# $B \to D^* \ell \nu_\ell$ semileptonic form factors from lattice QCD with Möbius domain-wall quarks

Y. Aoki,[1] B. Colquhoun,[2] H. Fukaya,[3] S. Hashimoto,[4,5] T. Kaneko,[4,5,6]
R. Kellermann,[5] J. Koponen,[7] and E. Kou[8]

(JLQCD Collaboration)

[1]*RIKEN Center for Computational Science (R-CCS), Kobe 650-0047, Japan*
[2]*SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom*
[3]*Department of Physics, Osaka University, Osaka 560-0043, Japan*
[4]*High Energy Accelerator Research Organization (KEK), Ibaraki 305-0801, Japan*
[5]*Graduate Institute for Advanced Studies, SOKENDAI (The Graduate University for Advanced Studies),
Ibaraki 305-0801, Japan*
[6]*Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University,
Aichi 464-8602, Japan*
[7]*PRISMA+ Cluster of Excellence and Institute für Kernphysik, Johannes Gutenberg-Universität Mainz,
D-55128 Mainz, Germany*
[8]*Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France*

We calculate the form factors for the $B \to D^* \ell \nu_\ell$ decay in $2 + 1$ flavor lattice QCD. For all quark flavors, we employ the Möbius domain-wall action, which preserves chiral symmetry to a good precision. Our gauge ensembles are generated at three lattice cutoffs $a^{-1} \sim 2.5$, 3.6, and 4.5 GeV with pion masses as low as $M_\pi \sim 230$ MeV. The physical lattice size $L$ satisfies the condition $M_\pi L \geq 4$ to control finite volume effects (FVEs), while we simulate a smaller size at the smallest $M_\pi$ to directly examine FVEs. The bottom quark masses are chosen in a range from the physical charm quark mass to $0.7a^{-1}$ to control discretization effects. We extrapolate the form factors to the continuum limit and physical quark masses based on heavy meson chiral perturbation theory at next-to-leading order. Then the recoil parameter dependence is parametrized using a model independent form leading to our estimate of the decay rate ratio between the tau ($\ell = \tau$) and light lepton ($\ell = e, \mu$) channels $R(D^*) = 0.252(22)$ in the Standard Model. A simultaneous fit with recent data from the Belle experiment yields $|V_{cb}| = 39.19(91) \times 10^{-3}$, which is consistent with previous exclusive determinations, and shows good consistency in the kinematical distribution of the differential decay rate between the lattice and experimental data.

## I. INTRODUCTION

The $B \to D^* \ell \nu_\ell$ semileptonic decay, where $\ell = e, \mu, \tau$ and $\nu_\ell$ represents the corresponding neutrino, plays a key role in stringent tests of the Standard Model (SM) and searches for new physics. The channels associated with light leptons $\ell = e, \mu$ provide a determination of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element $|V_{cb}|$, which is a fundamental parameter of the SM. On the other hand, the $\tau$

channel is expected to be a good probe of new physics, since, for instance, it is expected to strongly couple to the charged Higgs boson predicted by supersymmetric models. Indeed, there is an intriguing $\gtrsim 3\sigma$ tension between the SM and experiments on the decay rate ratio

$$R(D^*) = \frac{\Gamma(B \to D^* \tau \nu_\tau)}{\Gamma(B \to D^* \ell \nu_\ell)} \quad (\ell = e, \mu) \qquad (1)$$

describing the lepton-flavor universality violation [1].

However, there has been a long-standing tension between the $|V_{cb}|$ determinations from exclusive and inclusive decays $(B \to D\ell\nu_\ell + D^*\ell\nu_\ell + D^{**}\ell\nu_\ell + \cdots)$. The Heavy Flavor Averaging Group reported that analysis of the inclusive decay yields [1]

$$|V_{cb}| \times 10^3 = 41.98(45), \qquad (2)$$

which shows a $\lesssim 4\sigma$ (9%) tension with the determination from the exclusive decays

$$|V_{cb}| \times 10^3 = \begin{cases} 39.14(99) & (B \to D\ell\nu_\ell), \\ 38.46(68) & (B \to D^*\ell\nu_\ell). \end{cases} \qquad (3)$$

It has been argued [2] that such a tension can be explained by introducing a tensor interaction beyond the SM, which, however, largely distorts the $Z \to \bar{b}b$ decay rate measured precisely by experiments. Therefore, it is likely that the $|V_{cb}|$ tension is due to our incomplete understanding of the theoretical and/or experimental uncertainty.

The largest theoretical uncertainty in the exclusive determination comes from form factors, which describe nonperturbative QCD effects to the decay amplitude through the matrix elements

$$\sqrt{M_B M_{D^*}}^{-1} \langle D^*(\epsilon', p')|V_\mu|B(p)\rangle = i\varepsilon_{\mu\nu\rho\sigma}\epsilon'^{*\nu}v'^\rho v^\sigma h_V(w),$$

$$(4)$$

$$\sqrt{M_B M_{D^*}}^{-1} \langle D^*(\epsilon', p')|A_\mu|B(p)\rangle$$
$$= (w+1)\epsilon'^*_\mu h_{A_1}(w)$$
$$- (\epsilon'^* v)\{v_\mu h_{A_2}(w) + v'_\mu h_{A_3}(w)\}, \qquad (5)$$

where $w = vv'$ is the recoil parameter defined by four velocities $v = p/M_B$ and $v' = p'/M_{D^{(*)}}$, and $\epsilon'$ is the polarization vector of $D^*$, which satisfies $p'\epsilon' = 0$. Note that, here and in the following, kinematical variables (momentum, velocity, polarization vector, but not space-time coordinates) with and without the symbol " $'$ " represent those for $D^*$ and $B$, respectively.

Lattice QCD is a powerful method to provide a first-principles calculation of the form factors with systematically improvable accuracy [3,4]. However, until recently, only $h_{A_1}(1)$ at the zero-recoil limit $w = 1$ had been calculated in unquenched lattice QCD [5–7]. In the previous determination of $|V_{cb}|$, therefore, other information of the form factors in addition to $|V_{cb}|$ had to be determined by fitting experimental data of the differential decay rate to a theoretical expression [1,8–10]. While a model independent parametrization of the form factors by Boyd, Grinstein, and Lebed (the BGL parametrization) is available [11], its large number of parameters due to the poor knowledge on the form factors makes the fit unstable [12–15] even with experimental data of differential distribution with respect to all kinematical variables, namely the recoil parameter and three decay angles, from the KEKB/Belle experiment [16–18].

One of the largest problems in the lattice study of $B$ meson physics is that the simulation cost rapidly increases as the lattice spacing decreases. In spite of recent progress

in computer performance and development of simulation algorithms, it is still difficult to simulate lattice cutoffs $a^{-1}$ sufficiently larger than the physical bottom quark mass $m_{b,\text{phys}}$ to suppress $O((am_b)^n)$ discretization errors to, say, the 1% level. For the time being, a practical approach to $B$ physics on the lattice is to employ a heavy quark action based on an effective field theory, such as the heavy quark effective theory (HQET), to directly simulate $m_{b,\text{phys}}$ at currently available cutoffs $a^{-1} \lesssim m_{b,\text{phys}}$. Another straightforward approach is the so-called relativistic approach, where a lattice QCD action is used also for heavy quarks but with their masses sufficiently smaller than the lattice cutoff to suppress discretization effects. The Fermilab/MILC reported the first lattice calculation of all form factors at zero and nonzero recoils [19] using the Fermilab approach for heavy quarks, namely a HQET interpretation of the anisotropic Wilson-clover quark action [20], and staggered light quarks. It was recently followed by the HPQCD collaboration with a fully relativistic approach [21], where the highly improved staggered quark action was used both for light and heavy flavors.

In this paper, we present an independent study with a fully relativistic approach using the Möbius domain-wall quark action [22] for all relevant quark flavors. This preserves chiral symmetry to good precision at finite lattice spacing, reducing the leading discretization errors to second order, i.e. $O((am_b)^2)$. We can also use chiral perturbation theory in the continuum limit ($a = 0$) to guide the chiral extrapolation to the physical pion mass. Note also that we do not need explicit renormalization of the weak axial and vector currents on the lattice to calculate relevant form factors, because the renormalization constants of these currents coincide with each other and are canceled by taking appropriate ratios of relevant correlation functions [23]. Our preliminary analyses and discussions on these features can be found in our earlier reports [24–26].

This paper is organized as follows. After a brief introduction to our choice of simulation parameters, the form factors are extracted from correlation functions at the simulation points in Sec. II. These results are extrapolated to the continuum limit and physical quark masses in Sec. III. In Sec. IV, we generate synthetic data of the form factors and fit them to a model-independent parametrization in terms of the recoil parameter to make a comparison with the previous lattice and phenomenological studies as well as to estimate the SM value of $R(D^*)$. We also perform a simultaneous fit with experimental data to estimate $|V_{cb}|$. Our conclusions are given in Sec. V.

## II. FORM FACTORS AT SIMULATION POINTS

### A. Gauge ensembles

Our gauge ensembles are generated for $2 + 1$-flavor lattice QCD, where dynamical up and down quarks are degenerate and dynamical strange quark has a distinct mass.

We take three lattice spacings of $a \simeq 0.044$, 0.055, and 0.080 fm [27], which correspond to the lattice cutoffs $a^{-1} \sim 2.453(4)$, 3.610(9), and 4.496(9) GeV, respectively. They are determined using the Yang-Mills gradient flow [28]. The tree-level improved Symanzik gauge action [29–31] and the Möbius domain-wall quark action [22] are used to control discretization errors and to preserve chiral symmetry to a good precision. We refer the interested reader to Refs. [22,27,32] for the five-dimensional formulation of the quark action and our implementation. Its four-dimensional effective Dirac operator is given by

$$\frac{1+m_q}{2} + \frac{1-m_q}{2}\gamma_5 \epsilon[H_M], \tag{6}$$

where $m_q$ represents the quark mass. We employ the kernel operator $H_M = 2\gamma_5 D_W(-M)/\{2 + D_W(-M)\}$, where $D_W(-M)$ is the Wilson-Dirac operator with a negative mass with $M = 1$ and with stout smearing [33] applied three successive times to the gauge links. With our choice of the staple weight ($\rho = 0.1$), the smearing radius $\sim 1.5a$ is at the scale of the lattice spacing and vanishes in the limit of $a \to 0$. We therefore expect that the link smearing does not change the continuum limit of the $B \to D^*$ form factors. It may induce additional discretization effects, which, however, do not turn out to be large in our continuum and chiral extrapolation in Sec. III. The polar decomposition approximation in Eq. (6) is realized for the sign function $\epsilon[H_M]$ in the five-dimensional domain-wall implementation. With this choice, the residual quark mass, which is a measure of the chiral symmetry violation, is suppressed to the level of

1 MeV at $a \simeq 0.080$ fm, and 0.2 MeV or less on finer lattices, which are significantly smaller than the physical up and down quark masses [27]. With reasonably small values of the fifth-dimensional size $N_5 = 8$–12, the computational cost is largely reduced from that of our previous large-scale simulation [34,35] using the overlap formulation that also preserves chiral symmetry [36].

We employ the Möbius domain-wall action for all relevant quark flavors. Our choices of the degenerate up and down quark mass $m_{ud}$ correspond to pion masses from $M_\pi \sim 500$ MeV down to 230 MeV. Chiral symmetry allows us to use heavy meson chiral perturbation theory (HMChPT) [37,38] for our chiral extrapolation of simulation results to the physical pion mass without introducing terms to describe discretization effects. We take a strange quark mass $m_s$ close to its physical value fixed from $M_{\eta_s}^2 = 2M_K^2 - M_\pi^2$. The charm quark mass $m_c$ is set to its physical value fixed from the spin-averaged charmonium mass $(M_{\eta_c} + 3M_{J/\Psi})/4$ [39]. Depending on the lattice spacing $a$, we take three to six bottom quark masses $m_Q = 1.25^n m_c$ ($n = 0, 1, \ldots$) satisfying the condition $m_Q \leq 0.7a^{-1}$ to avoid large discretization errors. We note that, with chiral symmetry, the leading discretization error is reduced to $O((am_Q)^2)$. As we will see in the next section, $a$ and $m_Q$ dependences of the form factors are mild, and the extrapolation to the continuum limit and physical quark masses is under good control. Our simulation parameters are summarized in Tables I and II.

The spatial lattice size $L = N_s a$ is chosen to satisfy the condition $M_\pi L \gtrsim 4$ to suppress finite volume effects

TABLE I. Parameters of gauge ensembles. We denote the five-dimensional lattice size as $N_s^3 \times N_t \times N_5$. Quark masses are bare value in lattice units.

| Lattice parameters | $m_{ud}$ | $m_s$ | $M_\pi$[MeV] | $M_K$[MeV] | $\Delta t + \Delta t'$ | $N_{conf}$ | $N_{t_{src}}$ |
|---|---|---|---|---|---|---|---|
| $\beta = 4.17$, $a^{-1} = 2.453(44)$ GeV, $32^3 \times 64 \times 12$ | 0.0190 | 0.0400 | 499(1) | 618(1) | 12, 16, 24, 28 | 100 | 1 |
| | 0.0120 | 0.0400 | 399(1) | 577(1) | 12, 16, 22, 26 | 100 | 1 |
| | 0.0070 | 0.0400 | 309(1) | 547(1) | 12, 16, 22, 26 | 100 | 2 |
| | 0.0035 | 0.0400 | 230(1) | 527(1) | 8, 12, 16, 20 | 100 | 2 |
| $\beta = 4.17$, $48^3 \times 96 \times 12$ | 0.0035 | 0.0400 | 226(1) | 525(1) | 8, 12, 16, 20 | 100 | 2 |
| $\beta = 4.35$, $a^{-1} = 3.610(65)$ GeV, $48^3 \times 96 \times 8$ | 0.0120 | 0.0250 | 501(2) | 620(2) | 18, 24, 36, 42 | 50 | 1 |
| | 0.0080 | 0.0250 | 408(2) | 582(2) | 18, 24, 33, 39 | 50 | 1 |
| | 0.0042 | 0.0250 | 300(1) | 547(2) | 18, 24, 33, 39 | 50 | 2 |
| $\beta = 4.47$, $a^{-1} = 4.496(80)$ GeV, $64^3 \times 128 \times 8$ | 0.0030 | 0.0150 | 284(1) | 486(1) | 16, 24, 32, 40 | 50 | 2 |

TABLE II. Bare heavy quark masses in lattice units used to calculate relevant two- and three-point functions. The smallest value corresponds to the physical charm mass fixed from the spin averaged charmonium mass.

| $\beta$ | $m_Q$ |
|---|---|
| 4.17 | 0.44037, 0.55046, 0.68808 |
| 4.35 | 0.27287, 0.34109, 0.42636, 0.53295, 0.66619 |
| 4.47 | 0.210476, 0.263095, 0.328869, 0.4110859, 0.5138574, 0.6423218 |

(FVEs). At our smallest $M_\pi$ on the coarsest lattice, we also simulate a smaller physical lattice size $M_\pi L \sim 3$ to directly examine the FVEs.

