fit results have stabilized and errors have saturated. The kinematic error therefore includes the error associated with truncating the z-expansion. The extrapolation to values of q^2 for which we have no simulation data is controlled by the z-expansion. As a result, the growth in form factor errors away from the simulation region is due almost entirely to kinematic and statistical errors.

- **chiral:** This error component is the sum in quadrature of errors associated with priors for $c_i^{(k)}$ in Eq. (25). These terms are responsible for extrapolating to the physical light quark mass and for accommodating for the slight strange quark mistuning and the small mismatch in sea and valence quark masses due to the mixed action used in the

simulation. As shown in Fig. 7, these errors are subdominant and do not vary significantly with q^2.

- **discretization:** We account for momentum-dependent discretization effects via the $e_i^{(k)}$, and momentum-independent discretization effects via the $d_i^{(k)}$, terms of Eq. (25). In addition we allow for heavy- and light-quark mass-dependent discretization effects via the $h_i^{(k)}$ and $l_i^{(k)}$ terms in Eq. (30). The discretization error component, which is essentially independent of q^2, is the sum in quadrature of the error due to the priors for these fit parameters.

- **statistical:** The statistical component of the error is due to uncertainty in the data, *i.e.* the errors from form factor fit results of Table II. Simulation data exist for $q^2 \gtrsim 17$ GeV^2 for f_0 and over the range 17 GeV$^2 \lesssim q^2 \lesssim 22$ GeV^2 for f_+. Extrapolation beyond these regions leads to increasing errors.

- **matching:** The matching error is due to the uncertainty associated with the priors for $m_{\parallel,\perp}$ introduced in Eq. (32) and discussed in the surrounding text.

In addition to the largest sources of error, which we account for directly in the fit, there are remaining systematic uncertainties.

We simulate with degenerate light quarks and neglect electromagnetism. By adjusting the physical kaon mass ($M_{K^\pm} \to M_{K^0}$) used in the chiral, continuum and kinematic extrapolation, we estimate the "kinematic" effects of omitting electromagnetic and isospin symmetry breaking in our simulation to be $\lesssim 0.1\%$. It is more difficult to determine the size of the full effects. However, in general electromagnetic and isospin effects are expected to be sub-percent. We assume the error in our form factor calculation due to these effects is negligible relative to other sources of uncertainty.

Our simulations include up, down, and strange sea quarks and we assume omitted charm sea quark effects are negligible. This has been the case for processes in which it has been possible and appropriate to perturbatively estimate effects of charm quarks in the sea [20].

Our final form factor results, multiplied by the Blaschke factor $P_{0,+}$, are shown in Fig. 8 where they are compared with results from a model calculation using perturbative QCD (pQCD) [21] and a relativistic quark model (RQM) [22]. Our results provide significant clarification on the form factors at large q^2.

B. Reconstructing $B_s \to K\ell\nu$ Form Factors

In the physical limit our form factor results are parameterized in a BCL [17] form with coefficients $b_k^{(0,+)}$ (see Eq. (21)). Including the kinematic constraint and terms
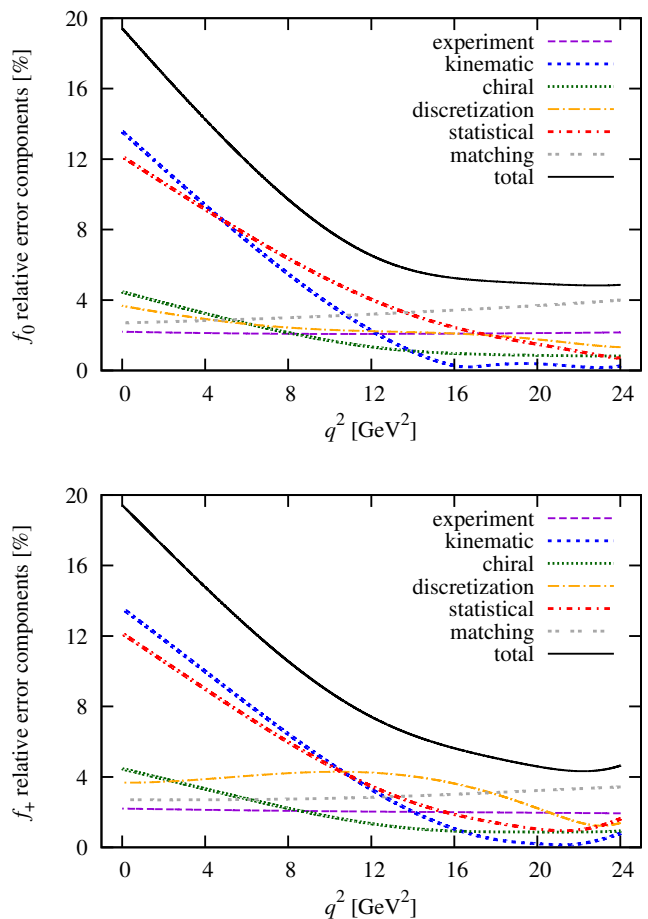

FIG. 7: (*color online*) $B_s \to K$ (*top*) f_0 and (*bottom*) f_+ relative error components. The total error (solid line) is the sum in quadrature of the components.

through order z^3, we have

$$P_0(q^2)f_0(q^2) = \sum_{k=1}^{3} b_k^{(0)}(z^k - z(0)^k)$$
$$+ \sum_{k=0}^{2} b_k^{(+)}\left[z(0)^k - (-1)^{k-3}\frac{k}{3}z(0)^3\right], \quad (35)$$

$$P_+(q^2)f_+(q^2) = \sum_{k=0}^{2} b_k^{(+)}\left[z^k - (-1)^{k-3}\frac{k}{3}z^3\right], \quad (36)$$

where

$$z(q^2) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}, \quad (37)$$

$$t_+ = (M_{B_s} + M_K)^2, \quad (38)$$

$$t_0 = (M_{B_s} + M_K)(\sqrt{M_{B_s}} - \sqrt{M_K})^2, \quad (39)$$

$$P_{0,+}(q^2) = 1 - q^2/M_{0,+}^2, \quad (40)$$

and the resonance masses are $M_0 = 5.6794(10)$ GeV and $M_+ = 5.32520(48)$ GeV. The values of the coefficients

coefficient	value
$b_1^{(0)}$	0.315(129)
$b_2^{(0)}$	0.945(1.305)
$b_3^{(0)}$	2.391(4.671)
$b_0^{(+)}$	0.3680(214)
$b_1^{(+)}$	-0.750(193)
$b_2^{(+)}$	2.720(1.458)

	$b_1^{(0)}$	$b_2^{(0)}$	$b_3^{(0)}$	$b_0^{(+)}$	$b_1^{(+)}$	$b_2^{(+)}$
$b_1^{(0)}$	1.676e-2	1.462e-1	4.453e-1	1.165e-3	2.140e-2	1.434e-1
$b_2^{(0)}$		1.702	5.852	9.481e-3	2.255e-1	1.539
$b_3^{(0)}$			2.181e+1	2.963e-2	7.472e-1	5.325
$b_0^{(+)}$				4.577e-4	1.157e-3	-1.309e-3
$b_1^{(+)}$					3.721e-2	1.858e-1
$b_2^{(+)}$						2.124

TABLE III: (*left*) Physical extrapolated coefficients of the HPChPT z-expansion for the $B_s \to K$ form factors, defined in Eqs. (35, 36) and (*right*) the associated covariance matrix.
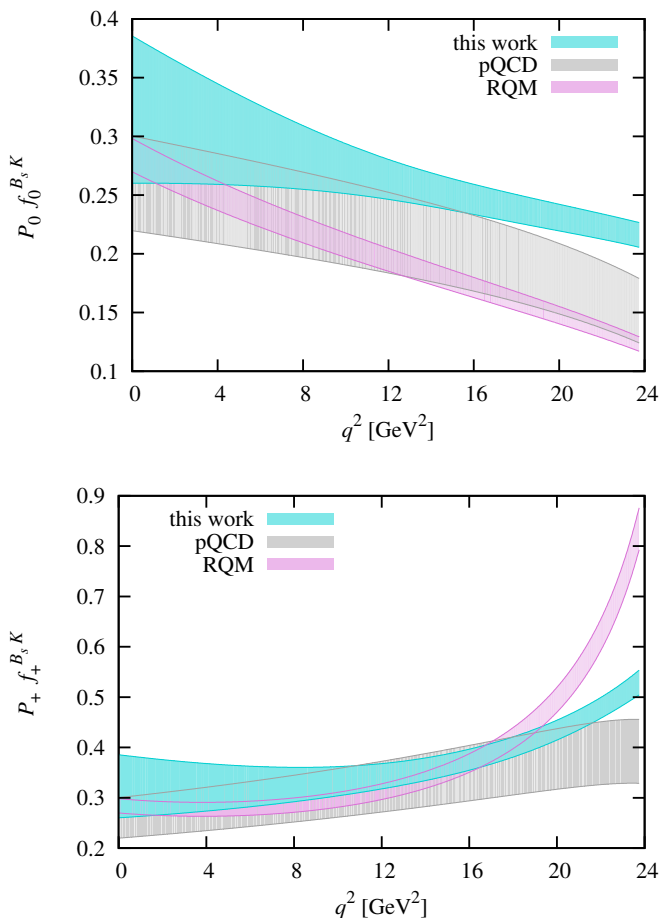

FIG. 8: (*color online*) Comparison of our $B_s \to K$ (*top*) f_0 and (*bottom*) f_+ form factors with those from a perturbative QCD model (pQCD) [21] and the Relativistic Quark Model (RQM) [22]. To ease comparison the vertical scale is reduced by multiplying the form factors by the Blaschke factor $P_{0,+}$.

$b_k^{(0,+)}$, derived from the extrapolation fit results of Sec. V, and the associated covariance matrix, are given in Table III. Note that it is necessary to take into account the correlations among the coefficients to correctly reproduce the form factor errors.

VII. PHENOMENOLOGY

With the benefit of *ab initio* form factors from lattice QCD, we explore the standard model implications of our results. In this section we make standard model predictions for several observables related to the $B_s \to K\ell\nu$ decay for $\ell = \mu$ and τ.

The standard model $B_s \to K\ell\nu$ differential decay rate is related to the form factors by

$$\frac{d\Gamma}{dq^2} = \frac{G_F^2 |V_{ub}|^2}{24\pi^3 M_{B_s}^2}\left(1-\frac{m_\ell^2}{q^2}\right)^2 |\mathbf{p}_K|\left[\left(1+\frac{m_\ell^2}{2q^2}\right)M_{B_s}^2\mathbf{p}_K^2|f_+|^2\right.$$
$$\left. + \frac{3m_\ell^2}{8q^2}(M_{B_s}^2 - M_K^2)^2|f_0|^2\right]. \quad (41)$$

In Fig. 9 we plot predicted differential decay rates for $B_s \to K\mu\nu$ and $B_s \to K\tau\nu$, divided by $|V_{ub}|^2$, over the full kinematic range of q^2. The ratio $\Gamma/|V_{ub}|^2$ can be combined with experimental results for the decay rates, typically differential decay rates integrated over q^2 bins, to allow the determination of $|V_{ub}|$. In Eqs. (42, 43) we give numerical results for $d\Gamma/dq^2$, integrated over the kinematically accessible regions of q^2,

$$\Gamma(B_s \to K\mu\nu)/|V_{ub}|^2 = 7.75(1.52)\text{ ps}^{-1}, \quad (42)$$
$$\Gamma(B_s \to K\tau\nu)/|V_{ub}|^2 = 4.92(0.60)\text{ ps}^{-1}. \quad (43)$$

Combining our form factor results with the current[3] inclusive and exclusive semileptonic determinations of $|V_{ub}|$,

$$\text{exclusive } |V_{ub}| = 3.47(22) \times 10^{-3}, \quad (44)$$
$$\text{inclusive } |V_{ub}| = 4.41(22) \times 10^{-3}, \quad (45)$$

---

[3] For inclusive $|V_{ub}|$ we take the value from the Particle Data Group [25]. For the exclusive determination we use the "global lattice + Belle" results reported by the FLAG-2 collaboration [26].
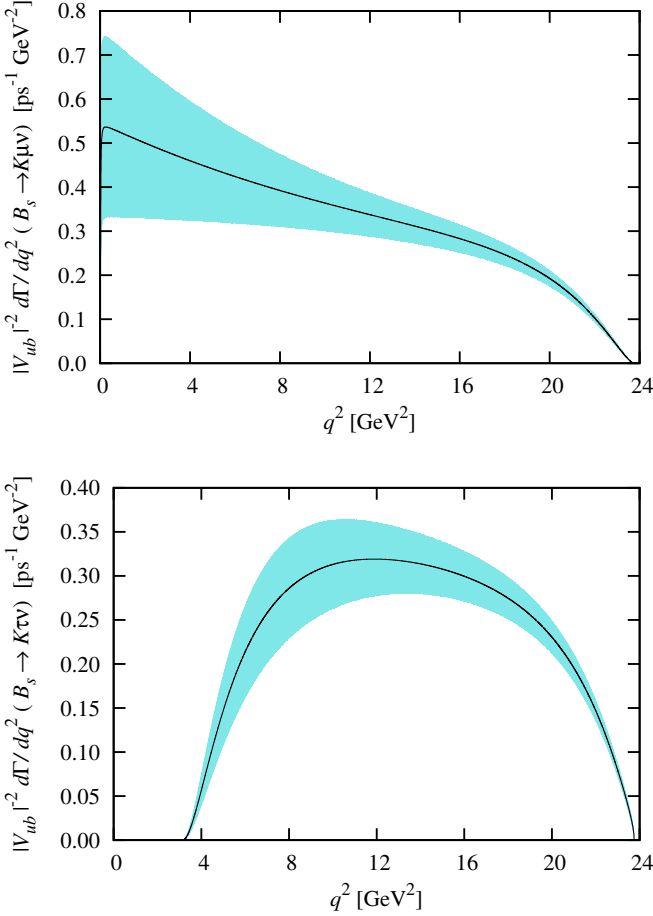
FIG. 9: Predicted differential decay rates, divided by $|V_{ub}|^2$, for (top) $B_s \to K\mu\nu$ and (bottom) $B_s \to K\tau\nu$.

FIG. 10: (*color online*) Predicted differential branching fractions for the (top) $B_s \to K\mu\nu$ and (bottom) $B_s \to K\tau\nu$ decays using inclusive and exclusive semileptonic determinations of $|V_{ub}|$. In each band, the light outer band includes all sources of error and the dark interior band neglects the uncertainty in $|V_{ub}|$.

we demonstrate in Fig. 10 the potential of this decay to shed light on this $\sim 3\sigma$ discrepancy. In this and subsequent figures, dark interior bands represent the error in the differential branching fractions omitting the error associated with $|V_{ub}|$. Experimental errors commensurate with these predictions, especially for the $B_s \to K\tau\nu$ decay or at large q^2 for the $B_s \to K\mu\nu$ decay, would allow differentiation between the current inclusive and exclusive values of $|V_{ub}|$.

Decays that couple to the τ have increased dependence on the scalar form factor and to new physics models with scalar states (see, *e.g.* Refs. [23, 24] for a discussion of new physics in the closely related decay $B \to \pi\tau\nu$). The ratio of the $B_s \to K\tau\nu$ differential branching fraction to that for $B_s \to K\mu\nu$,

$$R_\mu^\tau(q_{\text{low}}^2, q_{\text{high}}^2) = \frac{\int_{q_{\text{low}}^2}^{q_{\text{high}}^2} dq^2 \, d\mathcal{B}/dq^2(B_s \to K\tau\nu)}{\int_{q_{\text{low}}^2}^{q_{\text{high}}^2} dq^2 \, d\mathcal{B}/dq^2(B_s \to K\mu\nu)}, \quad (46)$$

is therefore a potentially sensitive probe of new physics. Integrating over the full kinematic range, we find

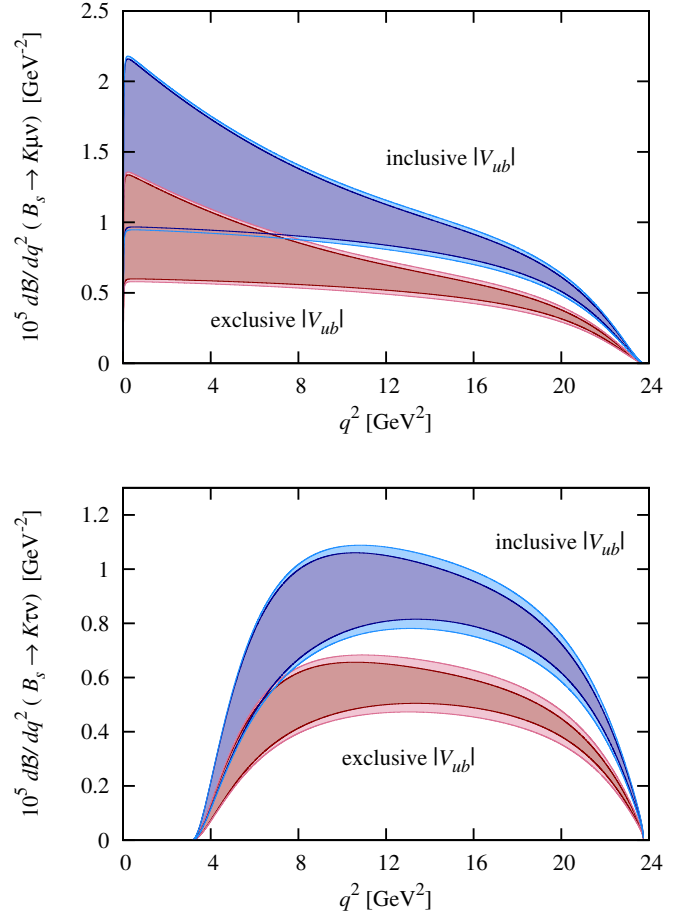$$R_\mu^\tau(m_\tau^2, q_{\text{max}}^2) = 0.800(44), \quad (47)$$

where $q_{\text{max}}^2 = (M_{B_s} - M_K)^2$. We plot the standard model prediction for this ratio, as a function of $q^2 = (q_{\text{low}}^2 + q_{\text{high}}^2)/2$, over the full kinematic range in Fig. 11.

The angular dependence of the differential decay rate, neglecting final state electromagnetic interactions, is given by

$$\frac{d^2\Gamma}{dq^2 \, d\cos\theta_\ell} = \frac{G_F^2 |V_{ub}|^2}{128\pi^3 M_{B_s}^2} \left(1 - \frac{m_\ell^2}{q^2}\right)^2 |\mathbf{p}_K| \times$$

$$\left[4M_{B_s}^2 |\mathbf{p}_K|^2 \left(\sin^2\theta_\ell + \frac{m_\ell^2}{q^2}\cos^2\theta_\ell \right) f_+^2 \right.$$

$$+ \frac{4m_\ell^2}{q^2}(M_{B_s}^2 - M_K^2)M_{B_s}|\mathbf{p}_K|\cos\theta_\ell f_0 f_+$$

$$\left. + \frac{m_\ell^2}{q^2}(M_{B_s}^2 - M_K^2)^2 f_0^2 \right], \quad (48)$$

where θ_ℓ is defined, in the q^2 rest frame (*i.e.* where $\mathbf{p}_\ell + \mathbf{p}_\nu$ is zero), as the angle between the final state lepton
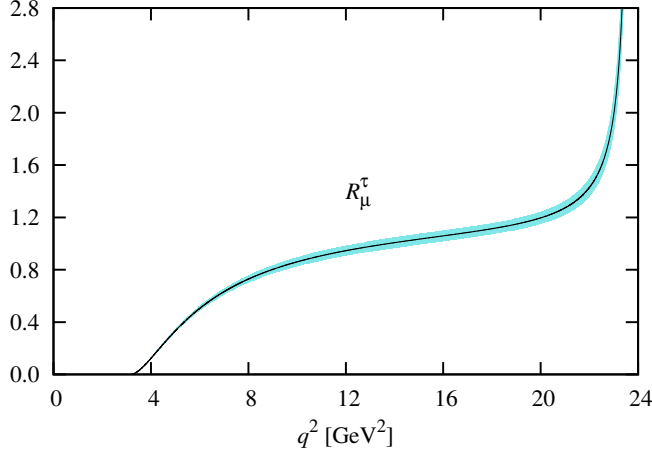
FIG. 11: Predicted differential branching fraction ratio.

and the B_s meson. From this angular dependence we can extract a forward-backward asymmetry [27],

$$\mathcal{A}_{\text{FB}}^{\ell}(q^2) = \left[\int_0^1 - \int_{-1}^0 \right] d\cos\theta_\ell \frac{d^2\Gamma}{dq^2\, d\cos\theta_\ell} \quad (49)$$

$$= \frac{G_F^2 |V_{ub}|^2}{32\pi^3 M_{B_s}} \left(1 - \frac{m_\ell^2}{q^2} \right)^2 |\mathbf{p}_K|^2 \times$$

$$\frac{m_\ell^2}{q^2} (M_{B_s}^2 - M_K^2) f_0 f_+, \quad (50)$$

which is suppressed in the standard model by a factor of m_ℓ^2/q^2. In Fig. 12 we show standard model predictions for the forward-backward asymmetry using the inclusive and exclusive values for $|V_{ub}|$. Integrating over the full kinematic range of q^2 gives

$$\int_{m_\mu^2}^{q_{\text{max}}^2} dq^2\, \mathcal{A}_{\text{FB}}^\mu(q^2)/|V_{ub}|^2 = 0.052(17)\ \text{ps}^{-1}, \quad (51)$$

$$\int_{m_\tau^2}^{q_{\text{max}}^2} dq^2\, \mathcal{A}_{\text{FB}}^\tau(q^2)/|V_{ub}|^2 = 1.40(20)\ \text{ps}^{-1}. \quad (52)$$