The statistics at the lattice cutoffs $a^{-1} \simeq 2.5$, 3.6, and 4.5 GeV, are 5,000, 2,500, and 2,500 hybrid Monte Carlo trajectories with a unit trajectory length of 1, 2, and 4 times the molecular dynamics time unit, respectively, which is increased to take account of the longer autocorrelation toward the continuum limit. On each ensemble, we take $N_{\mathrm{conf}}$ configurations in an equal trajectory interval to calculate correlation functions relevant to the $B \to D^* \ell \nu$ decay. Our choice of $N_{\mathrm{conf}}$ is listed in Table I. More details on our simulation setup can be found in Ref. [27].

## B. Ratio method

The $B \to D^*$ matrix elements can be extracted from the three- and two-point functions

$$C^{BD^*}_{\mathcal{O}_\Gamma}(\Delta t, \Delta t'; \mathbf{p}, \mathbf{p}', \epsilon') = \frac{1}{N_s^3 N_{t_{\mathrm{src}}}} \sum_{t_{\mathrm{src}}, \mathbf{x}_{\mathrm{src}}} \sum_{\mathbf{x}, \mathbf{x}'} \langle \mathcal{O}_{D^*}(\mathbf{x}', t_{\mathrm{src}} + \Delta t + \Delta t'; \epsilon') \mathcal{O}_\Gamma(\mathbf{x}, t_{\mathrm{src}} + \Delta t) \mathcal{O}_B(\mathbf{x}_{\mathrm{src}}, t_{\mathrm{src}})^\dagger \rangle e^{-i\mathbf{p}(\mathbf{x}-\mathbf{x}_{\mathrm{src}}) - i\mathbf{p}'(\mathbf{x}'-\mathbf{x})}, \quad (7)$$

$$C^P(\Delta t; \mathbf{p}) = \frac{1}{N_s^3 N_{t_{\mathrm{src}}}} \sum_{t_{\mathrm{src}}, \mathbf{x}_{\mathrm{src}}} \sum_{\mathbf{x}} \langle \mathcal{O}_P(\mathbf{x}, t_{\mathrm{src}} + \Delta t) \mathcal{O}_P(\mathbf{x}_{\mathrm{src}}, t_{\mathrm{src}})^\dagger \rangle e^{-i\mathbf{p}(\mathbf{x}-\mathbf{x}_{\mathrm{src}})} \quad (8)$$

through their ground state contribution

$$C^{BD^*}_{\mathcal{O}_\Gamma}(\Delta t, \Delta t'; \mathbf{p}, \mathbf{p}', \epsilon') \xrightarrow[\Delta t, \Delta t' \to \infty]{} \frac{Z^*_{D^*}(\mathbf{p}', \epsilon') Z_B(\mathbf{p})}{4 E_{D^*}(\mathbf{p}') E_B(\mathbf{p})} \langle D^*(p', \epsilon') | \mathcal{O}_\Gamma | B(p) \rangle e^{-E_{D^*}(\mathbf{p}')\Delta t' - E_B(\mathbf{p})\Delta t},$$

$$C^P(\Delta t; \mathbf{p}) \xrightarrow[\Delta t \to \infty]{} \frac{Z^*_P(\mathbf{p}) Z_P(\mathbf{p})}{2 E_P(\mathbf{p})} e^{-E_P(\mathbf{p})\Delta t}, \quad (9)$$

where $\mathcal{O}_\Gamma$ is the weak vector ($V_\mu$) or axial ($A_\mu$) current, and we omit the symbol $'$ on the momentum variable and the argument $\epsilon'$ for the $D^*$ two-point function $C^{D^*}(\Delta t; \mathbf{p}', \epsilon')$ for simplicity. The meson interpolating field is denoted by $\mathcal{O}_P$ ($P = B, D^*$), where Gaussian smearing is applied to enhance their overlap with the ground state $Z_P(\mathbf{p}) = \langle P | \mathcal{O}_P^\dagger \rangle$. The $B$ meson is at rest ($\mathbf{p} = \mathbf{0}$) throughout this paper. The $w$ dependence of the form factors is studied by varying the three-momentum of $D^*$ as $|\mathbf{p}'|^2 = 0$, 1, 2, 3, 4 (in this paper, we denote the momentum on the lattice in units of $2\pi/L$). To this end, we also calculate $D^*$ two-point functions with the local sink operator $\mathcal{O}^\dagger_{D^*, \mathrm{lcl}}$:

$$C^{D^*}_{\mathrm{sl}}(\Delta t; \mathbf{p}', \epsilon') = \frac{1}{N_s^3 N_{t_{\mathrm{src}}}} \sum_{t_{\mathrm{src}}, \mathbf{x}_{\mathrm{src}}} \sum_{\mathbf{x}} \langle \mathcal{O}_{D^*, \mathrm{lcl}}(\mathbf{x}, t_{\mathrm{src}} + \Delta t; \epsilon')$$
$$\times \mathcal{O}_{D^*}(\mathbf{x}_{\mathrm{src}}, t_{\mathrm{src}}; \epsilon')^\dagger \rangle e^{-i\mathbf{p}'(\mathbf{x}-\mathbf{x}_{\mathrm{src}})}, \quad (10)$$

which are used in a correlator ratio (13) below.

The three-point functions are calculated by using the sequential source method, where the total temporal separation $\Delta t + \Delta t'$ between the source and sink operators is fixed and the temporal location of the weak current is varied. In order to control the contamination from the excited states, we repeat our measurement for four values of the source-sink separation $\Delta t + \Delta t'$ listed in Table I.

The statistical accuracy of the two- and three-point functions are improved by averaging over the spatial location of the source operator $\mathbf{x}_{\mathrm{src}}$ as indicated in Eqs. (7) and (8). To this end, we employ a volume source with $Z_2$ noise. At $M_\pi \lesssim 300$ MeV, we repeat our measurement over two values of the source time slices $t_{\mathrm{src}} = 0$ and $N_t/2$ to take the average of the correlators. Namely, $N_{t_{\mathrm{src}}}$ in Eqs. (7) and (8) is 2 at $M_\pi \lesssim 300$ MeV, and 1 otherwise. The correlators with nonzero momentum $\mathbf{p}$ are also averaged over all possible $\mathbf{p}$'s based on parity and rotational symmetries on the lattice.

The number of measurements on each ensemble is given as $N_{\mathrm{conf}} N_{t_{\mathrm{src}}}$ in Table I. As mentioned above, correlation functions for each configuration are averaged over $N_{t_{\mathrm{src}}}$ values of the source time slice $t_{\mathrm{src}}$. Then, simulation data of $N_{\mathrm{conf}}$ configurations are divided into 50 bins: namely, the bin size is two (one) configurations for $\beta = 4.17$ (4.35 and 4.47). The statistical error is estimated by the bootstrap method with 5,000 replicas. Figure 1 shows the bin size dependence of the relative statistical error of a three-point function at our largest cutoff, where the topological charge $Q$ changes much less frequently leading to larger autocorrelation compared to the coarser lattices [27]. Our choice of the bin size on this finest lattice is one configuration, and we do not observe significant increase of the statistical error toward larger bin sizes. This suggests that our bin size is sufficiently large partly due to our choice of larger unit trajectory length towards the continuum limit.
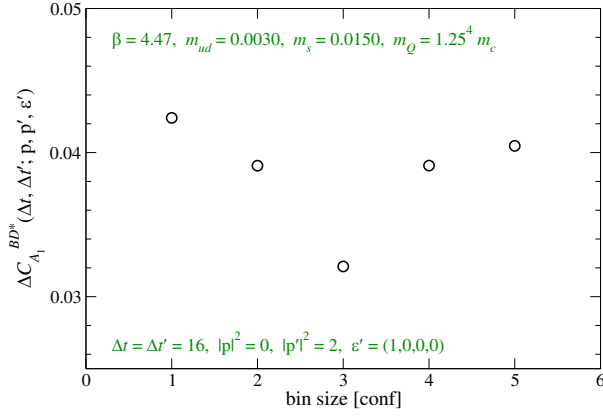
FIG. 1. Bin size dependence of relative statistical error of three-point function $C_{\mathcal{O}_\Gamma}^{BD^*}(\Delta t, \Delta t'; \mathbf{p}, \mathbf{p}', \epsilon')$. We plot data at the largest cutoff $a^{-1} \simeq 4.5$ GeV, $m_Q = 1.25^4 m_c$, $\Delta t = \Delta t' = 16$, $|\mathbf{p}|^2 = 0$, $|\mathbf{p}'|^2 = 2$, and $\epsilon' = (1, 0, 0, 0)$.

We also note that the error estimate is stable when we vary the number of bootstrap replicas as well as when we employ the jackknife method instead.

As observed in our previous simulation in the trivial topological sector [34], the local topological fluctuation are active even if we fix the "global" topological charge $Q$, namely the whole-volume integral of the topological charge density. In Ref. [40], we demonstrated that the topological susceptibility $\chi_t$ is calculable within the fixed topology setup. Our result for $\chi_t$ shows quark mass dependence consistent with ChPT, and its value extrapolated to the chiral limit is in good agreement with that from our simulation in the $\epsilon$ regime [41,42]. References [43,44] argued that the bias due to the freezing of the global topology can be considered as FVEs suppressed by the inverse space-time volume. In our study of the $B \to \pi \ell \nu$ decay [27], we confirmed that such FVEs on the pion effective mass are well suppressed, so that its $|Q|$ dependence is not significant. Figure 2 shows the statistical average of a $B \to D^*$ three-point function calculated for each $|Q|$. We do not observe any significant $|Q|$ dependence of the average at $|Q| \leq 4$. At larger $|Q|$, we do not have enough data for a reliable error estimate, but individual data plotted by crosses are consistent with the averages at $|Q| \leq 4$. We therefore conclude that the topology freezing effect is insignificant within our statistical accuracy.

To precisely extract the form factors, we construct ratios of the correlation functions, in which unnecessary overlap factors and exponential damping factors cancel for the ground-state contribution [23]. The statistical fluctuation is also expected to partly cancel in such a ratio. With $B$ and $D^*$ mesons at rest, the vector current matrix element (4) vanishes, and the axial vector one (5) is sensitive only to the axial vector form factor $h_{A_1}$ at zero recoil, which is the fundamental input to determine $|V_{cb}|$. In order to precisely determine $h_{A_1}(1)$, we employ a double ratio,
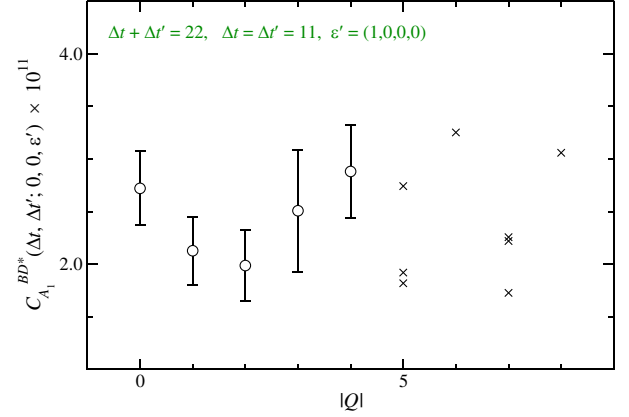


FIG. 2. Topological charge dependence of three-point function $C_{\mathcal{O}_\Gamma}^{BD^*}(\Delta t, \Delta t'; \mathbf{p}, \mathbf{p}', \epsilon')$ at $a^{-1} \simeq 2.5$ GeV and $M_\pi \simeq 300$ MeV. We plot data with $m_Q = 1.25^2 m_c$, $\Delta t = \Delta t' = 11$, zero momenta $|\mathbf{p}|^2 = |\mathbf{p}'|^2 = 0$, and $\epsilon' = (1, 0, 0, 0)$. The error is estimated by the jackknife method. At $|Q| \geq 5$, the number of data are smaller than 10, and it is difficult to reliably estimate the statistical error. We therefore plot individual data by crosses rather than their statistical average.

$$R_1(\Delta t, \Delta t') = \frac{C_{A_1}^{BD^*}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{0}, \epsilon') C_{A_1}^{D^*B}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{0}, \epsilon')}{C_{V_4}^{BB}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{0}) C_{V_4}^{D^*D^*}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{0}, \epsilon', \epsilon')}$$
$$\xrightarrow[\Delta t, \Delta t' \to \infty]{} |h_{A_1}(1)|^2, \qquad (11)$$

where the $D^*$ polarization is chosen as $\epsilon' = (1, 0, 0, 0)$ along the polarization of the current inserted, i.e. $A_1$. The left panel of Fig. 3 shows our results for this ratio at our smallest $M_\pi$ and $a^{-1}$ with $m_Q = 1.25 m_c$. Note that the ratio is symmetrized as $R_1(\Delta t, \Delta t') \to \{R_1(\Delta t, \Delta t') + R_1(\Delta t', \Delta t)\}/2$, since we use the same smearing function for both the source and sink operators. We observe a reasonable consistency among data at intermediate values of the source-sink separation $\Delta t + \Delta t' = 12$ (circles) and 16 (squares). The excited state contribution could potentially be significant but not large ($\lesssim 1\%$) for the smallest $\Delta t + \Delta t' = 8$ (diamonds). Data at the largest $\Delta t + \Delta t' = 20$ (triangles) show a long plateau and are consistent with those at smaller $\Delta t + \Delta t'$ within the large statistical errors. These observations suggest that the data at $\Delta t + \Delta t' = 12$, which is roughly 1 fm, are dominated by the ground state contribution at least at the midpoint $\Delta t \sim \Delta t'$. The situation is similar in the right panel of the same figure for our largest $a^{-1}$ and with $m_Q = 1.25^4 m_c$, where data around midpoint $\Delta t \sim \Delta t'$ are consistent among all simulated source-sink separations within $2\sigma$.