Normalizing the forward-backward asymmetry by the differential decay rate removes $|V_{ub}|$ ambiguity and most hadronic uncertainties,

$$\bar{\mathcal{A}}_{\text{FB}}^\ell(q_{\text{low}}^2, q_{\text{high}}^2) = \frac{\int_{q_{\text{low}}^2}^{q_{\text{high}}^2} dq^2\, \mathcal{A}_{\text{FB}}^\ell(q^2)}{\int_{q_{\text{low}}^2}^{q_{\text{high}}^2} dq^2\, d\Gamma/dq^2}, \quad (53)$$

and represents the probability the lepton will have a momentum component, in this frame, in the direction of motion of the parent B_s meson. Integrating over q^2 yields

$$\bar{\mathcal{A}}_{\text{FB}}^\mu(m_\mu^2, q_{\text{max}}^2) = 0.0066(10), \quad (54)$$

$$\bar{\mathcal{A}}_{\text{FB}}^\tau(m_\tau^2, q_{\text{max}}^2) = 0.284(17), \quad (55)$$
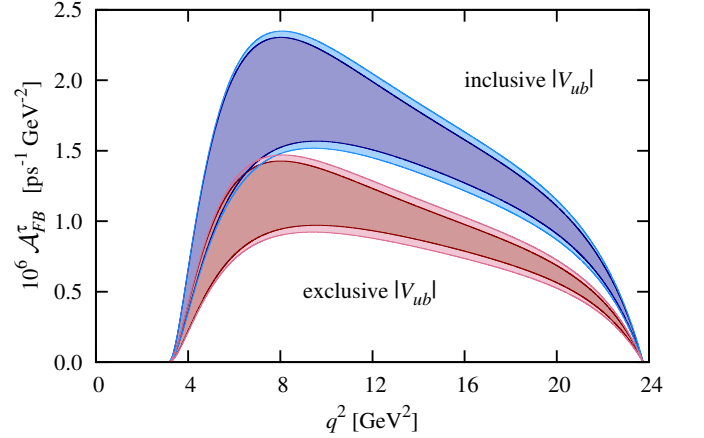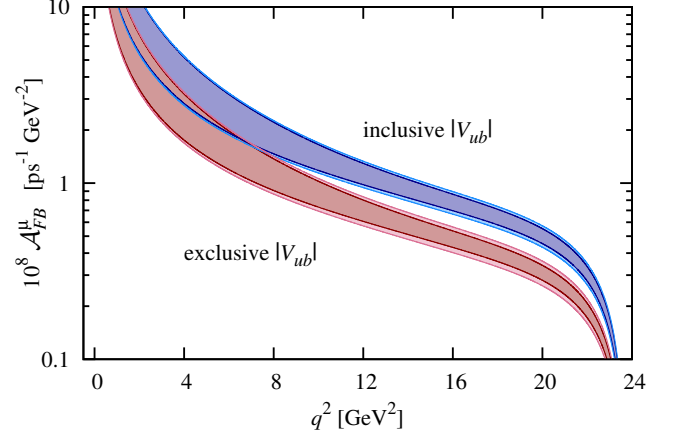

FIG. 12: (*color online*) Differential decay rate forward-backward asymmetries for the (*top*) $B_s \to K\mu\nu$ and (*bottom*) $B_s \to K\tau\nu$ decays using inclusive and exclusive semileptonic determinations of $|V_{ub}|$. Light outer bands includes all sources of error and the dark interior bands neglect uncertainty in $|V_{ub}|$.

with central values equal to those obtained by taking the ratio of results from Eqs. (51, 52) with those from Eqs. (42, 43). The errors, however, are $\sim 3\times$ smaller when correlations are accounted for. The normalized standard model asymmetries are plotted in Fig. 13 as a function of q^2.

The production of right-handed final state leptons is helicity-suppressed in the standard model, providing a probe of non standard model helicity-violating interactions. The standard model differential decay rates for left-handed (LH) and right handed (RH) polarized final
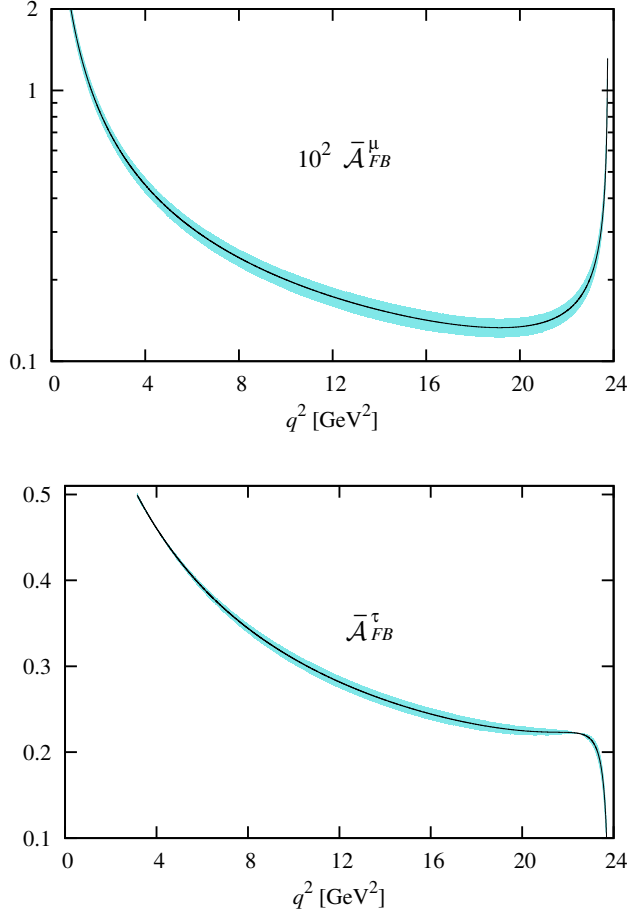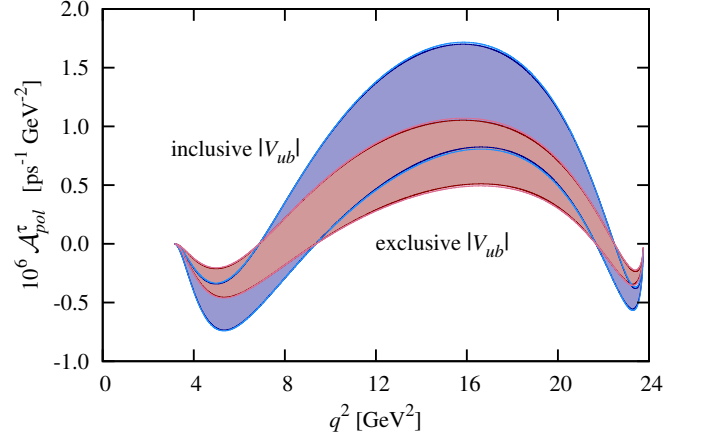
FIG. 14: Standard model τ-polarization distribution for the differential decay rate of $B_s \to K\tau\nu$.

q^2 gives

$$\int_{m_\mu^2}^{q_{\max}^2} dq^2 \; \mathcal{A}_{\mathrm{pol}}^\mu(q^2)/|V_{ub}|^2 \;=\; 7.61(1.60) \text{ ps}^{-1}, \quad (58)$$

$$\int_{m_\tau^2}^{q_{\max}^2} dq^2 \; \mathcal{A}_{\mathrm{pol}}^\tau(q^2)/|V_{ub}|^2 \;=\; 0.52(32) \text{ ps}^{-1}. \quad (59)$$

As with the forward-backward asymmetry, we normalize the ℓ-polarization distribution by the differential decay rate to remove ambiguity associated with $|V_{ub}|$ and hadronic uncertainties. The resulting polarization fraction [27] is defined by

$$\bar{\mathcal{A}}_{\mathrm{pol}}^\ell(q_{\mathrm{low}}^2, q_{\mathrm{high}}^2) = \frac{\int_{q_{\mathrm{low}}^2}^{q_{\mathrm{high}}^2} dq^2 \; \mathcal{A}_{\mathrm{pol}}^\ell(q^2)}{\int_{q_{\mathrm{low}}^2}^{q_{\mathrm{high}}^2} dq^2 \; d\Gamma/dq^2}. \quad (60)$$

Integrating over q^2 we find the standard model prediction for the fraction of polarized leptons to be

$$\bar{\mathcal{A}}_{\mathrm{pol}}^\mu(m_\mu^2, q_{\max}^2) \;=\; 0.982(^{+18}_{-79}), \quad (61)$$

$$\bar{\mathcal{A}}_{\mathrm{pol}}^\tau(m_\tau^2, q_{\max}^2) \;=\; 0.105(63), \quad (62)$$

where the error associated with the numerical integration of $\bar{\mathcal{A}}_{\mathrm{pol}}^\mu$ (± 0.079) has been truncated to satisfy the constraint that $\bar{\mathcal{A}}_{\mathrm{pol}}^\ell < 1$. The q^2 dependence of the ℓ-polarization fraction is plotted in Fig. 15.

VIII.   SUMMARY AND OUTLOOK

Using NRQCD b and HISQ light and strange valence quarks with the MILC $2+1$ dynamical asqtad configurations, we report on the first lattice QCD calculation of the form factors for the semileptonic decay $B_s \to K\ell\nu$.

With the help of a new technique, called chaining, we fit the $B_s \to K$ correlator data simultaneously with data

FIG. 13: Normalized differential decay rate forward-backward asymmetries for the (*top*) $B_s \to K\mu\nu$ and (*bottom*) $B_s \to K\tau\nu$ decays.

state leptons in $B_s \to K\ell\nu$ decays is [27]

$$\frac{d\Gamma(LH)}{dq^2} = \frac{G_F^2 |V_{ub}|^2 |\mathbf{p}_K|^3}{24\pi^3} \left(1 - \frac{m_\ell^2}{q^2}\right)^2 f_+^2 \;,$$

$$\frac{d\Gamma(RH)}{dq^2} = \frac{G_F^2 |V_{ub}|^2 |\mathbf{p}_K|}{24\pi^3} \frac{m_\ell^2}{q^2} \left(1 - \frac{m_\ell^2}{q^2}\right)^2 \times$$
$$\left[\frac{3}{8} \frac{(M_{B_s}^2 - M_K^2)^2}{M_{B_s}^2} f_0^2 + \frac{1}{2}|\mathbf{p}_K|^2 f_+^2 \right], \quad (56)$$

and the ℓ-polarization distribution is given by the difference

$$\mathcal{A}_{\mathrm{pol}}^\ell(q^2) = \frac{d\Gamma(LH)}{dq^2} - \frac{d\Gamma(RH)}{dq^2}. \quad (57)$$

We plot the τ-polarization distribution, again using the inclusive and exclusive values of $|V_{ub}|$ from Eqs. (44, 45), in Fig. 14. Because of their relatively small mass, muons produced in the decay are predominantly left-handed and the plot of $\mathcal{A}_{\mathrm{pol}}^\mu$ is equivalent to the total differential decay rate. Integrating the ℓ-polarization distributions over
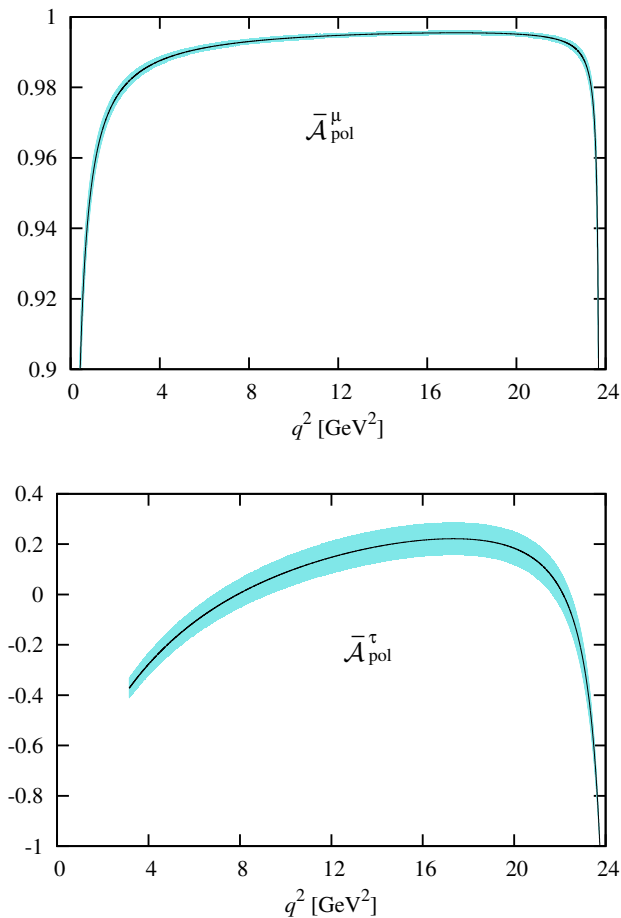
FIG. 15: Standard model ℓ-polarization fraction for the differential decay rate of $B_s \to K\ell\nu$, for $\ell = \mu, \tau$.

for the fictitious decay $B_s \to \eta_s$. Fitting these data simultaneously accounts for correlations — useful for constructing ratios of form factors. We extrapolate our lattice form factor results to the continuum, to physical quark mass, and over the full kinematic range of q^2 using a combination of the modified z-expansion and HPChPT that we refer to as the HPChPT z-expansion.

We then make standard model predictions for:

- differential decay rates divided by $|V_{ub}|^2$, an observable that, when combined with experiment, will allow an alternative semileptonic exclusive determination of $|V_{ub}|$,

- differential branching fractions using both the inclusive and exclusive semileptonic $B \to \pi\ell\nu$ determinations of $|V_{ub}|$,

- the ratio of differential branching fractions $R_\mu^\tau(q^2)$,

- the forward-backward asymmetry, using inclusive and exclusive values of $|V_{ub}|$,

- the normalized forward-backward asymmetry,

- the τ-polarization distribution in the differential decay rate for $B_s \to K\tau\nu$, and

- the ℓ-polarization fraction in the differential decay rate for $B_s \to K\ell\nu$, for $\ell = \mu, \tau$.

In Appendix D we construct ratios of form factors for $B_s \to K$ with those for $B_s \to \eta_s$. In combination with a future calculation of $B_s \to \eta_s$ using HISQ b, these ratios can provide a nonperturbative determination of the $b \to u$ current matching factor. This would be relevant for both $B_s \to K\ell\nu$ and $B \to \pi\ell\nu$ simulations using NRQCD b quarks.

Our results, built on first principles lattice QCD form factors, greatly clarify standard model expectations [27] based on model estimates of form factors [21, 22, 28], most notably at large q^2. Combining our form factors, which are most precise at large q^2, with model calculations, typically more reliable at low q^2, would result in a more precise determination of f_0 and f_+. We are studying the possibility of further refining $B_s \to K\ell\nu$ standard model predictions using such form factors.

Appendix A: Fitting Basics

Here we describe in more detail two aspects of our statistical analysis: 1) the definition of our error budgets for fit results; and 2) the technique for chained fits of multiple data sets. We also discuss a general procedure for testing fit procedures. These are general techniques applicable to many types of fitting problems [29]. Finally we illustrate these ideas with an example drawn from this paper.

1. Fits and Error Budgets

The formal structure of a least-squares problem involves fitting input data y_i with functions $f_i(p)$ by adjusting fit parameters p_α to minmize

$$\chi^2(p) = \sum_{ij} \Delta y(p)_i \left(\mathrm{cov}_y^{-1}\right)_{ij} \Delta y(p)_j, \qquad (A1)$$

where cov_{ij} is the covariance matrix for the input data and

$$\Delta y(p)_i \equiv f_i(p) - y_i. \qquad (A2)$$

There are generally two types of input data — actual data, and prior data for each fit parameter — but we lump these together here since they enter $\chi^2(p)$ in the same way. So the sums here over i and j are over all data and priors. Note that priors and data may be correlated in some problems.

The best-fit parameters $\overline{p}_\alpha$ are those that minimize χ^2:

$$\partial_\alpha \chi^2(\overline{p}) = 2\sum_{ij} \partial_\alpha f_i(\overline{p}) \left(\mathrm{cov}_y^{-1}\right)_{ij} \Delta y(\overline{p})_j = 0 \quad (A3)$$

where the derivative $\partial_\alpha \equiv \partial/\partial\overline{p}_\alpha$. The inverse covariance matrix, $\partial_\alpha\partial_\beta\chi^2(\overline{p})/2$, for the $\overline{p}_\alpha$ is then given by

$$\left(\mathrm{cov}_p^{-1}\right)_{\alpha\beta} = \sum_{ij} \partial_\alpha f_i(\overline{p}) \left(\mathrm{cov}_y^{-1}\right)_{ij} \partial_\beta f_j(\overline{p}) + \mathcal{O}(\Delta y),$$
$$(A4)$$

where we neglect terms proportional to Δy (which makes sense for reasonable fits to accurate data). This is the conventional result.

The uncertainties in the $\overline{p}_\alpha$ are due to the uncertainties in the input data y_i, and, for very accurate data, depend linearly upon cov_y. The relationship can be demonstrated by differentiating Eq. (A3) with respect to y_j to obtain

$$\sum_\beta \left(\mathrm{cov}_p^{-1}\right)_{\alpha\beta} \frac{\partial\overline{p}_\beta}{\partial y_j} = \sum_i \partial_\alpha f_i(\overline{p}) \left(\mathrm{cov}_y^{-1}\right)_{ij} + \mathcal{O}(\Delta y),$$
$$(A5)$$

where again we neglect terms proportional to Δy. Solving for $\partial\overline{p}_\beta/\partial y_j$ gives:

$$\frac{\partial\overline{p}_\beta}{\partial y_j} = \sum_{\alpha i} (\mathrm{cov}_p)_{\beta\alpha} \partial_\alpha f_i(\overline{p}) \left(\mathrm{cov}_y^{-1}\right)_{ij} \qquad (A6)$$

In the high-statistics, small-error limit the covariances in the $\overline{p}_\alpha$ are related to those in the y_i by the standard formula

$$(\mathrm{cov}_p)_{\alpha\beta} = \sum_{ij} \frac{\partial\overline{p}_\alpha}{\partial y_i} (\mathrm{cov}_y)_{ij} \frac{\partial\overline{p}_\beta}{\partial y_j}, \qquad (A7)$$

and, indeed, substituting Eq. (A6) into this equation reproduces Eq. (A4) for cov_p.

Eqs. (A6) and (A7) allow us to express the error σ_g for a function $g(\overline{p})$ of the best-fit parameter values in terms of the input errors:

$$\sigma_g^2 \equiv \sum_{\alpha\beta} \partial_\alpha g(\overline{p}) (\mathrm{cov}_p)_{\alpha\beta} \partial_\beta g(\overline{p}) = \sum_{ij} c_{ij} (\mathrm{cov}_y)_{ij} \;\; (A8)$$

where

$$c_{ij} \equiv \sum_{\alpha\beta} \partial_\alpha g(\overline{p}) \frac{\partial\overline{p}_\alpha}{\partial y_i} \frac{\partial\overline{p}_\beta}{\partial y_j} \partial_\beta g(\overline{p}). \qquad (A9)$$

and Eq. (A6) is used to evaluate $\partial\overline{p}_\alpha/\partial y_i$. We can then decompose σ_g^2 into separate contributions coming from the different block-diagonal submatrices of cov_y. These contributions to σ_g constitute the error budget for $g(\overline{p})$.

The c_{ij}s in Eq. (A8) depend upon both the y_i and their covariance matrix, but that dependence can be neglected to leading order in cov_y. Consequently Eq. (A8) can be used to estimate the impact on σ_g of possible modifications to any element of cov_y.

Note that the data's covariance matrix cov_y can be quite singular if there are strong correlations in the data. This can make it numerically difficult to invert the matrix for use in $\chi^2(p)$. This problem is typically dealt with by using a *singular value decomposition* (SVD) to regulate the most singular components of the covariance matrix. In our fits we rescale the covariance matrix by its diagonal elements to obtain the correlation matrix, which we then diagonalize. We introduce a minimum eigenvalue by setting any smaller eigenvalue equal to the minimum. We then reconstitute the correlation matrix, and rescale it back into a (less singular) covariance matrix which we use in the fit. This procedure, in effect, increases the error in the data and so increases the uncertainties in the final fit results; it is a conservative move.

It is common when using SVD to discard eigenmodes corresponding to the small eigenvalues. This is equivalent to setting the variance associated with these modes to infinity in the fit. In our implementation, all eigenmodes are retained, but the small eigenvalues are replaced by a (larger) minimum eigenvalue. This is a more realistic estimate for the variances of these modes — that is, more realistic than setting them to infinity — and gives more accurate fit results.

2. Chained Fits

Chained fits simplify fits of multiple data sets whose fit functions share fit parameters by allowing us to fit each

data set separately. To illustrate, consider two sets of data, $y_i(A)$ and $y_j(B)$, that we fit with functions $f_i(A, p)$ and $f_j(B, p)$, respectively — both functions of the same fit parameters p_α (unlike the previous section, here we do not lump the priors in with the ys). The fit procedure is straightforward in a Bayesian framework if $y(A)$ and $y(B)$ are statistically uncorrelated. We first fit, say, data set $y(A)$ to obtain best-fit estimates $\overline{p}(A)$ for the parameters and an estimate $\mathrm{cov}_{p(A)}$ for the parameters' covariance matrix. We then fit data set $y(B)$, but using $\overline{p}(A)$ and $\mathrm{cov}_{p(A)}$ to form the prior for the fit parameters.

This two-step fit merges the information contained in $y(A)$ with that from $y(B)$ by feeding the information from the first fit into the second fit as prior information. The order in which the data sets are fit doesn't matter in the high-statistics (Gaussian) limit; with larger errors, it is better to fit the more accurate data set first. The χ^2 for the two-step fit is the sum of the χ^2s for each step.

The situation is slightly more complicated if $y(A)$ and $y(B)$ are correlated. Then the best-fit parameters $\overline{p}(A)$ from the first fit above are correlated with the second data set $y(B)$. The $p(A)$-$y(B)$ covariance can be computed from

$$\mathrm{cov}_{p(A)y(B)} \equiv \sum_{y(A)} \frac{\partial \overline{p}(A)}{\partial y(A)} \mathrm{cov}_{y(A)y(B)} \qquad (A10)$$

using Eq. (A6) in the previous section. This correlation must be included in the second fit, to data set $y(B)$. So the second fit uses the best-fit parameters $\overline{p}(A)$ from the first fit to construct the prior, together with $\mathrm{cov}_{p(A)}$ for parameter-parameter covariances and $\mathrm{cov}_{p(A)y(B)}$ for parameter-data covariances.