We find that all data are well described by the following fitting form including effects from the first excited states:

$$R_1(\Delta t, \Delta t') = |h_{A_1}(1)|^2 \big(1 + a e^{-\Delta M_B \Delta t} + b e^{-\Delta M_{D^*} \Delta t}$$
$$+ a e^{-\Delta M_B \Delta t'} + b e^{-\Delta M_{D^*} \Delta t'}\big) \qquad (12)$$

with $\chi^2/\text{d.o.f.} \lesssim 1$ for a wide range of the values of $\Delta t$ and $\Delta t'$. Here $\Delta M_{B(D^*)}$ represents the energy difference between the $B(D^*)$ meson ground state and the first excited state of the same quantum numbers, and is set to the value estimated from a two-exponential fit to the two-point function $C^{B(D^*)}(\Delta t; \mathbf{0})$. The fit range is chosen such that all the data of $R_1(\Delta t, \Delta t')$ with the source-to-current (current-to-sink) separation $\Delta t^{(\prime)}$ equal to or larger than a lower cut $\Delta t_{\min}^{(\prime)}$ are included irrespective of the source-sink separation $\Delta t + \Delta t'$. As shown in the left panel of Fig. 4, our result for $h_{A_1}(1)$ is stable against the choice of the lower cuts $\Delta t_{\min}$ and $\Delta t'_{\min}$.

The right panels of Figs. 3 and 4 for our largest $a^{-1}$ also show ground state dominance with $\Delta t + \Delta t' \gtrsim 23a \sim 1$ fm

and stability of $h_{A_1}(1)$ against the choice of the fit range, respectively. The situation is similar for other simulation parameters.

In order to extract form factors at nonzero recoils, the $D^*$ momentum and polarization vector need to be chosen appropriately. The axial vector matrix element (5) is sensitive only to $h_{A_1}$ with a polarization vector $\boldsymbol{\epsilon}' = (1, 0, 0)$ and a momentum $\mathbf{p}'_\perp$ that satisfies $\epsilon'^* v = 0$. We note that, here and in the following, three-dimensional polarization vector $\boldsymbol{\epsilon}$ is accompanied by its temporal component to be fixed from the convention $\epsilon'^* p'_\perp = 0$, which is assumed for Eqs. (4) and (5). Then, the recoil parameter dependence of $h_{A_1}$ may be studied from the following ratio:
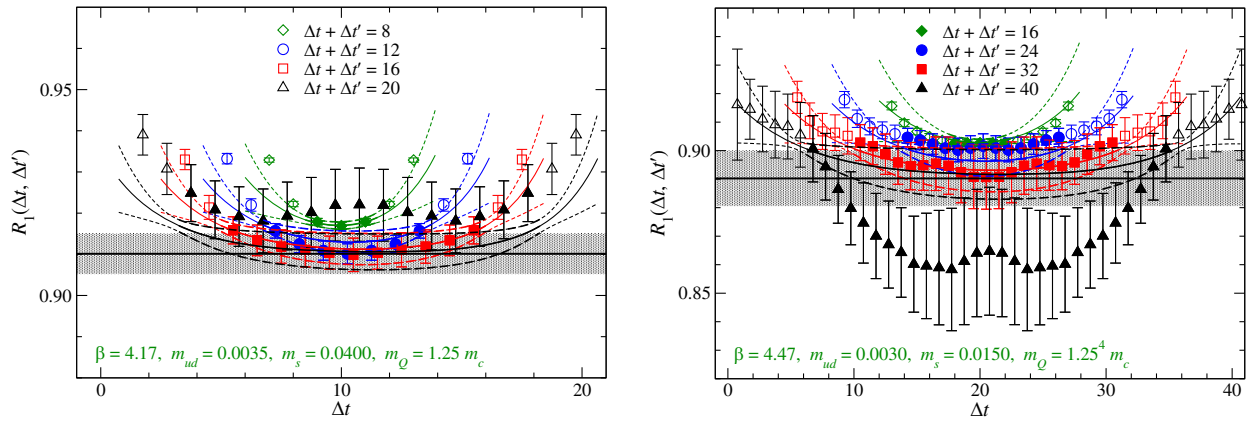


FIG. 3.    Double ratio (11) to estimate $h_{A_1}(1)$ as a function of temporal location of weak currents $\Delta t$. The open symbols show all data at different choices of the source-sink separations $\Delta t + \Delta t'$, whereas the filled symbols are data used in the fit (12) to extract the ground state matrix element shown by the horizontal black band. The thick (thin) curves are fit curves inside (outside) the fit range. The left panel shows the data at our smallest pion mass $M_\pi \simeq 230$ MeV ($a^{-1} \simeq 2.5$ GeV, $m_Q = 1.25 m_c$), whereas the right panel is for our largest cutoff $a^{-1} \simeq 4.5$ GeV ($M_\pi \simeq 300$ GeV, $m_Q = 1.25^4 m_c$). All data are shifted along the horizontal axis so that the mid-point $\Delta t = \Delta t'$ is located at the center of the panel ($\Delta t = 10$ and 20 in the left and right panels, respectively). The symbols are further but slightly shifted along the horizontal axis for clarity.
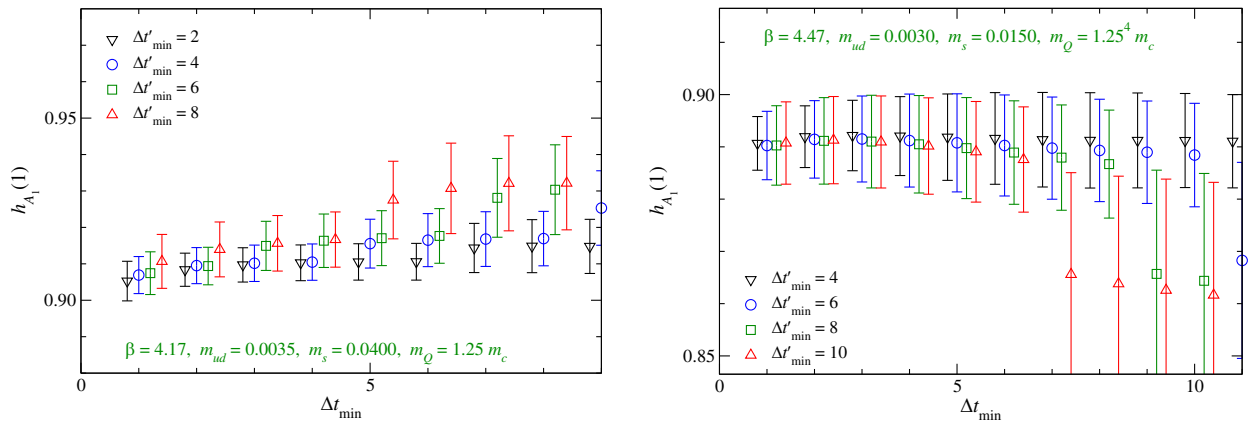


FIG. 4.    Axial form factor $h_{A_1}(1)$ from fit (12) as a function of lower cut $\Delta t_{\min}$. Left and right panels show data at $(a^{-1}, M_\pi) \simeq$ (2.5 GeV, 230 MeV) and (4.5 GeV, 300 MeV), respectively. Different symbols show data with different values for $\Delta t'_{\min}$. Symbols are slightly shifted along the horizontal axis for clarity.

$$R'_1(\mathbf{p}'_\perp; \Delta t, \Delta t') = \frac{C^{BD^*}_{A_1}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}'_\perp, \epsilon') C^{D^*}_{sl}(\Delta t'; \mathbf{0}, \epsilon')}{C^{BD^*}_{A_1}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{0}, \epsilon') C^{D^*}_{sl}(\Delta t', \mathbf{p}'_\perp, \epsilon')}$$

$$\xrightarrow[\Delta t, \Delta t' \to \infty]{} \frac{w+1}{2} \frac{h_{A_1}(w)}{h_{A_1}(1)}. \quad (13)$$

Here we use the two-point function with the local sink $C^{D^*}_{sl}$ (10) as in our study of the $K \to \pi \ell \nu_\ell$ decay [45]. The vector form factor $h_V$ can be extracted from the following ratio through the vector current matrix element (4):

$$R_V(\mathbf{p}''_\perp, \mathbf{p}'_\perp; \Delta t, \Delta t') = \frac{C^{BD^*}_{V_1}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}''_\perp, \epsilon'')}{C^{BD^*}_{A_1}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}'_\perp, \epsilon')}$$

$$\xrightarrow[\Delta t, \Delta t' \to \infty]{} \frac{i\varepsilon_{1ij} \epsilon''^*_i v''_{\perp j}}{w+1} \frac{h_V(w)}{h_{A_1}(w)}. \quad (14)$$

Here we use two different polarization vectors, $\epsilon' = (1, 0, 0)$ and $\epsilon'' = (0, 1, 0)$, and momenta $\mathbf{p}'_\perp$ and $\mathbf{p}''_\perp$ that satisfy $\mathbf{p}'^{(\prime\prime)}_\perp \perp \epsilon'^{(\prime\prime)}$. Note that $|\mathbf{p}''_\perp| = |\mathbf{p}'_\perp|$ to share the same recoil parameter $w$, and $v''_{\perp j} = p''_{\perp j}/M_{D^*}$. We emphasize that renormalization factors of the weak vector and axial currents cancel thanks to chiral symmetry preserved in our simulations.

Figures 5 and 6 show an example of our results for $R'_1$ and $R_V$ at our largest cutoff $a^{-1} \simeq 4.5$ GeV and $m_Q = 1.25^4 m_c$. Around the midpoint $\Delta t \sim \Delta t'$, we observe good consistency in these ratios among all simulated values of the source-sink separation $\Delta t + \Delta t' \gtrsim 1$ fm. A similar ground state dominance is also observed at other simulation points. We carry out a simultaneous fit using a fitting form that takes account of the first excited state contribution as

$$R'_1(\mathbf{p}'_\perp; \Delta t, \Delta t') = R'_1(\mathbf{p}'_\perp)\left(1 + ae^{-\Delta M_B \Delta t} + be^{-\Delta M_{D^*} \Delta t'}\right) \quad (15)$$

to extract the form factor ratio $h_{A_1}(w)/h_{A_1}(1)$ from the ground state contribution $R'_1(\mathbf{p}'_\perp)$, and a similar form for $R_V(\mathbf{p}''_\perp, \mathbf{p}'_\perp; \Delta t, \Delta t')$ to extract $h_V(w)/h_{A_1}(w)$. Our data are well described by this fitting form with $\chi^2/\text{d.o.f.} \lesssim 1$. From these form factor ratios and $h_{A_1}(1)$ from $R_1$, we calculate $h_{A_1}(w)$ and $h_V(w)$ at simulated values of $w$.

There is a $2\sigma$ bump of $R'_1$ for $\Delta t + \Delta t' = 40$ (black triangles) around $\Delta t = 28$ ($\Delta t' = 12$) in Fig. 5. Since the global topological charge does not fluctuate very much at the corresponding cutoff, one may expect bump(s) in the exponential decay of relevant correlators due to instantons frozen at the temporal separation $\Delta t^{(\prime)}$ for a certain period of our simulation time. However, Fig. 7 shows that there is no statistically significant bump of the relevant correlators at $\Delta t \sim 28$ and $\Delta t' \sim 12$. As mentioned above, the local topological fluctuation is active even when the global topological charge is fixed. We attribute $1 - 2\sigma$ bumps of correlator ratios in Figs. 3, 5–6, 8, and 9 to large statistical fluctuations of correlators at the largest temporal separation $\Delta t + \Delta t' = 40$, and accidental anticorrelation of the statistical fluctuations. Since these data have larger statistical error than those at smaller $\Delta t + \Delta t'$, these bumps do not change results of the fits (12) and (15) to extract the ground state contribution.