We refer to a sequential fit of multiple data sets, where the best-fit parameters and covariance matrix from one fit are used as the prior for the next fit, as a *chained fit*. It is essential in such fits to account for possible correlations between the priors (from previous fits) and the data being fit at each stage. The results of a chained fit should agree with those of a simultaneous fit in the limit of large (*i.e.*, Gaussian) statistics.

3. Testing Fits

It is generally useful to have ways of testing particular fit strategies. One simple approach to testing is to create multiple fake data sets that are very similar to the actual data being fit, but where the exact values for the fit parameters are known ahead of time. Running several such data sets through an analysis code tells you very quickly whether, for example, your analysis code gives results that are correct to within one sigma 68% of the time, as is desired.

It is easy to create fake data sets of this sort. One simple recipe is the following:

1. Fit the actual data to obtain a set of parameter values p_α^* such that the fit function $f_i(p^*)$ closely matches the mean values y_i of the actual data. Calculate the difference between the actual means of the data and the fit values for $p = p^*$:

$$\delta y_i \equiv f_i(p^*) - y_i. \qquad (A11)$$

2. Create a bootstrap copy y_i^{bs} of the original data and replace its mean values by:

$$y_i^* = y_i^{\mathrm{bs}} + \delta y_i \qquad (A12)$$

The fake data set then consists of the mean values y_i^* and the covariance matrix cov_y of the original data. The role of the bootstrap here is to generate fluctuations in the means with the same distribution as the original data. These data sets will fluctuate around central values $f_i(p^*)$ rather than the original means of the data.

3. Repeat the second step to create any number of additional fake data sets.

Each fake data set is fit using the same procedure that was used to analyze the original data. The results for the fit parameters are compared with the parameter values p^* used to define the correction δy_i (Eq. (A11)), since, by construction, these are the *correct* values for the parameters in the fake data.

Typically only a handful of parameters from a fit are of interest. Their best-fit values from different fake data sets will differ, but they should all agree with the p^* values to within the errors generated by the fake fit (that is, to within one sigma 68% of the time, two sigma 95% of the time, and so on). Such tests can reveal, for example, potential problems coming from poor priors or inadequate SVD cuts, or biases in particular combinations of fit parameters.

4. Example

We compare chained and unchained fit results in Fig. 16. Because unchained fits to very large data sets are unreliable, for purposes of comparison we divide the data into the smallest subsets that allow the extraction of individual matrix elements. Such fits are uncorrelated in that they neglect correlations among data at different momenta, for different currents, and among the two decays. The uncorrelated fits include only one decay mode ($B_s \to K$ or $B_s \to \eta_s$), data for only one simulation momentum (000, 100, 110, or 111), and only one current (V_t or V_k). These fits are still complicated, however, as they require the minimum amount of data needed to extract a single matrix element. This minimum number of correlators consists of parent and daughter two point and three point data, *i.e.* $B_s \to B_s$, $\eta_s \to \eta_s(000)$, and $B_s \to V_t \to \eta_s(000)$. Including correlations results in marked improvement in the accuracy of matrix elements obtained from the noisiest data — that for V_k at
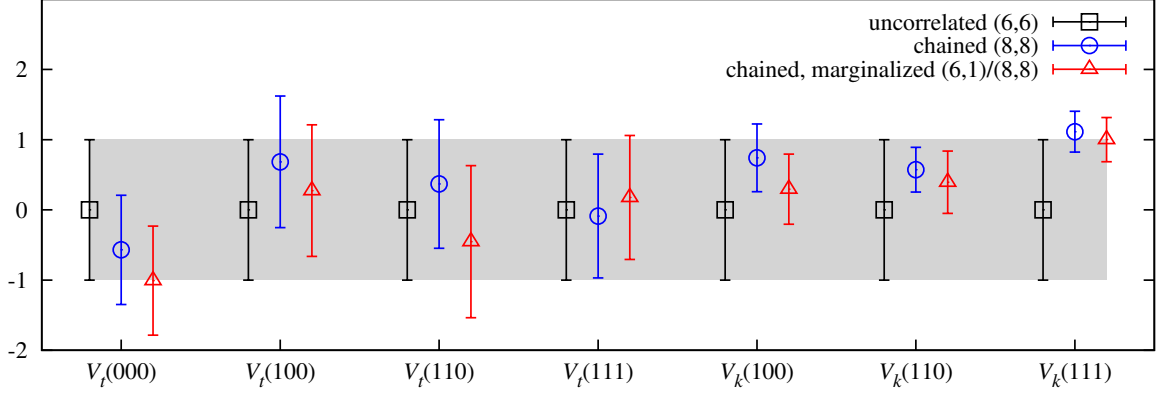
FIG. 16: $B_s \to \eta_s$ three point ground state amplitudes, for varying currents and momenta, as obtained from different fitting strategies described in the text. Plotted central values indicate the number of standard deviations by which a fit result differs from an "uncorrelated" fit. The size of the error bars is the ratio of the plotted fit error to that from an uncorrelated fit.

large momenta. This improvement can be traced to correlations of these data with the more precise data for V_t (for the same decay and at a common momentum), as demonstrated in Fig. 3.

In addition to properly accounting for correlations in the data, chaining reduces the time required to perform the fits. While the uncorrelated fits required a total of 1 hour 14 minutes, the chained (8,8) fit required only 24 minutes. The use of marginalization significantly reduces the time required. The chained and marginalized (6,1)/(8,8) fit required only 57 seconds.

ens	prior	2pt	2+3pt
C1	0.537(53)	0.53780(72)	0.53801(31)
C2	0.54(6)	0.54360(84)	0.54234(35)
C3	0.54(8)	0.5362(15)	0.53575(36)
F1	0.405(55)	0.4081(13)	0.40869(21)
F2	0.407(60)	0.40770(64)	0.40710(23)

TABLE IV: B_s priors and fit results for $aE_{B_s}^{\mathrm{sim}(0)}$.

ens	$aM_K^{(0)}$	$aE_{K(100)}^{(0)}$	$aE_{K(110)}^{(0)}$	$aE_{K(111)}^{(0)}$
C1	0.312(17)	0.41(11)	0.48(23)	0.55(28)
	0.31211(15)	0.40657(58)	0.48461(76)	0.5511(16)
	0.31195(14)	0.40661(49)	0.48408(63)	0.5513(13)
C2	0.329(24)	0.45(15)	0.55(15)	0.61(31)
	0.32863(18)	0.45406(85)	0.5511(16)	0.6261(75)
	0.32870(16)	0.45434(73)	0.5506(11)	0.6273(35)
C3	0.356(25)	0.475(75)	0.58(20)	0.65(30)
	0.35717(22)	0.47521(85)	0.5723(11)	0.6524(30)
	0.35744(21)	0.47507(71)	0.57218(80)	0.6539(18)
F1	0.229(60)	0.32(24)	0.39(34)	0.43(40)
	0.22865(11)	0.32024(66)	0.39229(86)	0.4515(25)
	0.22861(12)	0.32020(61)	0.39192(82)	0.4528(16)
F2	0.246(36)	0.33(23)	0.40(30)	0.47(37)
	0.24577(13)	0.33322(52)	0.40214(73)	0.4623(14)
	0.24566(13)	0.33310(50)	0.40184(72)	0.4624(11)

TABLE V: K priors and fit results. For each ensemble, the first row lists priors, the second row gives two point correlator fit results, and the third row shows simultaneous two and three point correlator fit results.

ens	$aM_{\eta_s}^{(0)}$	$aE_{\eta_s(100)}^{(0)}$	$aE_{\eta_s(110)}^{(0)}$	$aE_{\eta_s(111)}^{(0)}$
C1	0.411(9)	0.487(12)	0.553(50)	0.61(11)
	0.41111(12)	0.48736(23)	0.55311(29)	0.61148(60)
	0.41107(11)	0.48726(23)	0.55294(29)	0.61135(52)
C2	0.415(12)	0.52(5)	0.61(11)	0.68(23)
	0.41445(17)	0.51949(46)	0.6063(12)	0.6797(31)
	0.41446(15)	0.51934(44)	0.60647(67)	0.6794(18)
C3	0.412(20)	0.518(40)	0.61(12)	0.69(35)
	0.41180(23)	0.51757(63)	0.60723(78)	0.6831(23)
	0.41175(20)	0.51742(57)	0.60720(67)	0.6843(14)
F1	0.294(24)	0.37(10)	0.43(23)	0.48(34)
	0.294109(93)	0.36965(31)	0.43278(45)	0.4867(13)
	0.294066(88)	0.36988(26)	0.43301(38)	0.48729(88)
F2	0.293(30)	0.369(89)	0.43(18)	0.49(30)
	0.29315(12)	0.36939(35)	0.43259(45)	0.48810(87)
	0.29310(12)	0.36927(35)	0.43197(48)	0.48729(97)

TABLE VI: Like Table V but for the η_s.

Appendix B: Correlator Fit Results

The method for selecting priors for correlator fits was described in detail in Appendix B of Ref. [10]. We use the same method in this analysis. Tables IV, V, and VI tabulate priors and fit results for ground state energies. They compares results obtained from fits to two point correlation function data to those from simultaneous fits to two and three point correlation function data, as described in Sec. IV. The combined fits show improved precision for the B_s meson mass the larger momenta daughter meson energies, suggesting that the three point correlation function data provide additional information to the fit. Within errors, the two point and simultaneous two and three point fit results are consistent.

Appendix C: HPChPT z-Expansion Fit Results

Group I parameters listed in Table VII insert error in the fit based on uncertainty associated with input parameters – quantities not determined by the data. Priors for r_1 and $M_{\eta_s^{\text{phys}}}$ are taken from Ref. [30]. We base our prior choice for the $BB^*\pi$ coupling $g_{BB^*\pi}$ on the combined works in Ref. [31]. Resonance masses for the Blaschke factors $P_{0,+}$ introduced in Eq. (20) are calculated relative to the M_{B_s} meson mass in our simulations,

$$M_0^{B_s K} = M_{B_s} - (M_{B_s} - M_B) + 400(1)\,\text{MeV}, \quad \text{(C1)}$$
$$M_+^{B_s K} = M_{B_s} - (M_{B_s} - M_B) + \Delta_B^{\text{hyperfine}}, \quad \text{(C2)}$$
$$M_0^{B_s \eta_s} = M_{B_s} + 400(1)\,\text{MeV}, \quad \text{(C3)}$$
$$M_+^{B_s \eta_s} = M_{B_s} + \Delta_{B_s}^{\text{hyperfine}}, \quad \text{(C4)}$$

and we refer to the shift relative to M_{B_s} as $\Delta_{0,+}^{B_s K, B_s \eta_s}$. The $M_{B_s} - M_B$ and hyperfine splittings are taken from the PDG [25]. We tested increasing the uncertainty in the location of the scalar pole, which we have taken to be $400(1)$ MeV above the $J^P = 0^-$ state. A splitting of $400(50)$ MeV gives identical results for the form factors, in both the central value and error, but accommodates for part of the error in f_0 via allowed uncertainty in M_0. To reconstruct the form factors in this case, correlations between P_0 and the coefficients of the z-expansion must be accounted for. By effectively fixing M_0 we arrive at the same fit results and can neglect uncertainty in P_0 and correlations with the coefficients. The 4% uncertainty associated with the perturbative matching is accounted for by $m_{\parallel}$ and $m_{\perp}$, where we use prior central values of zero and width 0.04, as explained by Eq. (32) and surrounding text. Matrix elements for $B_s \to K$ and $B_s \to \eta_s$ use the same matching factors so we use common $m_{\parallel,\perp}$ for both data sets. We use values for r_1/a from Ref. [6] and $M_{\pi,K}^{\text{asqtad}}$ from Ref. [32]. We use values for $M_{\pi,K,\eta_s}^{\text{HISQ}}$ and M_{B_s} from best fit results in this and an ongoing $B \to \pi$ analysis using HISQ valence quarks.

The Group II parameters of Table VII are quantities determined by the fit. We choose priors for a_k to be 0 ± 5, based roughly on the unitarity constraint, and verified that fit results are insensitive to variations in the prior width from 1 to 10. Chiral analytic terms are written in terms of dimensionless parameters that are naturally $\mathcal{O}(1)$. For this reason we use priors of zero with width one for c_1 and c_3. Based on previous analyses using the same ensembles we know that sea-quark effects are smaller than those of the valence quarks, so we choose priors for c_2 to be 0 ± 0.3. The leading order HISQ discretization effects are $\mathcal{O}(\alpha_s a^2)$, so for the coefficients d_1 and e_1 which characterize the $\mathcal{O}(a^2)$ discretization effects, we choose priors of 0 ± 0.3. Coefficients d_2 and e_2 characterize $\mathcal{O}(a^4)$ effects and we use 0 ± 1. The coefficients h and l characterize light- and heavy-quark mass-dependent discretization effects. These terms are written in terms of $\mathcal{O}(1)$ quantities and we take the coefficients to have priors of 0 ± 1.

Group I	prior	fit
r_1 [fm]	0.3133(23)	0.3133(23)
$g_{B^*B\pi}$	0.51(20)	0.53(20)
$M_{\eta_s^{\text{phys}}}$ [GeV]	0.6858(40)	0.6858(40)
$\Delta_0^{B_s K}$ [GeV]	0.3127(10)	0.3126(10)
$\Delta_+^{B_s K}$ [GeV]	-0.04157(42)	-0.04157(42)
$\Delta_0^{B_s \eta_s}$ [GeV]	0.4000(10)	0.4000(10)
$\Delta_+^{B_s \eta_s}$ [GeV]	0.0487(22)	0.0487(22)
$m_{\parallel}$	0.00(4)	0.000(40)
$m_{\perp}$	0.00(4)	0.001(40)
r_1/a	2.647(3)	2.6465(30)
r_1/a	2.618(3)	2.6186(30)
r_1/a	2.644(3)	2.6438(30)
r_1/a	3.699(3)	3.6992(30)
r_1/a	3.712(4)	3.7117(40)
aM_{B_s}	3.2303(12)	3.2300(12)
aM_{B_s}	3.2663(13)	3.2668(12)
aM_{B_s}	3.2336(13)	3.2333(12)
aM_{B_s}	2.30849(89)	2.30841(87)
aM_{B_s}	2.30035(90)	2.30048(88)
aM_K^{HISQ}	0.31195(14)	0.31196(14)
aM_K^{HISQ}	0.32870(17)	0.32868(17)
aM_K^{HISQ}	0.35744(21)	0.35746(21)
aM_K^{HISQ}	0.22861(12)	0.22861(12)
aM_K^{HISQ}	0.24566(13)	0.24565(13)
aM_K^{asqtad}	0.36530(29)	0.36532(29)
aM_K^{asqtad}	0.38331(24)	0.38331(24)
aM_K^{asqtad}	0.40984(21)	0.40983(21)
aM_K^{asqtad}	0.25318(19)	0.25316(19)
aM_K^{asqtad}	0.27217(21)	0.27219(21)
aM_π^{HISQ}	0.15988(12)	0.15988(12)
aM_π^{HISQ}	0.21097(16)	0.21097(16)
aM_π^{HISQ}	0.29309(22)	0.29309(22)
aM_π^{HISQ}	0.13453(11)	0.13453(11)
aM_π^{HISQ}	0.18737(13)	0.18736(13)
aM_π^{asqtad}	0.15971(20)	0.15971(20)
aM_π^{asqtad}	0.22447(17)	0.22447(17)
aM_π^{asqtad}	0.31125(16)	0.31125(16)
aM_π^{asqtad}	0.14789(18)	0.14789(18)
aM_π^{asqtad}	0.20635(18)	0.20365(18)
$aM_{\eta_s}^{\text{HISQ}}$	0.41107(11)	0.41109(11)
$aM_{\eta_s}^{\text{HISQ}}$	0.41447(15)	0.41442(15)
$aM_{\eta_s}^{\text{HISQ}}$	0.41176(20)	0.41177(20)
$aM_{\eta_s}^{\text{HISQ}}$	0.294066(89)	0.294053 (89)
$aM_{\eta_s}^{\text{HISQ}}$	0.29310(12)	0.29312(12)

TABLE VII: Group I priors for the HPChPT z-expansion for $f_{0,+}^{B_s K}$ and $f_{0,+}^{B_s \eta_s}$. Quantities listed in five consecutive rows have ensemble-dependent values corresponding to C1, C2, C3, F1, and F2.

Group II	prior	$f_0^{B_s K}$	$f_+^{B_s K}$	$f_0^{B_s \eta_s}$	$f_+^{B_s \eta_s}$	Group II	prior	$f_0^{B_s K}$	$f_+^{B_s K}$	$f_0^{B_s \eta_s}$	$f_+^{B_s \eta_s}$
			fit result						fit result		
a_0	0(5)	0.24(10)	0.284(32)	0.04(12)	0.293(30)	$h_1^{(0)}$	0(1)	0.31(92)	0.37(92)	0.0(1.0)	0.22(95)
a_1	0(5)	0.7(1.0)	-0.58(16)	0.0(1.2)	-0.99(18)	$h_1^{(1)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
a_2	0(5)	1.9(3.6)	2.1(1.1)	2.1(4.3)	3.2(1.7)	$h_1^{(2)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$c_1^{(0)}$	0(1)	0.01(60)	0.07(11)	-0.23(99)	-0.16(15)	$h_2^{(0)}$	0(1)	0.20(99)	0.02(99)	0.0(1.0)	-0.15(99)
$c_1^{(1)}$	0(1)	0.11(90)	-0.16(38)	0.0(1.0)	-0.39(25)	$h_2^{(1)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$c_1^{(2)}$	0(1)	-0.04(99)	-0.62(85)	0.11(98)	-1.25(86)	$h_2^{(2)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$c_2^{(0)}$	0(0.3)	-0.24(27)	0.05(29)	-0.03(27)	0.15(29)	$h_3^{(0)}$	0(1)	0.0(1.0)	0.1(1.0)	0.0(1.0)	0.0(1.0)
$c_2^{(1)}$	0(0.3)	0.00(30)	-0.02(30)	0.00(30)	-0.01(30)	$h_3^{(1)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$c_2^{(2)}$	0(0.3)	0.00(30)	-0.01(30)	0.00(30)	-0.01(30)	$h_3^{(2)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$c_3^{(0)}$	0(1)	0.37(99)	-1.22(74)	0.0(1.0)	-0.19(70)	$h_4^{(0)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$c_3^{(1)}$	0(1)	0.0(1.0)	0.34(97)	0.0(1.0)	-0.24(94)	$h_4^{(1)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$c_3^{(2)}$	0(1)	0.1(1.0)	-0.20(99)	0.0(1.0)	0.00(99)	$h_4^{(2)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$d_1^{(0)}$	0(0.3)	0.16(18)	-0.20(22)	-0.02(21)	-0.15(24)	$l_1^{(0)}$	0(1)	0.64(0.97)	0.18(98)	0.0(1.0)	0.24(98)
$d_1^{(1)}$	0(0.3)	-0.01(30)	-0.06(30)	0.00(30)	-0.05(29)	$l_1^{(1)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$d_1^{(2)}$	0(0.3)	0.01(30)	-0.03(30)	0.00(30)	-0.01(30)	$l_1^{(2)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$d_2^{(0)}$	0(1)	-0.22(92)	-0.32(94)	-0.02(85)	-0.21(94)	$l_2^{(0)}$	0(1)	0.1(1.0)	0.0(1.0)	0.0(1.0)	0.1(1.0)
$d_2^{(1)}$	0(1)	0.0(1.0)	-0.1(1.0)	0.0(1.0)	-0.07(99)	$l_2^{(1)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$d_2^{(2)}$	0(1)	0.0(1.0)	-0.1(1.0)	0.0(1.0)	0.0(1.0)	$l_2^{(2)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$e_1^{(0)}$	0(0.3)	-0.21(17)	0.13(24)	-0.09(16)	0.15(23)	$l_3^{(0)}$	0(1)	-0.1(1.0)	0.0(1.0)	0.0(1.0)	0.1(1.0)
$e_1^{(1)}$	0(0.3)	-0.01(30)	0.00(29)	0.00(30)	-0.06(28)	$l_3^{(1)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$e_1^{(2)}$	0(0.3)	0.00(30)	-0.02(30)	0.00(30)	-0.02(30)	$l_3^{(2)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$e_2^{(0)}$	0(1)	0.40(24)	0.12(30)	0.26(19)	0.02(25)	$l_4^{(0)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$e_2^{(1)}$	0(1)	-0.1(1.0)	0.25(94)	0.0(1.0)	-0.04(83)	$l_4^{(1)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)
$e_2^{(2)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	-0.03(99)	$l_4^{(2)}$	0(1)	0.0(1.0)	0.0(1.0)	0.0(1.0)	0.0(1.0)

TABLE VIII: Group II priors and fit results for the simultaneous HPChPT z-expansion for $f_{0,+}^{B_s K}$ and $f_{0,+}^{B_s \eta_s}$.

ens	$f_0^{B_s\eta_s}(000)$	$f_0^{B_s\eta_s}(100)$	$f_0^{B_s\eta_s}(110)$	$f_0^{B_s\eta_s}(111)$
C1	0.8135(17)	0.7352(22)	0.6813(19)	0.6381(21)
C2	0.8205(21)	0.7127(33)	0.6475(39)	0.5921(70)
C3	0.8140(26)	0.7095(32)	0.6504(31)	0.6069(39)
F1	0.8179(20)	0.7107(23)	0.6410(26)	0.5862(47)
F2	0.8229(24)	0.7096(31)	0.6383(33)	0.5874(51)

ens		$f_+^{B_s\eta_s}(100)$	$f_+^{B_s\eta_s}(110)$	$f_+^{B_s\eta_s}(111)$
C1		1.843(10)	1.5476(62)	1.3400(63)
C2		1.742(13)	1.3885(99)	1.150(17)
C3		1.6802(95)	1.3855(84)	1.1771(85)
F1		1.6928(71)	1.3497(55)	1.134(10)
F2		1.7012(97)	1.3588(72)	1.155(11)

TABLE IX: Fit results for the scalar and vector $B_s \to \eta_s$ form factors on each ensemble and for each simulated momentum.