The axial vector form factors $h_{A_2}$ and $h_{A_3}$ can be extracted from the axial vector matrix element (5) with the $D^*$ spatial momentum $\mathbf{p}'_{\not\perp}$ not perpendicular to the spatial polarization vector. The matrix element of $A_1$, for instance, has nonzero sensitivity to $h_{A_1}$ and $h_{A_3}$. We use the following ratio to extract $h_{A_3}$:

$$R_3(\mathbf{p}'_{\not\perp}, \mathbf{p}'_\perp; \Delta t, \Delta t') = r_Z \frac{C^{BD^*}_{A_1}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}'_{\not\perp}, \epsilon'')}{C^{BD^*}_{A_1}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}'_\perp, \epsilon')}$$

$$\xrightarrow[\Delta t, \Delta t' \to \infty]{} \epsilon''^*_1 - \frac{\epsilon''^*_4 v'_{\not\perp, 1}}{w+1} \frac{h_{A_3}(w)}{h_{A_1}(w)}, \quad (16)$$
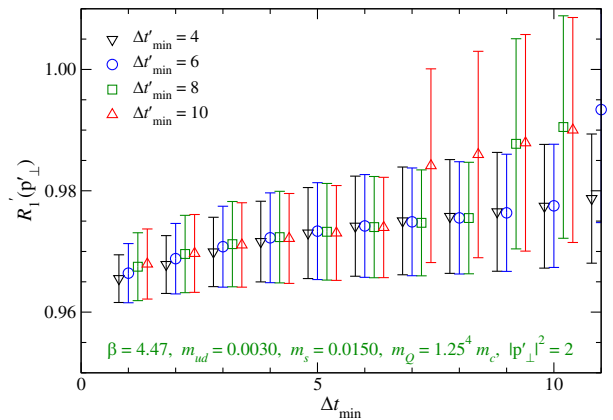


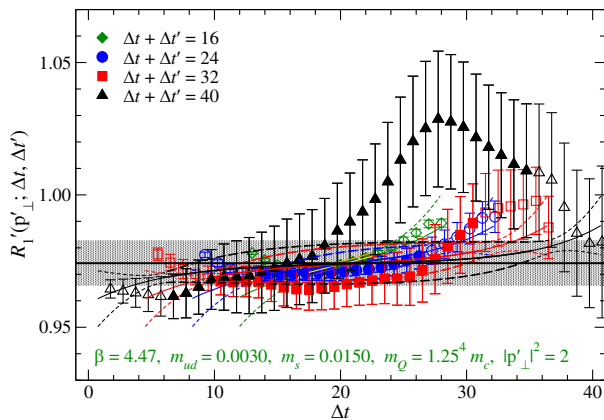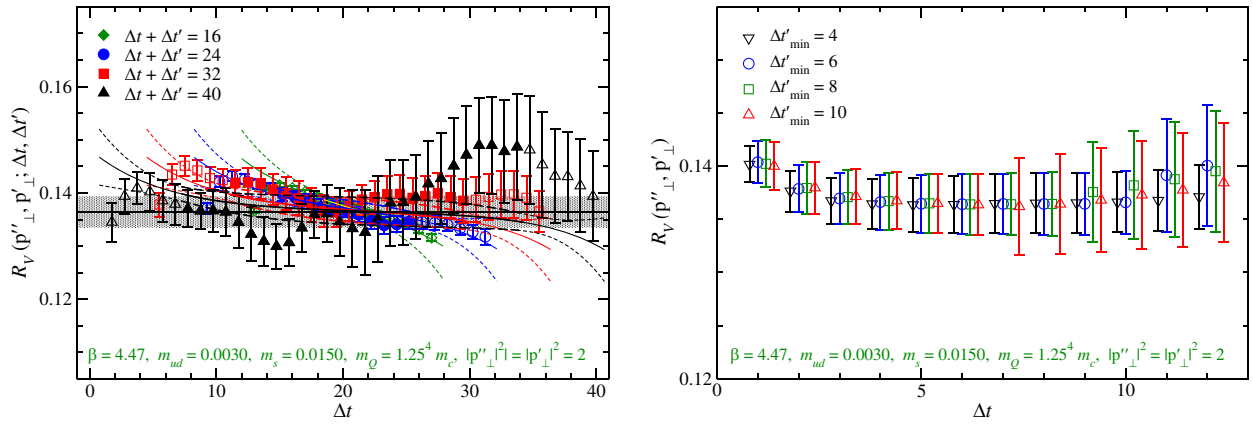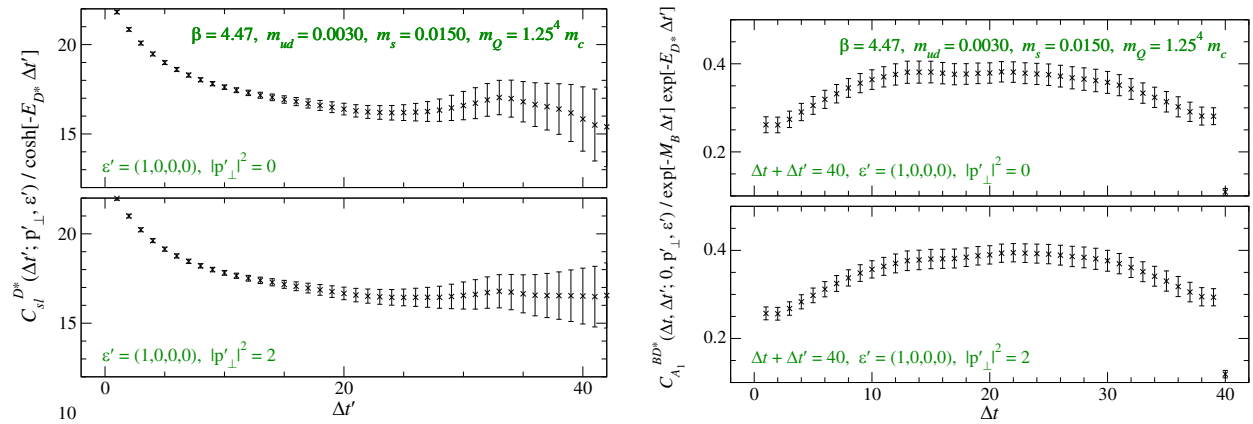FIG. 5. Left: ratio $R'_1(\mathbf{p}'_\perp; \Delta t, \Delta t')$ (13) to study $w$ dependence of $h_{A_1}$. Right: fitted value of $R'_1(\mathbf{p}'_\perp; \Delta t, \Delta t')$ with different choices of the lower cuts $\Delta t_{\min}$ and $\Delta t'_{\min}$. Both panels show data at $a^{-1} \simeq 4.5$ GeV and $M_\pi \simeq 300$ MeV with the $D^*$ momentum $|\mathbf{p}_\perp| = \sqrt{2}$.

FIG. 6. Same as Fig. 5 but for ratio $R_V(\mathbf{p}''_\perp, \mathbf{p}'_\perp; \Delta t, \Delta t')$ (14) to study $h_V(w)$.



FIG. 7. Effective plot of two- and three-point functions, $C_{sl}^{D^*}(\Delta t', \mathbf{p}'_\perp, \epsilon')$ (left panels) and $C_{A_1}^{BD^*}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}'_\perp, \epsilon')$ (right panels). Data divided by the exponential damping factor(s) for the ground state(s) are plotted as a function of temporal separations $\Delta t'$ and $\Delta t$, respectively. These are used to calculate $R'_1(\mathbf{p}'_\perp; \Delta t, \Delta t')$ for $\Delta t + \Delta t' = 40$ shown in Fig. 5. The top and bottom panels show data with the $D^*$ momentum $|\mathbf{p}'_\perp|^2 = 0$ and 2, respectively.
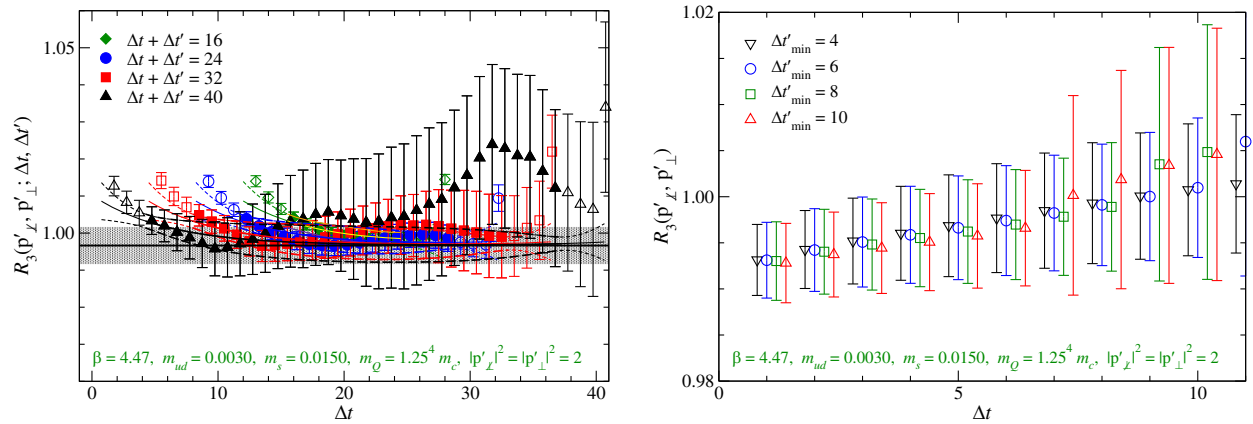


FIG. 8. Same as Fig. 5 but for ratio $R_3(\mathbf{p}_{\measuredangle}, \mathbf{p}_\perp; \Delta t, \Delta t')$ (16) to study $h_{A_3}(w)$.
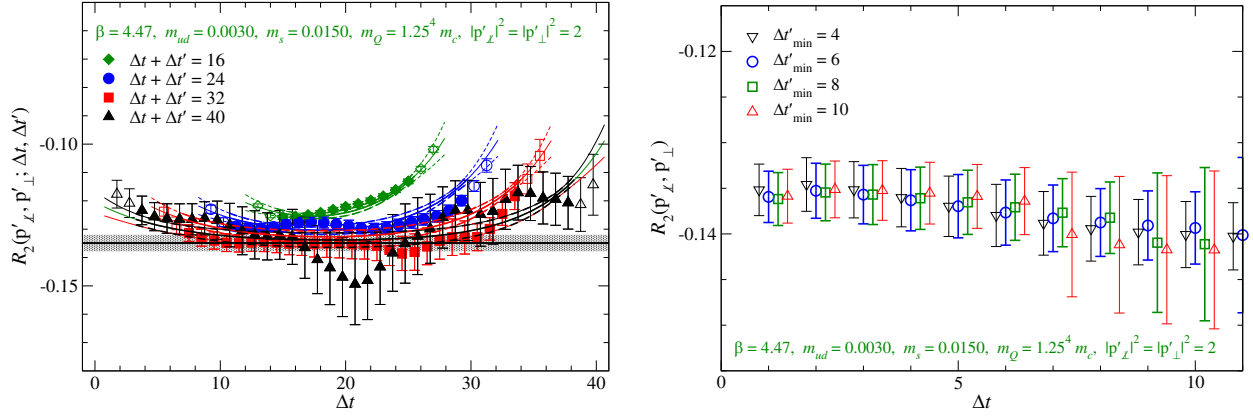
FIG. 9.    Same as Fig. 5 but for ratio $R_2(\mathbf{p}_{\not\perp}, \mathbf{p}_\perp; \Delta t, \Delta t')$ (16) to study $h_{A_2}(w)$.

where $\quad \epsilon' = (1,0,0), \quad \epsilon'^* \mathbf{p}'_\perp = 0, \quad \epsilon''^* \mathbf{p}'_{\not\perp} \neq 0, \quad$ and $|\mathbf{p}'_{\not\perp}| = |\mathbf{p}'_\perp|$. The factor $r_Z = Z_{D^*}(\mathbf{p}'_\perp, \epsilon')/Z_{D^*}(\mathbf{p}'_{\not\perp}, \epsilon'')$ appears, since the overlap factor of $D^*$ depends on whether the polarization vector is perpendicular to the momentum, even if $|\mathbf{p}'_{\not\perp}| = |\mathbf{p}'_\perp|$. While $r_Z$ can be estimated from individual fit to the relevant two-point functions, we employ

a ratio $r_Z = \sqrt{C^{D^*}(\Delta t_{\mathrm{ref}}, \mathbf{p}'_\perp, \epsilon')/C^{D^*}(\Delta t_{\mathrm{ref}}, \mathbf{p}'_{\not\perp}, \epsilon'')}$, which shows better stability against the temporal separation. The reference temporal separation is chosen as $\Delta t_{\mathrm{ref}} \sim T/4$ by inspecting the statistical accuracy and the ground state saturation of $r_Z$. A similar ratio but with $A_4$ in the numerator is also sensitive to $h_{A_2}$:

$$R_2(\mathbf{p}'_{\not\perp}, \mathbf{p}'_\perp; \Delta t, \Delta t') = r_Z \frac{C_{A_4}^{BD^*}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}'_{\not\perp}, \epsilon'')}{C_{A_1}^{BD^*}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}'_\perp, \epsilon')} \xrightarrow[\Delta t, \Delta t' \to \infty]{} \epsilon''^*_4 \left[ 1 - \frac{1}{w+1} \left\{ \frac{h_{A_2}(w)}{h_{A_1}(w)} + v_{\not\perp,4} \frac{h_{A_3}(w)}{h_{A_1}(w)} \right\} \right]. \quad (17)$$

We plot our results for $R_3$ and $R_2$ on our finest lattice in Figs. 8 and 9, respectively. Similar to other ratios, the excited state contamination is not large. The fit similar to (15) yields the ground state contributions, $R_3(\mathbf{p}'_{\not\perp}, \mathbf{p}'_\perp)$ and $R_2(\mathbf{p}'_{\not\perp}, \mathbf{p}'_\perp)$, which are stable against the choice of the fit range: we do not observe $\Delta t_{\mathrm{min}}^{(\prime)}$ dependence beyond $1\sigma$ level when $\Delta t_{\mathrm{min}} + \Delta t'_{\mathrm{min}} \lesssim 16$, which is our smallest value of the source-sink separation $\Delta t + \Delta t'$. The error rapidly increases at larger values of $\Delta t_{\mathrm{min}}^{(\prime)}$, because less data are available for the fit. From $R_2$, $R_3$, and $h_{A_1}$ extracted above, we can calculate $h_{A_2}$ and $h_{A_3}$.

For one of the $D^*$ momenta $\mathbf{p}' = (1,1,1)$, it is not straightforward to use the above-mentioned correlator ratios, which use $C_{A_1}^{BD^*}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}', \epsilon')$ with $\epsilon'^* \mathbf{p}' = 0$ as a correlator sensitive only to $h_{A_1}(w)$. In this case, we use a linear combination

$$C_{A_1}^{BD^*}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}', \epsilon') - C_{A_1}^{BD^*}(\Delta t, \Delta t'; \mathbf{0}, \mathbf{p}', \epsilon'')$$
$$\propto (w+1)h_{A_1}(w) \quad (18)$$

which only involves $h_{A_1}$ with $\epsilon' = (1,0,0)$ and $\epsilon'' = (0,1,0)$.