Appendix D: $B_s \to \eta_s$ Form Factors and Ratios

The results of $B_s \to \eta_s$ correlator fits are tabulated in Table IX and plotted as data points in the top two panels of Fig. 17. From these plots one sees that simulation data exhibit very small light sea quark mass and lattice spacing dependence. These fit results are obtained from a single fit to both the $B_s \to K$ and $B_s \to \eta_s$ data described in Sec. IV. As a result, the $B_s \to \eta_s$ fit results of Table IX are correlated with the $B_s \to K$ results of Table II, as shown in Fig. 3.

The $B_s \to \eta_s$ form factor data of Table IX is extrapolated to the physical quark mass, the continuum limit, and over the entire kinematic range using the HPChPT z-expansion described in Sec. V. This fit is also done simultaneously with the extrapolation of the $B_s \to K$ data. The fit functions for the simultaneous chiral, continuum, and kinematic extrapolation of $B_s \to \eta_s$ are equivalent to those of Sec. V, with Eqs. (25, 26) modified as follows:

$$
\begin{aligned}
D_k &= 1 + c_1^{(k)} x_K + c_2^{(k)} \delta x_K + c_3^{(k)} \delta x_{\eta_s} \\
&+ d_1^{(k)} (a/r_1)^2 + d_2^{(k)} (a/r_1)^4 \\
&+ e_1^{(k)} (aE_{\eta_s})^2 + e_2^{(k)} (aE_{\eta_s})^4, \quad \text{(D1)}
\end{aligned}
$$

$$
[\text{logs}] = -\frac{1+3g^2}{2} x_K \log x_K - \frac{1+3g^2}{6} x_\eta \log x_\eta, \quad \text{(D2)}
$$

with implicit indices in Eq. (D1) specifying scalar or vector form factor. Results of this fit for the $B_s \to \eta_s$ form factors are shown relative to data, and extrapolated over the full kinematic range of q^2, in Fig. 17. The HPChPT z-expansion stability analysis outlined in Sec. V involved simultaneous fits to both $B_s \to K$ and $B_s \to \eta_s$ data. The $B_s \to \eta_s$ fit results for each of the modifications discussed in that analysis are shown in Fig. 18. Because these results are from a simultaneous fit, the values of χ^2 in Fig. 5 are applicable here as well and are reproduced
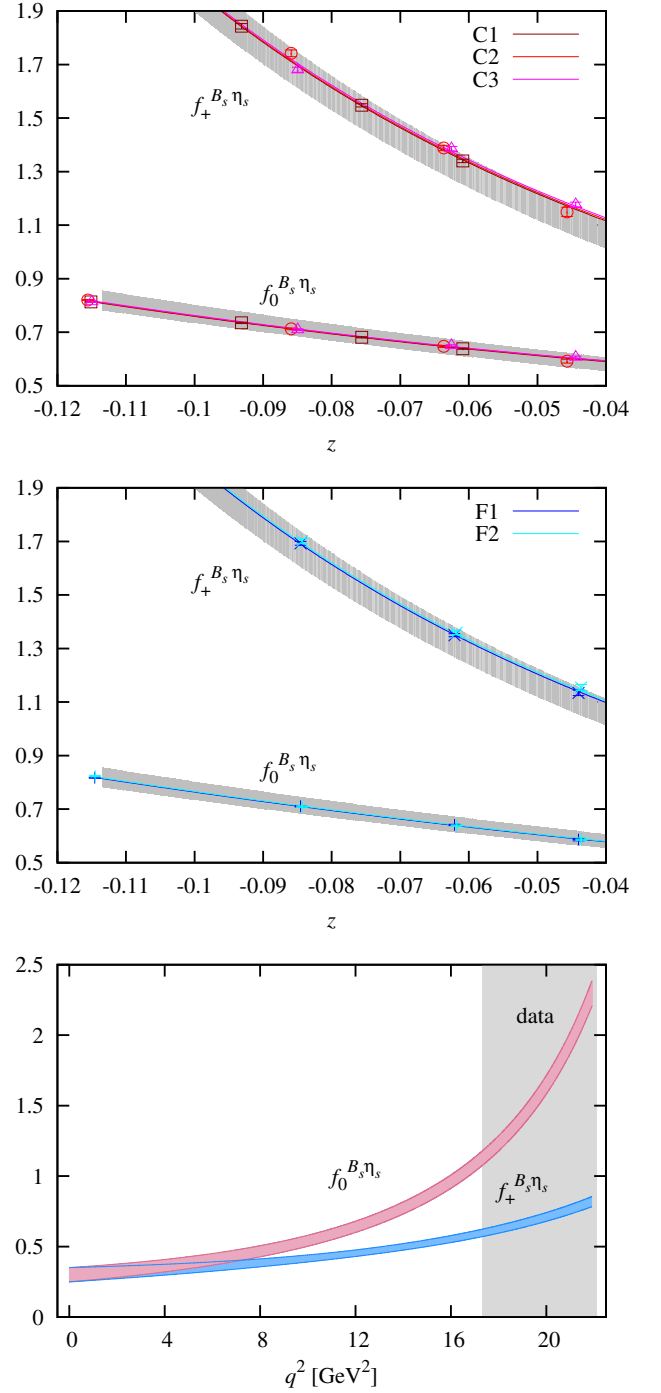

FIG. 17: (*color online*) $B_s \to \eta_s$ form factor results from a simultaneous HPChPT z-expansion are shown (*top*) relative to coarse ensemble data (C1, C2, and C3), (*middle*) relative to fine ensemble data (F1 and F2), and (*bottom*) in the continuum limit with physical masses, extrapolated over the full kinematic range.

for convenience in Fig. 18. Note that the chiral analytic terms for $B_s \to \eta_s$ differ slightly from those for $B_s \to K$, *c.f.* Eqs. (25, D1). As a result, the NNLO analytic terms
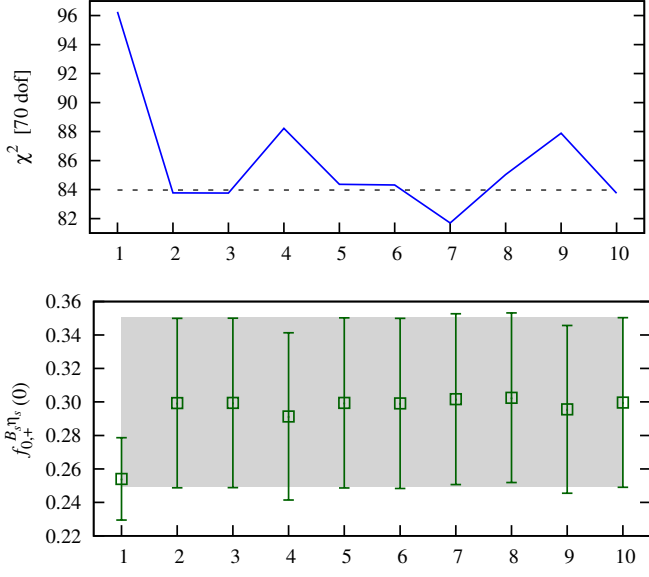
FIG. 18: The stability of the HPChPT z-expansion is demonstrated by studying the fit results under various modifications, discussed in Sec. V of the text.

added to the $B_s \to \eta_s$ fit function in modification 7 differ from those listed in Eq. (34). Error breakdown plots for the $B_s \to \eta_s$ form factors are shown in Fig. 19. In the ratios of form factors,

$$R_\parallel(q^2) = \frac{f_\parallel^{B_s K}(q^2)}{f_\parallel^{B_s \eta_s}(q^2)}, \qquad (D3)$$

$$R_\perp(q^2) = \frac{f_\perp^{B_s K}(q^2)}{f_\perp^{B_s \eta_s}(q^2)}, \qquad (D4)$$

the leading systematic error, that due to one-loop perturbative matching, largely cancels. Fig. 20 plots the ratios as functions of q^2 and shows that they are most precisely determined at $q^2 = (M_{B_s} - M_{\eta_s})^2$, where

$$R_\parallel\big((M_{B_s} - M_{\eta_s})^2\big) = 0.821(22), \qquad (D5)$$

$$R_\perp\big((M_{B_s} - M_{\eta_s})^2\big) = 0.931(30). \qquad (D6)$$

Neglecting correlations among the $B_s \to K$ and $B_s \to \eta_s$ decays yields ratios at this q^2 with $\sim 30\%$ larger errors. When combined with lattice results for $f_{\parallel,\perp}^{B_s \eta_s}$ using HISQ b quarks, these ratios will provide a nonperturbative determination of the NRQCD $b \to u$ current matching factor, applicable to both $B_s \to K$ and $B \to \pi$.
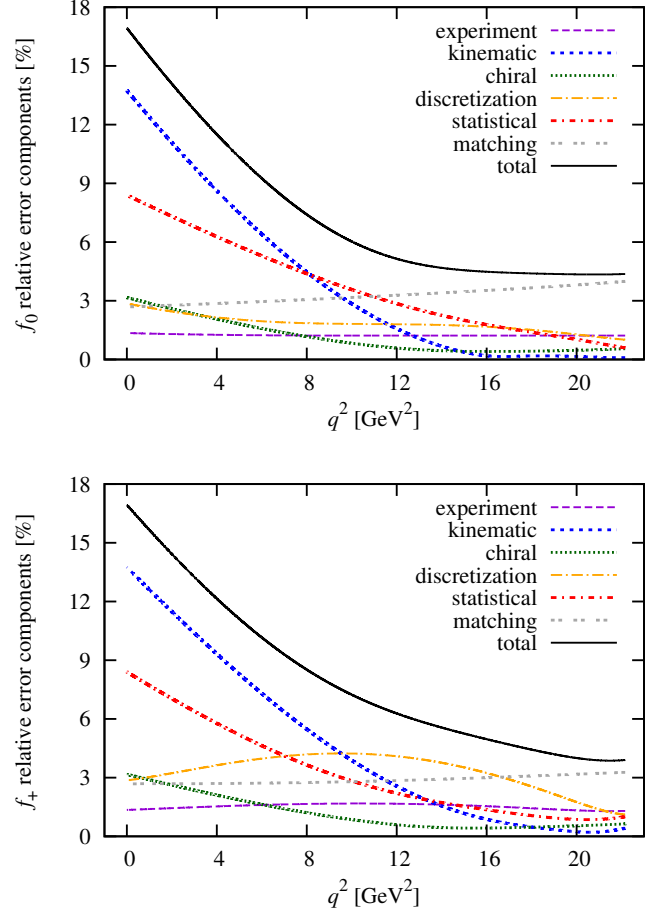

FIG. 19: (*color online*) $B_s \to \eta_s$ (*top*) f_0 and (*bottom*) f_+ relative error components. The total error (solid line) is the sum in quadrature of the components.
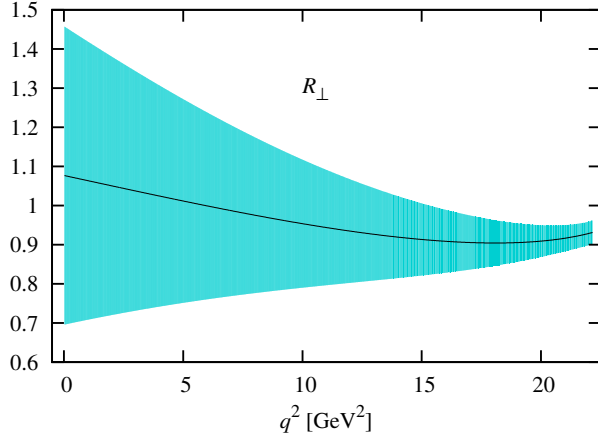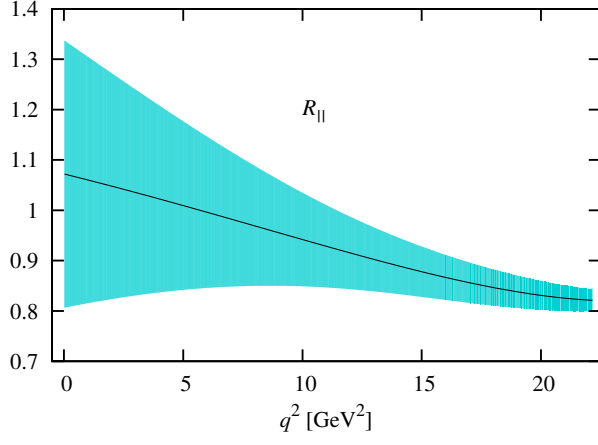
FIG. 20: Ratios of $B_s \to K$ to $B_s \to \eta_s$ form factors $R_{\parallel, \perp}$ as functions of q^2.
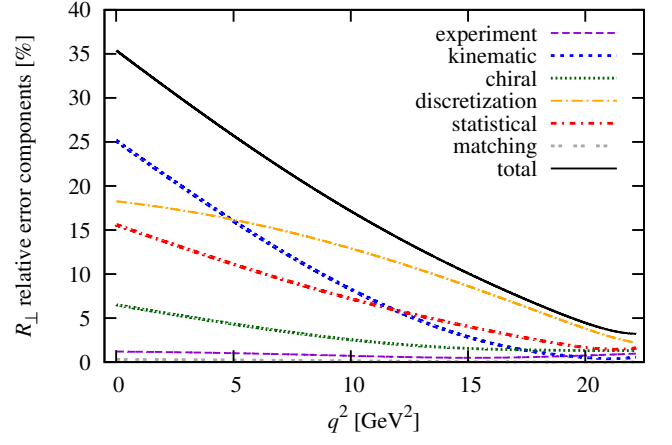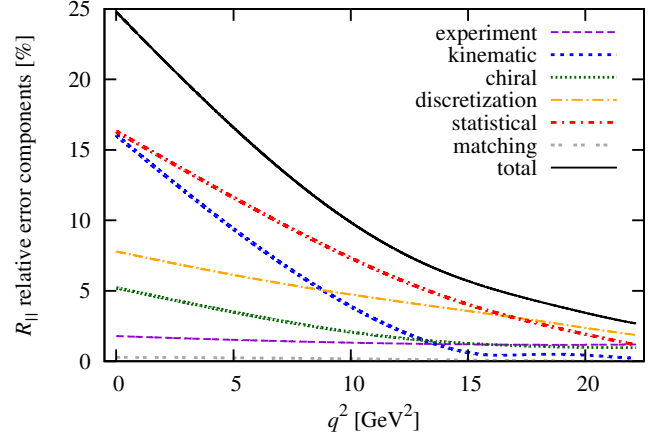

FIG. 21: (*color online*) Relative error components for (*top*) $R_{\parallel}$ and (*bottom*) $R_{\perp}$ as a function of q^2. The total error is the sum in quadrature of the components.

[1] K. Hornbostel, G.P. Lepage, C.T.H. Davies, R.J. Dowdall, H. Na, and J. Shigemitsu (HPQCD), Phys. Rev. D 85, 031504 (2012) [arXiv:1111.1363 [hep-lat]]

[2] H. Na, C. T. H. Davies, E. Follana, G. P. Lepage, and J. Shigemitsu (HPQCD), Phys. Rev. D 82, 114506 (2010) [arXiv:1008.4562 [hep-lat]]

[3] H. Na, C. T. H. Davies, E. Follana, J. Koponen, G. P. Lepage, and J. Shigemitsu (HPQCD), Phys. Rev. D 84, 114505 (2011) [arXiv:1109.1501 [hep-lat]]

[4] J. Bijnens and I. Jemos, Nucl. Phys. B 840, 54 (2010); Erratum-ibid. B 844, 182 (2011) [arXiv:1006.1197 [hep-ph]]

[5] J. Bijnens and I. Jemos, Nucl. Phys. B 846, 145 (2011) [arXiv:1011.6531 [hep-ph]]

[6] A. Bazavov, C. Bernard, C. DeTar, S. Gottlieb, U. M. Heller, J. E. Hetrick, J. Laiho, L. Levkova, P. B. Mackenzie, M. B. Oktay, R. Sugar, D. Toussaint, and R. S. Van de Water (MILC), Rev. Mod. Phys. 82, 1349 (2010) [arXiv:0903.3598 [hep-lat]]

[7] G. P. Lepage, L. Magnea, C. Nakhleh, U. Magnea, and K. Hornbostel (HPQCD), Phys. Rev. D 46, 4052 (1992) [arXiv:hep-lat/9205007]

[8] H. Na, C. J. Monahan, C. T. H. Davies, R. Horgan, G. P. Lepage and J. Shigemitsu (HPQCD), Phys. Rev. D 86, 034506 (2012) [arXiv:1202.4914 [hep-lat]]

[9] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD), Phys. Rev. D 75, 054502 (2007) [arXiv:hep-lat/0610092]

[10] C. M. Bouchard, G. P. Lepage, C. Monahan, H. Na, and J. Shigemitsu (HPQCD), Phys. Rev. D 88, 054509 (2013); Erratum-ibid. D 88, 079901 (2013) [arXiv:1306.2384 [hep-lat]]

[11] C. Monahan, J. Shigemitsu, and R. Horgan (HPQCD), Phys. Rev. D 87, 034017 (2013) [arXiv:1211.6966 [hep-lat]]

[12] E. Gulez, A. Gray, M. Wingate, C. T. H. Davies, G. P. Lepage, and J. Shigemitsu (HPQCD), Phys. Rev. D 73, 074502 (2006); Erratum-ibid D 75, 119906 (2007) [arXiv:hep-lat/0601021]

[13] E. B. Gregory, C. T. H. Davies, I. D. Kendall, J. Koponen, K. Wong, E. Follana, E. Gámiz, G. P. Lepage, E. H. Müller, H. Na, and J. Shigemitsu (HPQCD), Phys. Rev. D 83, 014506 (2011) [arXiv:1010.3848 [hep-lat]]

[14] G. P. Lepage, B. Clark, C. T. H. Davies, K. Hornbostel, P. B. Mackenzie, C. Morningstar and H. Trottier, Nucl. Phys. Proc. Suppl. 106, 12 (2002) [arXiv:hep-lat/0110175]

[15] C. G. Boyd, B. Grinstein, and R. F. Lebed, Phys. Rev. Lett. 74, 4603 (1995) [arXiv:hep-ph/9412324]

[16] M. C. Arnesen, B. Grinstein, I. Z. Rothstein, and I. W. Stewart, Phys. Rev. Lett. 95, 071802 (2005) [arXiv:hep-ph/0504209]

[17] C. Bourrely, I. Caprini, and L. Lellouch, Phys. Rev. D 79, 013008 (2009); Erratum-ibid. D 82, 099902 (2010) [arXiv:0807.2722 [hep-ph]]

[18] G. Colangelo, M. Procura, L. Rothen, R. Stucki, and J. Tarrus JHEP 09 (2012) 081 [arXiv:1208.0498 [hep-ph]]

[19] C. Bernard (MILC), Phys. Rev. D 65, 054031 (2002) [arXiv:hep-lat/0111051]

[20] C. T. H. Davies, C. McNeile, E. Follana, G. P. Lepage, H. Na, and J. Shigemitsu (HPQCD), Phys. Rev. D 82 114504 (2010) [arXiv:1008.4018 [hep-lat]]

[21] W.-F. Wang and Z.-J. Xiao, Phys. Rev. D 86, 114025 (2012) [arXiv:1207.0265 [hep-ph]]

[22] R. N. Faustov and V. O. Galkin, Phys. Rev. D 87, 094028 (2013) [arXiv:1304.3255 [hep-ph]]

[23] Y. S. Tsai, Nucl. Phys. B, Proc. Suppl. 55, 293 (1997)

[24] C.-H. Chen and C.-Q. Geng, JHEP 10 (2006) 053 [arXiv:hep-ph/0608166]

[25] J. Beringer et al. (Particle Data Group), Phys. Rev. D 86, 010001 (2012) and 2013 partial update for the 2014 edition [http://pdg.lbl.gov]

[26] S. Aoki et al. (FLAG Working Group), arXiv:1310.8555 [hep-lat] [http://itpwiki.unibe.ch/flag]; J. A. Bailey et al. (FNAL Lattice and MILC), Phys. Rev. D 79, 054507 (2009) [arXiv:0811.3640]; E. Gulez, A. Gray, M. Wingate, C. T. H. Davies, G. P. Lepage, and J. Shigemitsu (HPQCD), Phys. Rev. D 73, 074502 (2006); Erratum-ibid. D 75, 119906 (2007) [hep-lat/0601021]

[27] Ulf-G. Meißner and W. Wang, JHEP 01 (2014) 107 [arXiv:1311.5420 [hep-ph]]

[28] R.C. Verma, J. Phys. G 39, 025005 (2012) [arXiv:1103.2973 [hep-ph]]

[29] The fitting software used in this paper is available online: see G. P. Lepage (2012). lsqfit v4.8.5.1. ZENODO. 10.5281/zenodo.10236 for a general package for nonlinear least squares fitting, and G. P. Lepage (2012). corrfitter v3.7.1. ZENODO. 10.5281/zenodo.10237 for a general package for fitting 2-point and 3-point correlators. This software implements the strategies discussed in Appendix A of this paper.

[30] C. T. H. Davies, E. Follana, I. D. Kendall, G. P. Lepage, and C. McNeile (HPQCD), Phys. Rev. D 81, 034506 (2010) [arXiv:0910.1229 [hep-lat]]

[31] H. Ohki, H. Matsufuru, and T. Onogi, Phys. Rev. D 77, 094509 (2008) [arXiv:0802.1563 [hep-lat]]; W. Detmold, C.-J. D. Lin, and S. Meinel, Phys. Rev. D 85, 114508 (2012) [arXiv:1203.3378 [hep-lat]] F. Bernardoni, J. Bulava, M. Donnellan, and R. Sommer (ALPHA), arXiv:1404.6951 [hep-lat];

[32] C. Aubin, C. Bernard, C. DeTar, J. Osborn, Steven Gottlieb, E. B. Gregory, D. Toussaint, U. M. Heller, J. E. Hetrick, and R. Sugar (MILC), Phys. Rev. D 70, 114501 (2004) [arXiv:hep-lat/0407028]