The double ratio $R_1$ enables us to calculate $h_{A_1}(1)$ with 1% statistical error or better, as it only involves correlators with zero momentum and the statistical fluctuation partially cancels between the numerator and the denominator. At nonzero recoil, the numerator of $R'_1$ ($R_V$) is exclusively sensitive to $h_{A_1}$ ($h_V$), and the statistical error of $h_{A_1}$ and $h_V$ is typically a few %. This is not the case for $h_{A_2}$ and $h_{A_3}$ with our setup where the $B$ meson is at rest. The numerator of $R_3$ involves $h_{A_1}$ and $h_{A_3}$, and the typical precision for $h_{A_3}$ is 10%–20%. The statistical error increases to $\gtrsim 40\%$ for $h_{A_2}$, since the central value is suppressed by heavy quark symmetry and $R_2$ is a linear combination of all axial form factors.

We note that, thanks to chiral symmetry, finite renormalization factors for the weak vector and axial currents cancel in the above correlator ratios (11), (13), (14), (16), and (17). While discretization effects to the wave function renormalization are not small [46], these also cancel in the ratios.

## III. EXTRAPOLATION TO CONTINUUM LIMIT AND PHYSICAL QUARK MASSES

### A. Choice of fitting form

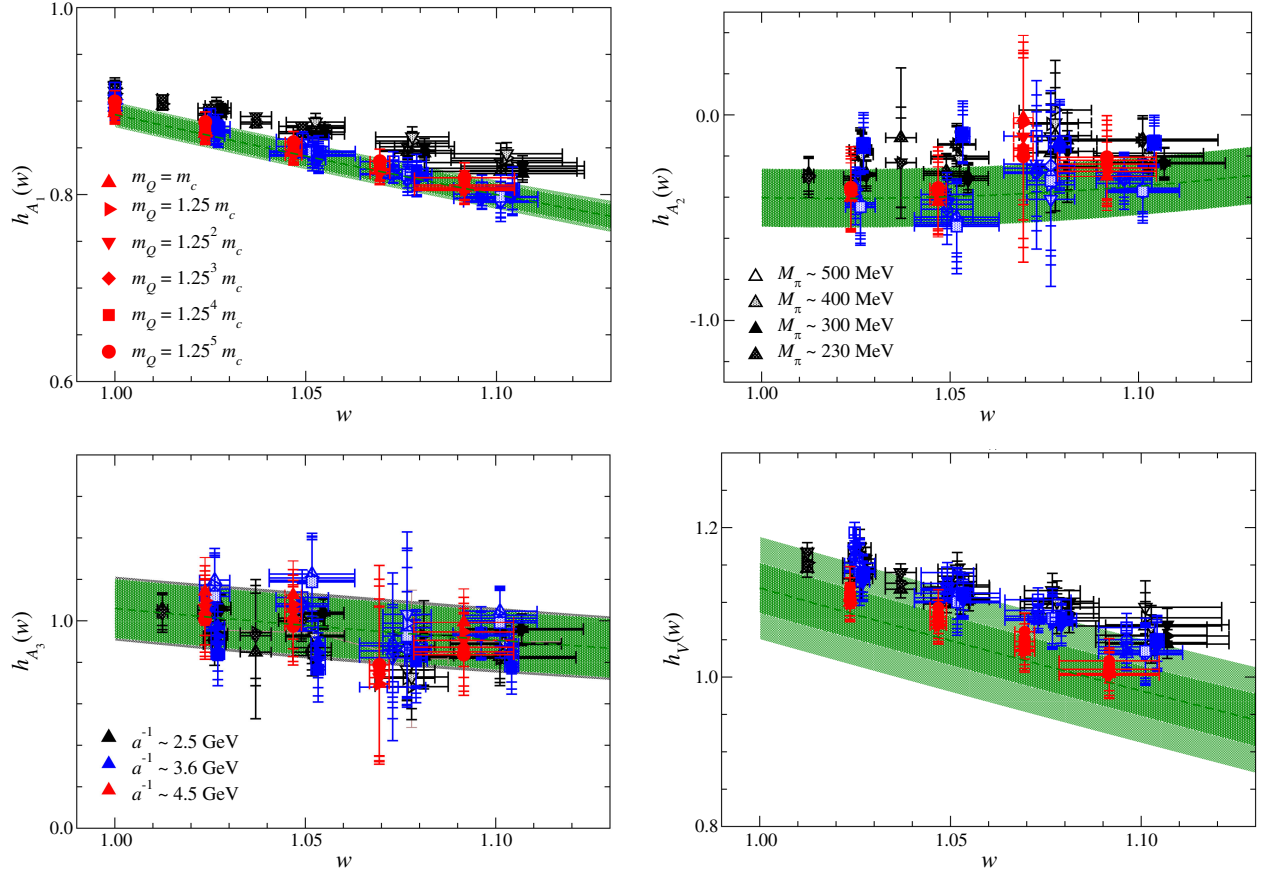Our results for the form factors obtained at various simulation parameters are plotted as a function of the recoil

FIG. 10. Form factors as a function of recoil parameter $w$. The top-left, top-right, bottom-left, and bottom-right panels show results for $h_{A_1}$, $h_{A_2}$, $h_{A_3}$, and $h_V$, respectively. The black, blue, and red symbols show data at $a^{-1} \sim 2.5$, 3.6, and 4.5 GeV, whereas the open, pale-shaded, filled, and dark-shaded ones are at $M_\pi \sim 500$, 400, 300, and 230 MeV, respectively. Symbols with different shapes are obtained with different bottom quark masses. (See the legend in the plots.) The green bands show the form factors extrapolated to the continuum limit and physical quark masses. The dark and pale shaded bands represent their statistical and total uncertainties, respectively.

parameter $w$ in Fig. 10. We find that the results with different heavy and light quark masses are all close to each other. It also turns out that the discretization effects are not large.

We extrapolate these results to the continuum limit and physical quark masses based on next-to-leading order heavy meson chiral perturbation theory (NLO HMChPT) [37,38]. The extrapolation form is generically written as

$$h_X/\eta_X = c + \frac{g_{D^*D\pi}^2}{2\Lambda_\chi^2} \bar{F}_{\log}(\xi_\pi, \Delta_D, \Lambda_\chi) + c_\pi\xi_\pi + c_{\eta_s}\xi_{\eta_s} + c_Q\epsilon_Q$$

$$+ c_a\xi_a + c_{am_Q}\xi_{am_Q} + c_w(w-1) + d_w(w-1)^2 \quad (X = A_1, A_2, A_3 \text{ and } V), \tag{19}$$

which comprises the constant term, chiral logarithm at $w = 1$, and leading corrections in small expansion parameters

$$\xi_\pi = \frac{M_\pi^2}{\Lambda_\chi^2}, \qquad \xi_{\eta_s} = \frac{M_{\eta_s}^2}{\Lambda_\chi^2}, \qquad \epsilon_Q = \frac{\bar{\Lambda}}{2m_Q}, \qquad \xi_a = (\Lambda a)^2, \qquad \xi_{am_Q} = (am_Q)^2, \tag{20}$$

with $M_{\eta_s}^2 = 2M_K^2 - M_\pi^2$, $\Lambda_\chi = 4\pi f_\pi$, and nominal values of $\bar{\Lambda} = \Lambda = 500$ MeV.

We denote the one-loop integral function for the chiral logarithm by $\bar{F}_{\log}(\xi_\pi, \Delta_D, \Lambda_\chi)$, where $\Delta_D = M_{D^*} - M_D$ is the $D^*$-$D$ mass splitting. While the loop function can be approximated as $\bar{F}_{\log}(\xi_\pi, \Delta_D, \Lambda_\chi) = \Delta_D^2 \ln[\xi_\pi] + O(\Delta_D^3)$,

we use its explicit form in Ref. [38]. For the $D^*D\pi$ coupling, we take a value of $g_{D^*D\pi} = 0.53(8)$, which was employed in the Fermilab/MILC paper [6] and covers previous lattice estimates [47–51]. This choice does not lead to a large systematic uncertainty, because the chiral logarithm is suppressed by the small mass splitting squared $\Delta_D^2$.
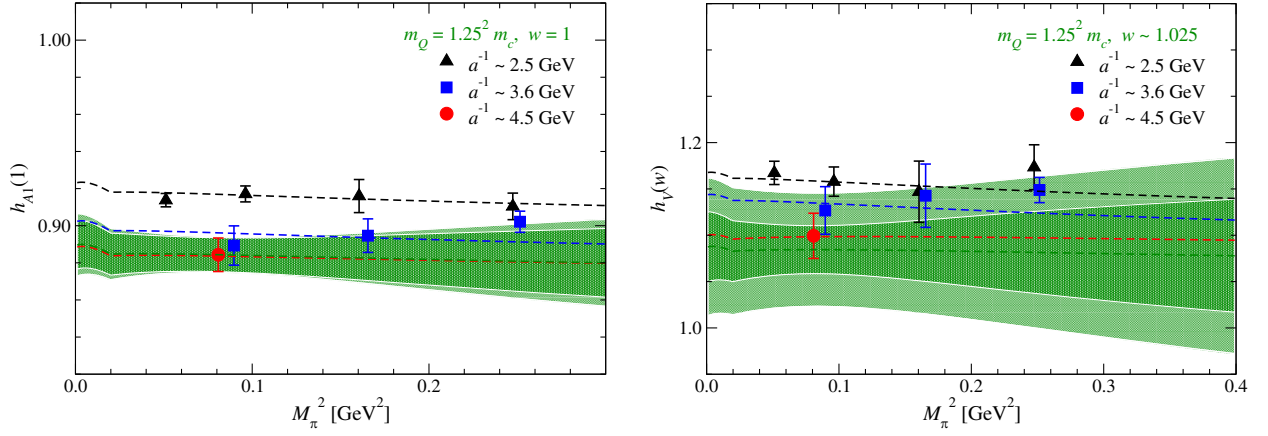
FIG. 11. Axial vector form factor $h_{A_1}$ (left panel) and vector form factor $h_v$ (right panel) as a function of $M_\pi^2$. Both panels show data at $m_Q = 1.25^2 m_c$ and the smallest value of $w$: namely, $w = 1.0$ for $h_{A_1}$ and $\sim 1.205$ for $h_v$, respectively. For the latter, $w$ is not exactly equal to 1.025 due to the slight difference in $L$ among simulated lattice spacings as well as discretization effects to $w = E_D^*/M_D^*$ itself. The black, blue, and red symbols and lines show data and fit curve at $a^{-1} \sim 2.5$, 3.6, and 4.5 GeV, respectively. The green dashed line shows the form factors extrapolated to the continuum limit and physical strange and bottom quark masses. The dark and pale shaded green bands represent the statistical and total error, respectively.

In Fig. 11, we plot $h_{A_1}$ and $h_V$ at our smallest value of $w$ as a function of $M_\pi^2$. The $M_\pi^2$ dependence is rather mild, possibly because there is no valence pion in this decay. We note that there is a nontrivial concave structure at small $M_\pi^2$ due to the opening of the $D^* \to D\pi$ decay. The effect estimated in NLO HMChPT is, however, not large compared to our statistical accuracy.

For our chiral extrapolation, we employ the so-called $\xi$ expansion in $\xi_\pi$ and $\xi_{\eta_s}$: namely, we use the measured meson masses squared, $M_\pi^2$ and $M_{\eta_s}^2$, instead of quark masses $m_{ud}$ and $m_s$, respectively, as well as quark-mass dependent $f_\pi$ for $\Lambda_\chi$ instead of the decay constant $f$ in the chiral limit. This is motivated by our observation [52] that the $\xi$ expansion shows better convergence of the chiral expansion for light meson observables compared to the $x$ expansion with $x = 2Bm_q/(4\pi f)^2$, where $B$ gives the lowest order relation $M_\pi^2 = 2Bm_{ud}$.

The extrapolation form explicitly includes the one-loop radiative correction $\eta_X$ for the matching of the heavy-heavy currents between QCD and HQET [53–55]. Remaining heavy quark mass dependence of the form factors is then parametrized by polynomials of $\epsilon_Q \propto 1/m_Q$ and $\xi_{am_Q} \propto m_Q$. The dependence is mild as shown in Fig. 12. Our results are well described by the linear terms $c_Q \epsilon_Q$ and $c_{am_Q} \xi_{am_Q}$ in Eq. (19). We note that, for $h_{A_1}$, the term $c_Q \epsilon_Q$ is modified as $c_Q(w-1)\epsilon_Q$ to be consistent with Luke's theorem [56], and we add the quadratic term $d_Q \epsilon_Q^2$ to take account of possible $m_Q$ dependence at $w = 1$.

The discretization effects in the form factors also turn out to be mild with our choice of parameters, $a^{-1} \gtrsim 2.5$ GeV and $m_Q \leq 0.7a^{-1}$. These are described by an expansion, i.e. linear in the $m_Q$ independent and dependent parameters $\xi_a$ and $\xi_{am_Q}$, respectively. Figure 12 shows that the $m_Q$
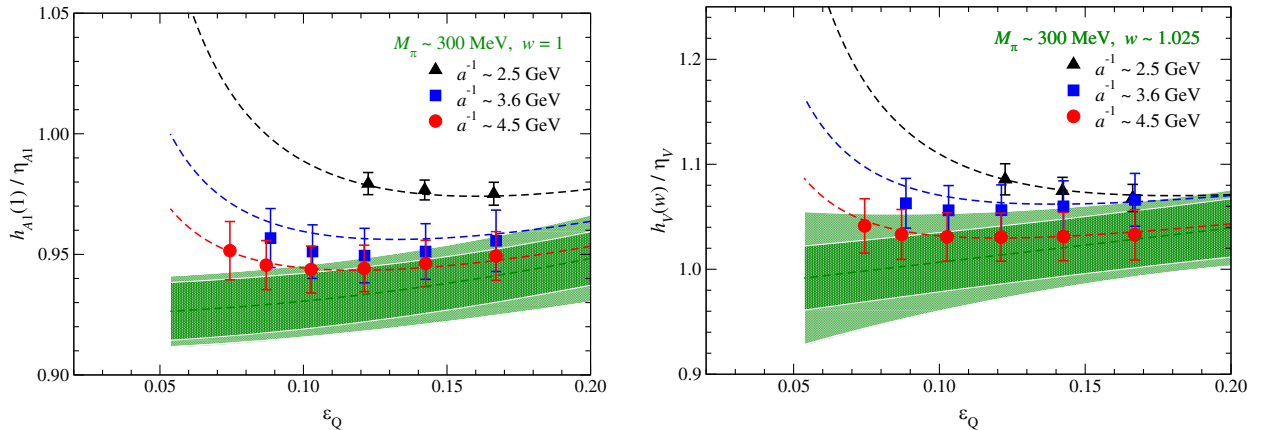


FIG. 12. Same as Fig. 11 but as a function of HQET expansion parameter $\epsilon_Q$. To cancel nontrivial $m_Q$ dependence due to the matching factor $\eta_{\{A_1,V\}}$, we plot a ratio to this factor, which is fitted to the HMChPT-based polynomial form (19) in $\epsilon_Q$ and $\xi_{am_Q}$.

TABLE III. Fit parameters from continuum and chiral extrapolation of form factors (19).

| $h_X$ | $\chi^2/\text{d.o.f.}$ | $c$ | $c_\pi$ | $c_{\eta_s}$ | $c_Q$ | $d_Q$ | $c_a$ | $c_{am_Q}$ | $c_w$ | $d_w$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $h_{A_1}$ | 0.14(11) | 0.898(54) | 0.68(19) | 0.08(16) | $-0.18(32)$ | 0.59(17) | 0.37(26) | 0.0376(89) | $-0.953(70)$ | 0.67(22) |
| $h_{A_2}$ | 0.18(13) | $-0.5(1.0)$ | 0.9(3.6) | $-0.0(2.6)$ | 0.13(49) | $\cdots$ | 3.3(3.0) | $-0.022(67)$ | $-0.41(86)$ | 10.3(7.1) |
| $h_{A_3}$ | 0.18(13) | 0.91(90) | $-0.4(3.3)$ | 0.3(2.4) | 1.25(47) | $\cdots$ | $-3.4(2.7)$ | 0.117(62) | $-1.72(75)$ | 1.9(6.1) |
| $h_V$ | 0.15(14) | 0.87(13) | 1.47(45) | 0.39(37) | 0.33(13) | $\cdots$ | $-0.51(63)$ | 0.092(18) | $-1.32(14)$ | 0.6(1.3) |

dependent discretization error is significant in our results. However, the net $m_Q$ dependence is described reasonably well with our fitting form (19). We note that the $m_Q$ dependence for $h_{A_2}$ and $h_{A_3}$ is insignificant due to their large relative uncertainty. As we will mention later, fit results do not change significantly if we exclude data at $0.5 \leq am_Q \leq 0.7$ in our continuum and chiral extrapolation.

In this study, we first extrapolate the form factors to the continuum limit and physical quark masses, and then parametrize the $w$ dependence by using the model independent BGL parametrization, because it may not be simply combined with the chiral and heavy quark expansions. Concerning the $w$ dependence, therefore, the fit form (19) with an expansion in $w - 1$ is used to interpolate our results. Figure 10 shows that our data do not show strong curvature in the simulation region of $1 \leq w \lesssim 1.1$, and can be well described with only up to the quadratic term.

Numerical results of our continuum and chiral extrapolation are summarized in Table III. The polynomial coefficients turn out to be $O(1)$ or less with our choice of the dimensionless expansion parameters (20), which are normalized by appropriate scales. Many of them are consistent with zero reflecting the mild dependence on the lattice spacing and quark masses mentioned above. As a result, our choice of the fitting forms (19) results in small values of $\chi^2/\text{d.o.f.} \lesssim 0.2$, which, however, are only a rough measure of the goodness of the fits, because we employ approximated estimators of the covariance matrix as discussed in the next subsection.

## B. Systematic uncertainties

### 1. Correlated fit

With limited statistics, the covariance matrix $C$ of our form factor data has exceptionally small eigenvalues with large statistical error. In this study, we test two methods to take the correlation into account in the combined continuum and chiral extrapolation. First, as in the study of $B \rightarrow \pi \ell \nu_\ell$ [27], we introduce a lower cut of the eigenvalue (singular value) $\lambda_{\text{cut}}$. Effects of the exceptionally small eigenvalues are suppressed by replacing eigenvalues smaller than $\lambda_{\text{cut}}$ by $\lambda_{\text{cut}}$ in the singular value decomposition of $C$ as

$$C \sim \sum_{k \leq n_{\text{cut}}} \lambda_{\text{cut}} u_k u_k^T + \sum_{k > n_{\text{cut}}} \lambda_k u_k u_k^T, \quad (21)$$

where $u_k$ represents the eigenvector corresponding to the eigenvalue $\lambda_k$. Here eigenvalues are sorted in ascending order ($\lambda_k < \lambda_{k+1}$) and $n_{\text{cut}}$ satisfies $\lambda_{n_{\text{cut}}} < \lambda_{\text{cut}} < \lambda_{n_{\text{cut}}+1}$. We choose $\lambda_{\text{cut}}$ such that all eigenvalues below $\lambda_{\text{cut}}$ have statistical error larger than 100%. The second method is the so-called shrinkage estimate of the covariance matrix [57,58], which amounts to replacing $C_{ij}$ by

$$\rho \delta_{ij} C_{ii} + (1 - \rho) C_{ij}. \quad (22)$$

With $0 \leq \rho \leq 1$, the shrinkage is an interpolation between the uncorrelated ($\rho = 1$) and correlated ($\rho = 0$) fits. An optimal value of $\rho$ can be estimated by minimizing a loss function [58]. We observed, however, that results of the continuum and chiral extrapolation do not strongly depend on the choice of $\lambda_{\text{cut}}$ and $\rho$. In this study, we therefore employ the singular value cut for our main results, since this is a conservative approach, which only tends to increase the final error. The shift of the fit results obtained with the shrinkage estimator is treated as a systematic error due to the estimate of $C$.

With the approximated estimators of $C$, $\chi^2/\text{d.o.f.}$ is only a rough measure of the goodness of the fit. The impact is shown in Fig. 13. We observe that, in the case of our analysis, $\chi^2/\text{d.o.f.}$ tends to be slightly smaller than that for
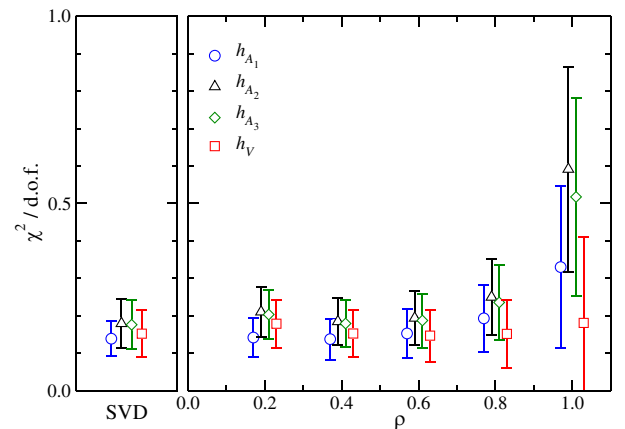


FIG. 13. Values of $\chi^2/\text{d.o.f.}$ for our continuum and chiral extrapolation of form factors. The left panel shows $\chi^2/\text{d.o.f.}$ with the singular value cut for the covariance matrix, whereas the right panel shows that with the shrinkage estimator as a function of the parameter $\rho$ in Eq. (22). The error is estimated by the bootstrap method.

the unquenched fit, and rather stable against the choice of the estimator (singular value cut, or shrinkage) and parameter $\rho$. While $C$ is not invertible for the correlated fit (with $\rho = 0$ for the shrinkage estimator), we may expect that $\chi^2/\text{d.o.f.}$ does not change drastically and serves as a "rough" measure of the goodness of the fit even with the approximated estimators of $C$. We also note that, in this paper, the approximated estimators are used only for the continuum and chiral extrapolation of form factors due to the large size of $C$ with different choices of the quark masses (including $m_Q$ for the relativistic approach), lattice cutoff, and recoil parameter.

### 2. Fitting form

The highest order is quadratic in the recoil parameter $w - 1$ expansion and heavy quark expansion for $h_{A_1}$, and linear for other expansions including chiral and finite $a$ expansions. The systematic error due to this choice is estimated by repeating the continuum and chiral extrapolation without one of the highest order terms, if the corresponding coefficient is poorly determined, otherwise by adding a yet higher order term.

Possible higher order corrections in the chiral expansion are also examined by testing the $x$ expansion. We also test a multiplicative form,

$$
\begin{aligned}
h_X/\eta_X = c &+ \left(1 + \frac{g_{D^*D\pi}^2}{2\Lambda_\chi^2}\bar{F}_{\log}(\xi_\pi, \Delta_D, \Lambda_\chi) + c_\pi\xi_\pi + c_{\eta_s}\xi_{\eta_s}\right)(1 + c_Q\epsilon_Q) \\
&\times (1 + c_a\xi_a + c_{am_Q}\xi_{am_Q})\{1 + c_w(w - 1) + d_w(w - 1)^2\},
\end{aligned}
\tag{23}
$$

to examine the potential effect of cross terms. We find that these alternative fits yield a change of the fit results well below the statistical error.

### 3. Inputs

The lattice cutoff $a^{-1}$, physical meson masses $M_\pi$, $M_K$, $M_{\eta_b}$, and the $D^*D\pi$ coupling $g_{D^*D\pi}$ are inputs to the continuum and chiral extrapolation. Since the masses have been measured very precisely, we take account of the uncertainty due to $a^{-1}$ and $g_{D^*D\pi}$ by repeating the extrapolation with each input shifted by $\pm 1\sigma$.

We set $M_\pi$ and $M_K$ to those in the isospin limit as in our determination of $a$ [27]. To examine the isospin breaking effects to the form factors, we test the continuum and chiral extrapolation to $M_\pi = M_{\pi^0}$ and $M_\pi = M_{\pi^\pm}$ as well as $M_K = M_{K^0}$ and $M_{K^\pm}$. We also test the heavy quark expansion parameter replaced with $\epsilon_Q = \bar{\Lambda}/2M_B$ and compare the extrapolations using $M_B = M_{B^0}$ and $M_{B^\pm}$. The difference in the form factor can be considered as an estimate of the strong isospin breaking effect. It turns out to be small as suggested by the mild quark mass dependence of the form factors. We include this in the systematic uncertainty so that our synthetic data for the form factors can be used for both the neutral and charged $B$ meson decays.

### 4. Finite volume effects (FVEs)

Our spatial lattice size satisfies $M_\pi L \geq 4$. At our smallest pion mass on the coarsest lattice, we also simulate a smaller size $L \simeq 3.0 M_\pi^{-1}$ to directly examine FVEs. Comparison between the two volumes at that simulation point does not show any significant deviation: for instance, $h_{A_1}$ at $w = 1$ is consistent within statistical error of $\lesssim 1\%$. We also repeat the continuum and chiral extrapolation by replacing the form factors at the smallest $M_\pi$ with those on the smaller $L$. The change in the fit results is included in the systematic uncertainty, but turns out to be well below the statistical uncertainty. We conclude that finite volume effects are not significant in our simulation region of $M_\pi \gtrsim 230$ MeV.

### 5. Accuracy of continuum and chiral extrapolation

Figure 14 shows the estimate of various uncertainties for the form factors in the continuum limit and at the physical quark masses as a function of $w$. As discussed in Sec. II, the axial [vector] matrix element is exclusively sensitive to $h_{A_1}(w)$ [$h_V(w)$] with an appropriate choice of the $D^*$ meson polarization vector and momentum. Combined with a precise determination at zero recoil, the total uncertainty of $h_{A_1}$ is about 2% in our simulation region of $w$, and the largest uncertainty comes from statistics and discretization effects. On the other hand, either $D^*$ or $B$ needs to have nonzero momentum to determine $h_V$, which is, therefore, less accurate. The total accuracy is $\lesssim 7\%$ with 3%–4% statistical and $\sim 5\%$ discretization errors.

Since the simulated bottom quark masses are not far below the lattice cutoff ($m_Q \leq 0.7a^{-1}$), we repeat the continuum and chiral extrapolation only using the data with $m_Q \leq 0.5a^{-1}$. The change in the fit results is well below the statistical uncertainty. We therefore conclude that our continuum extrapolation in the relativistic approach is controlled.

On the other hand, the axial form factors, $h_{A_2}$ and $h_{A_3}$, have larger uncertainty, i.e. 40% and 15%, respectively, dominated by the statistical error. Towards a more precise determination, it would be advantageous to simulate moving $B$ mesons so as to have matrix elements more sensitive to these form factors.
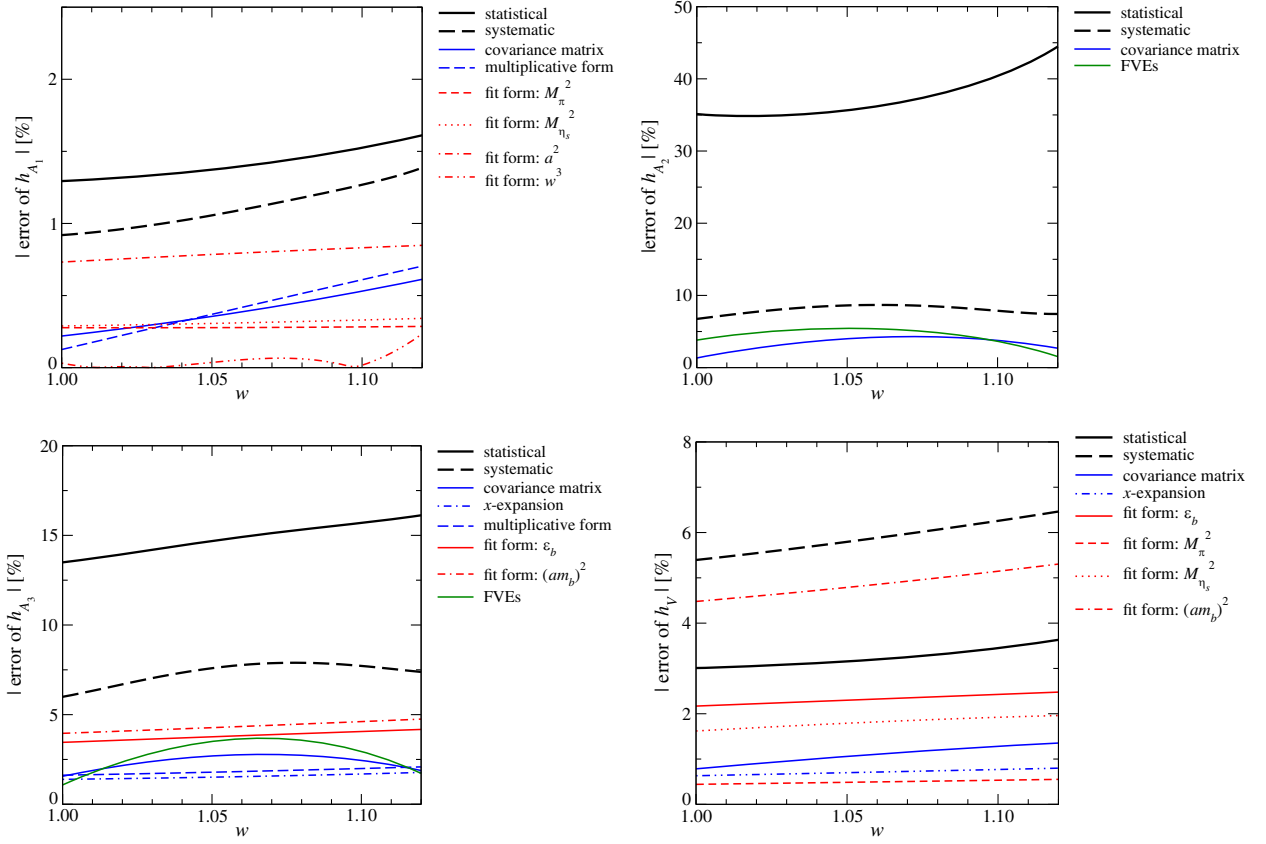
FIG. 14. Uncertainties of form factors as a function of $w$. The top-left, top-right, bottom-left, and bottom-right panels show uncertainties of $h_{A_1}$, $h_{A_2}$, $h_{A_3}$, and $h_V$, respectively. The thick solid and dashed lines show the statistical and total systematic uncertainties, while the thin lines show a breakdown of the latter. Just for clarity, we omit small systematic errors below 0.2% ($h_{A_1}$), 4% ($h_{A_2}$), 1% ($h_{A_3}$), and 0.5% ($h_V$) in the whole simulated region of $w$.

Previous determinations of $|V_{cb}|$ often employed a HQET-based parametrization by Caprini, Lellouch, and Neubert (CLN) of $h_{A_1}$ and form factor ratios

$$R_1(w) = \frac{h_V(w)}{h_{A_1}(w)}, \qquad R_2(w) = \frac{rh_{A_2}(w) + h_{A_3}(w)}{h_{A_1}(w)}, \quad (24)$$

where higher order HQET corrections are expected to be partially canceled [59]. Figure 15 shows that these ratios also have mild parameter dependences and can be safely extrapolated to the continuum limit and physical quark masses. We, however, do not employ the model-dependent CLN parametrization in the following analysis.

## IV. FORM FACTORS AS A FUNCTION OF THE RECOIL PARAMETER

### A. Fit to lattice data

From the continuum and chiral extrapolation in the previous section, we generate synthetic data of the form factors in the relativistic convention,

$$\langle D^*(\epsilon', p')|V_\mu|B(p)\rangle = i\varepsilon_{\mu\nu\rho\sigma}\epsilon'^{*\nu}p'^{\rho}p^{\sigma}g(w), \quad (25)$$

$$\langle D^*(\epsilon', p')|A_\mu|B(p)\rangle = \epsilon_\mu'^* f(w) - (\epsilon'^* p)\{(p + p')_\mu a_+(w) \\ + (p - p')_\mu a_-(w)\}, \quad (26)$$

with which the BGL parametrization was originally proposed. A kinematic invariant $k = \sqrt{(t_+ - q^2)(t_- - q^2)/4t}$ is defined with the maximum value of the momentum transfer $q^2_{\max} = t_- = (M_B - M_{D^*})^2$, which corresponds to the zero recoil limit, and the threshold $t_+ = (M_B + M_{D^*})^2$ for the $W \to BD^*$ production channel. Instead of $a_+$ and $a_-$, it is common to use linear combinations

$$\mathcal{F}_1(w) = \frac{1}{M_{D^*}}\left\{\frac{1}{2}(M_B^2 - M_{D^*}^2 - q^2)f(w) + 2k^2q^2a_+(w)\right\}, \quad (27)$$

$$\mathcal{F}_2(w) = \frac{1}{M_{D^*}}\{f(w) + (M_B^2 - M_{D^*}^2)a_+(w) + q^2a_-(w)\}. \quad (28)$$

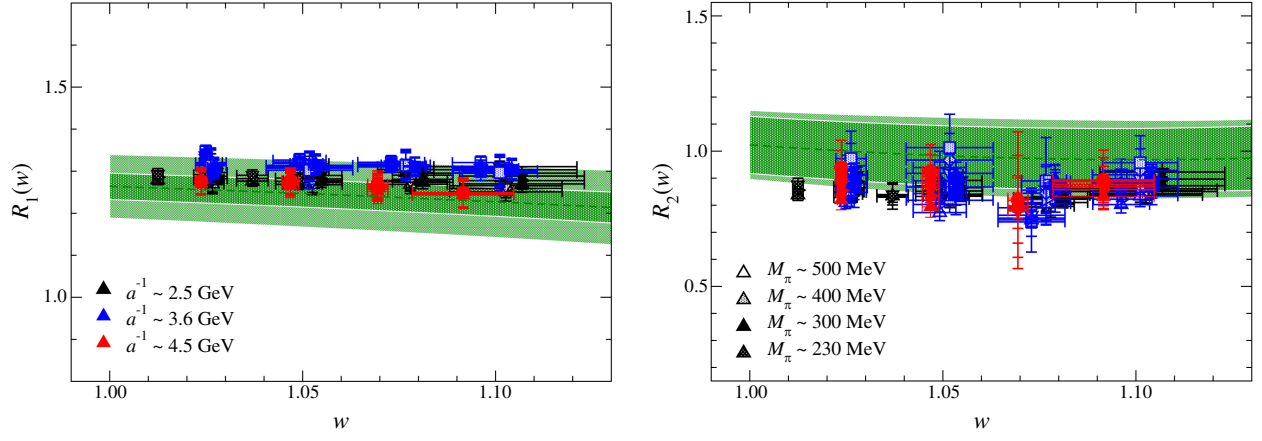These are related to the form factors in the HQET convention as

FIG. 15. Form factor ratios $R_1(w)$ (left panel) and $R_2(w)$ (right panel) as a function of recoil parameter $w$. Results at simulated parameters are plotted in symbols, which have the same meaning as in Fig. 10. The green bands show these ratios in the continuum limit and at physical quark masses reproduced by the continuum and chiral extrapolation of form factors (19).

$$g(w) = \frac{1}{M_B \sqrt{r}} h_V(w), \tag{29}$$

$$f(w) = M_B \sqrt{r}(w+1) h_{A_1}(w), \tag{30}$$

$$\mathcal{F}_1(w) = M_B^2 \sqrt{r}(w+1)\big[(w-r)h_{A_1}(w) - (w-1)\{rh_{A_2}(w) + h_{A_3}(w)\}\big], \tag{31}$$

$$\mathcal{F}_2(w) = \frac{1}{\sqrt{r}}\big\{(w+1)h_{A_1}(w) + (rw-1)h_{A_2}(w) + (r-w)h_{A_3}(w)\big\}. \tag{32}$$

We note that there are two kinematical constraints to be satisfied at the minimum ($w = 1$) and maximum values of $w$ [$w_{\max} = (1+r^2)/2r$ with $m_\ell = 0$]:

$$\mathcal{F}_1(1) = M_B(1-r)f(1), \tag{33}$$

$$\mathcal{F}_2(w_{\max}) = \frac{1+r}{M_B^2 r(1-r)(w_{\max}+1)} \mathcal{F}_1(w_{\max}). \tag{34}$$

With these form factors, we can write the helicity amplitudes for a given helicity of the intermediate $W$ boson, and the differential decay rate with respect to $w$ in simple forms

$$H_\pm(w) = f(w) \mp M_B^2 r \sqrt{w^2-1}\, g(w), \tag{35}$$

$$H_0(w) = \frac{1}{\sqrt{q^2}} \mathcal{F}_1(w), \tag{36}$$

$$H_S(w) = M_B r \sqrt{\frac{w^2-1}{1-2rw+r^2}} \mathcal{F}_2(w), \tag{37}$$

and

$$\frac{d\Gamma}{dw} = \frac{G_F^2}{16\pi^3} |V_{cb}|^2 |\eta_{\rm EW}|^2 M_B r^2 \sqrt{w^2-1}\left(1 - \frac{m_\ell^2}{q^2}\right)^2 \\ \times \left[\frac{q^2}{3}\left(1 + \frac{m_\ell^2}{2q^2}\right)\{|H_+(w)|^2 + |H_-(w)|^2 + |H_0(w)|^2\} + \frac{m_\ell^2}{2}|H_S(w)|^2\right]. \tag{38}$$

From the continuum and chiral extrapolation presented in Sec. III, synthetic data of $f$, $g$, and $\mathcal{F}_{1,2}$ are calculated at reference values of the recoil parameter $w_{\rm ref}$ through the relations (29)–(32). We take $w_{\rm ref} = 1.025$, $1.060$, and $1.100$, which span the whole simulated region of $w$, since our continuum and chiral extrapolation interpolates $h_{A_{\{1,2,3\}}}$ and $h_V$ in $w$. *Three* values are chosen, because the statistical covariance matrix of the synthetic data develops ill-determined eigenvalues with four or more values of $w_{\rm ref}$. This might be partly because, in our continuum and chiral extrapolation, we parametrize the $w$ dependence of each form factor using a quadratic form with three independent parameters. The numerical values together with their covariance matrix are presented in Table IV.

TABLE IV. Synthetic data of form factors $f$, $\mathcal{F}_{1,2}$, and $g$ in the relativistic convention. The central value and total error are in the second column, and the correlation matrix is in the third to fourteenth columns.

| | | $f(1.025)$ | $f(1.060)$ | $f(1.100)$ | $\mathcal{F}_1(1.025)$ | $\mathcal{F}_1(1.060)$ | $\mathcal{F}_1(1.100)$ | $\mathcal{F}_2(1.025)$ | $\mathcal{F}_2(1.060)$ | $\mathcal{F}_2(1.100)$ | $g(1.025)$ | $g(1.060)$ | $g(1.100)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f(1.025)$ | 5.6928(936) | 1.00000 | 0.98385 | 0.93944 | 0.94354 | 0.71378 | 0.49823 | 0.24518 | 0.22123 | 0.21755 | 0.22911 | 0.22641 | 0.22699 |
| $f(1.060)$ | 5.5879(992) | 0.98385 | 1.00000 | 0.98338 | 0.92328 | 0.70780 | 0.49984 | 0.22221 | 0.20401 | 0.20620 | 0.22112 | 0.22194 | 0.22694 |
| $f(1.100)$ | 5.473(109) | 0.93944 | 0.98338 | 1.00000 | 0.87531 | 0.68095 | 0.48820 | 0.19269 | 0.18105 | 0.18953 | 0.20747 | 0.21181 | 0.22233 |
| $\mathcal{F}_1(1.025)$ | 18.612(324) | 0.94354 | 0.92328 | 0.87531 | 1.00000 | 0.88532 | 0.70711 | 0.36140 | 0.35455 | 0.37432 | 0.20808 | 0.20305 | 0.20067 |
| $\mathcal{F}_1(1.060)$ | 18.308(410) | 0.71378 | 0.70780 | 0.68095 | 0.88532 | 1.00000 | 0.94630 | 0.58476 | 0.60631 | 0.65922 | 0.15762 | 0.15344 | 0.15170 |
| $\mathcal{F}_1(1.100)$ | 17.983(545) | 0.49823 | 0.49984 | 0.48820 | 0.70711 | 0.94630 | 1.00000 | 0.66784 | 0.70838 | 0.78248 | 0.11626 | 0.11328 | 0.11262 |
| $\mathcal{F}_2(1.025)$ | 2.1704(962) | 0.24518 | 0.22221 | 0.19269 | 0.36140 | 0.58476 | 0.66784 | 1.00000 | 0.98622 | 0.96645 | 0.05204 | 0.04696 | 0.04049 |
| $\mathcal{F}_2(1.060)$ | 2.097(106) | 0.22123 | 0.20401 | 0.18105 | 0.35455 | 0.60631 | 0.70838 | 0.98622 | 1.00000 | 0.98356 | 0.04418 | 0.03974 | 0.03438 |
| $\mathcal{F}_2(1.100)$ | 1.992(109) | 0.21755 | 0.20620 | 0.18953 | 0.37432 | 0.65922 | 0.78248 | 0.96645 | 0.98356 | 1.00000 | 0.04246 | 0.03872 | 0.03398 |
| $g(1.025)$ | 0.3326(212) | 0.22911 | 0.22112 | 0.20747 | 0.20808 | 0.15762 | 0.11626 | 0.05204 | 0.04418 | 0.04246 | 1.00000 | 0.99764 | 0.98991 |
| $g(1.060)$ | 0.3178(213) | 0.22641 | 0.22194 | 0.21181 | 0.20305 | 0.15344 | 0.11328 | 0.04696 | 0.03974 | 0.03872 | 0.99764 | 1.00000 | 0.99622 |
| $g(1.100)$ | 0.3013(215) | 0.22699 | 0.22694 | 0.22233 | 0.20067 | 0.15170 | 0.11262 | 0.04049 | 0.03438 | 0.03398 | 0.98991 | 0.99622 | 1.00000 |

The synthetic data are fitted to the model-independent BGL parametrization

$$F(z) = \frac{1}{P_F(z)\phi_F(z)} \sum_{k=0}^{N_F} a_{F,k} z^k \quad (F = f, g, \mathcal{F}_1, \mathcal{F}_2), \quad (39)$$

where the $z$ parameter

$$z(w) = \frac{\sqrt{w+1} - \sqrt{2}}{\sqrt{w+1} + \sqrt{2}} \quad (40)$$

maps the whole semileptonic kinematical region $0 \leq q^2[\text{GeV}^2] \lesssim 11 \, (1 \leq w \lesssim 1.5) \, (m_\ell = 0)$ into a small parameter region of $z$, $0 \leq z \lesssim 0.06$.

The Blaschke factors

$$P_f(z) = P_{\mathcal{F}_1}(z) = P_{1^+}(z), \qquad P_g(z) = P_{1^-}(z),$$
$$P_{\mathcal{F}_2}(z) = P_{0^-}(z) \quad (41)$$

factor out the pole singularities outside the semileptonic region but below the threshold $t_+$ due to the resonances for a given channel $J^P$ of $b\bar{c}$ mesons. The explicit form is given as

$$P_{J^P}(z) = \prod_k^{n_{\text{pole}}} \frac{z - z_{\text{pole},k}}{1 - zz_{\text{pole},k}}, \quad (42)$$

where

$$z_{\text{pole},k} = \frac{\sqrt{t_+ - M_{\text{pole},k}^2} - \sqrt{t_+ - t_-}}{\sqrt{t_+ - M_{\text{pole},k}^2} + \sqrt{t_+ - t_-}} \quad (43)$$

is the $z$ parameter corresponding to the $k$th resonance (in term of $q^2$ rather than $w$). In our main analysis, we employ resonance masses denoted as "Input-A" in Table V. This choice consists of slightly old experimental measurement [60] ($M_{\text{pole},\{0,1\}}$ for $J^P = 0^-$) and estimates based on lattice QCD [61,62] ($M_{\text{pole},0}$ for $1^+$ and $M_{\text{pole},\{0,1\}}$ for $1^-$) and quark potential models [63,64] (other $M_{\text{pole},n}$). This input has been used in some of the previous phenomenological analysis using the BGL parametrization [9,14] and lattice calculations of form factors [19,21]. Another choice ("Input-B" in Table V) in the literature is used to estimate the systematics uncertainty due to the use of input-A (see below).

The outer function $\phi_F(z)$ is chosen such that a constraint on the expansion coefficients derived from unitarity of the theory has a simple form:

$$\sum_{k=0}^{\infty} |a_{g,k}|^2 \leq 1, \qquad \sum_{k=0}^{\infty} |a_{f,k}|^2 + \sum_{k=0}^{\infty} |a_{\mathcal{F}_1,k}|^2 \leq 1,$$
$$\sum_{k=0}^{\infty} |a_{\mathcal{F}_2,k}|^2 \leq 1. \quad (44)$$

TABLE V. Two sets of input resonance masses with quantum number $J^P$. All values are in GeV units. The input-A is a compilation of slightly old experimental measurement [60] and lattice QCD [61,62] and phenomenological [63,64] studies. A different phenomenological estimate [65] for the (axial) vector resonances is used for input-B, where the pseudoscalar masses are set from a recent edition of the Particle Data Book [66] ($n = 0$, 1) and another phenomenological study [67] ($n = 2$). We employ input-A in our main analysis, whereas input-B is used to test the stability of the BGL parametrization against the choice of the input. See the text for more details.

| | Input-A | | Input-B | |
|---|---|---|---|---|
| $J^P$ | $n_{\text{pole}}$ | $M_{\text{pole},n}$ | $n_{\text{pole}}$ | $M_{\text{pole},n}$ |
| $1^+$ | 4 | 6.739, 6.750, 7.145, 7.150 | 4 | 6.730, 6.736, 7.135, 7.142 |
| $1^-$ | 3 | 6.329, 6.920, 7.020 | 4 | 6.337, 6.899, 7.012, 7.280 |
| $0^-$ | 3 | 6.275, 6.842, 7.250 | 3 | 6.274, 6,871, 7220 |

The explicit form is given as

$$\phi_g(z) = 16r^2 \sqrt{\frac{n_I}{3\pi \tilde{\chi}^T_{1^-}(0)}} \frac{(1+z)^2 (1-z)^{-1/2}}{\{(1+r)(1-z) + 2\sqrt{r}(1+z)\}^4},$$

(45)

$$\phi_f(z) = \frac{4r}{M_B^2} \sqrt{\frac{n_I}{3\pi \chi^T_{1^+}(0)}} \frac{(1+z)(1-z)^{3/2}}{\{(1+r)(1-z) + 2\sqrt{r}(1+z)\}^4},$$

(46)

$$\phi_{\mathcal{F}_1}(z) = \frac{4r}{M_B^3} \sqrt{\frac{n_I}{6\pi \chi^T_{1^+}(0)}} \frac{(1+z)(1-z)^{5/2}}{\{(1+r)(1-z) + 2\sqrt{r}(1+z)\}^5},$$

(47)

$$\phi_{\mathcal{F}_2}(z) = 8\sqrt{2} r^2 \sqrt{\frac{n_I}{\pi \chi^L_{1^+}(0)}} \frac{(1+z)^2 (1-z)^{-1/2}}{\{(1+r)(1-z) + 2\sqrt{r}(1+z)\}^4},$$

(48)

where $n_I = 2.6$ is the number of the spectator quarks with a correction due to SU(3) breaking. By $\chi^{\{T,L\}}_{J^P}(0)$, we denote the derivative of the vacuum polarization function of the weak current at zero momentum $q^2 = 0$ for a given channel $J^P$ and polarization $\{T, L\}$. As in the literature, we replace $\chi^T_{1^-}$ by $\tilde{\chi}^T_{1^-}$, from which the one-particle state contribution due to $B_c^*$ is subtracted for a better saturation of the unitarity bound (44). While a lattice QCD estimate is available [68], we employ a perturbative estimate [12,69,70] summarized in Table VI. We note that our choice of the resonance masses (input-A) and derivatives are the same as previous lattice studies to allow a direct comparison of the BGL

parametrization with them. It is known that the unitarity bound (44) is not well saturated by the $B \to D^* \ell \nu_\ell$ channel, and hence is rather weak. In this study, we do not impose this bound, but confirm that fit results satisfy it within the error.

Figure 16 shows the regular part of the axial form factor $P_f(z)\phi_f(z)f(z)$, which is to be expanded in $z$. We observe rather mild $z$ dependence of all form factors. While there is no clear sign of curvature in our simulation region, we employ a quadratic form $N_f = N_g = N_{\mathcal{F}_1} = N_{\mathcal{F}_2} = 2$ to safely suppress the truncation error. Among twelve coefficients, constants $a_{\mathcal{F}_1,0}$ and $a_{\mathcal{F}_2,0}$ are fixed to fulfill the kinematical constraints (33) at $w = 1$, where $z = 0$ and the linear and quadratic terms in the $z$-parameter expansion (39) vanish, and (34) at $w_{\text{max}}$, where $z \sim 0.05$ is well below unity, and hence the constant term gives rise to a large contribution to $\mathcal{F}_2$. To be consistent with the unitarity bound (44), we impose the bound $|a_{F,k}| \le 1$ ($k = 0$, 1, 2) for all form factors $F = g, f, \mathcal{F}_1$, and $\mathcal{F}_2$. Table VII shows fit results, which describe the synthetic data with an acceptable value of $\chi^2/\text{d.o.f.} = 0.51$.

We test the stability of the fit with different choices of the input and reference values of $w$. Input-B in Table V of the resonance masses for the Blaschke factors consists of the pseudoscalar masses $M_{\text{pole},\{0,1\}}$ from recent experimental measurements [66], and other resonance masses from phenomenological studies different from input-A: namely, Ref. [67] for $J^P = 0^-$ and $n = 2$, and Ref. [65] for $J^P = 1^{\{+,-\}}$. We note that these vector and axial masses have been used in a phenomenological analysis [10] and Belle analyses [16,17] of the $B \to D^* \ell \nu$ decay. For the derivatives of the vacuum polarization functions $\chi^{T,L}_{J^P}$, we test the lattice QCD estimate in Ref. [68]. Shift of the

TABLE VI. Derivatives of vacuum polarization functions entering other functions (45)–(48). Perturbative estimate (upper line) is employed in our main analysis.

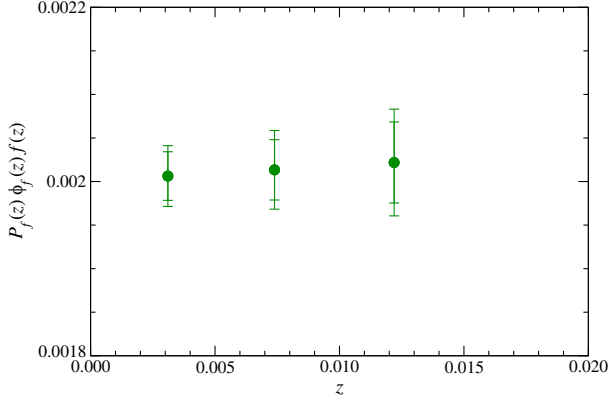| | $\chi^T_{1^+}(0)$ [GeV$^{-2}$] | $\chi^T_{1^-}(0)$ [GeV$^{-2}$] | $\chi^L_{1^+}(0)$ |
|---|---|---|---|
| Perturbation [12,69,70] | $3.894 \times 10^{-4}$ | $5.131 \times 10^{-4}$ | $1.9421 \times 10^{-2}$ |
| Lattice QCD [68] | $4.69(20) \times 10^{-4}$ | $5.84(44) \times 10^{-4}$ | $2.19(19) \times 10^{-2}$ |

FIG. 16. Regular part of axial form factor $P_f(z)\phi_f(z)f(z)$ as a function of $z$. The two error bars in this figure and Fig. 17 show the statistical and total errors, respectively.

expansion coefficients $a_{\{g,f,\mathcal{F}_1,\mathcal{F}_2\},\{0,1\}}$ in units of their uncertainty is summarized in Table VIII, where we test four choices of the input: namely, (i) vector and axial resonance masses from Ref. [65], (ii) ground and first excited pseudoscalar masses from Ref. [66], (iii) second excited pseudo-scalar mass from Ref. [67], and (iv) lattice estimate of $\chi_{J^P}^{T,L}$. The shift of $a_{\{g,f,\mathcal{F}_1,\mathcal{F}_2\},\{0,1\}}$ is at the 2-$\sigma$ level or typically well below it. A large shift ($\sim 4\sigma$) in $a_{f,0}$ with the lattice QCD estimate of $\chi_{J^P}^{T,L}$ can be attributed to about 20% shift of $\chi_{1^+}^T$ for the axial channel. As suggested in Eq. (46), this simply leads to rescaling of all $a_{\{f,\mathcal{F}_1\},\{0,1,2\}}$ for the corresponding channel. Also for other coefficients, the changes in the Blaschke factors and outer functions due to the different inputs are largely absorbed into the shift in $a_{\{g,f,\mathcal{F}_1,\mathcal{F}_2\},\{0,1,2\}}$ to reproduce the synthetic data of the form factors. As a result, the shift in physical observables, $R(D^*)$ and $|V_{cb}|$ in this paper, is well below $1\sigma$ as we will see later. This also means that our results for $a_{\{g,f,\mathcal{F}_1,\mathcal{F}_2\},\{0,1\}}$ in Table VII should be used with our choice of the input, namely input-A for resonance masses in Table V and perturbative estimate of $\chi_{J^P}^{T,L}$ in Table VI, otherwise the uncertainty suggested in Table VIII should be taken into account.

In Table VIII, we also test three different choices of the reference values $w_{\text{ref}}$: namely, (v) decreased maximum value $w_{\text{ref,max}} = 1.080$, (vi) increased minimum value $w_{\text{ref,min}} = 1.040$, and (vii) two values of $w_{\text{ref}} = 1.025, 1.100$ against our main choice of $w_{\text{ref}} = 1.025, 1.060, 1.100$. Table VIII shows that the shift of the coefficients $a_{\{g,f,\mathcal{F}_1,\mathcal{F}_2\},\{0,1\}}$ is well below $1\sigma$. Their relative accuracy is slightly worse (typically by 5%–10%) with the above choices of $w_{\text{ref}}$ suggesting that our main choice is reasonably good.

We confirm the unitarity bound for the axial vector channel as $\sum_{k=0}^{N_f} |a_{f,k}|^2 + \sum_{k=0}^{N_{\mathcal{F}_1}} |a_{\mathcal{F}_1,k}|^2 = 0.012(96)$, which is poorly saturated by $B \to D^* \ell \nu_\ell$. The bounds for other channels are satisfied as $\sum_{k=0}^{N_g} |a_{g,k}|^2 = 1.0(3.5)$

TABLE VII. Expansion coefficients and their correlation for BGL parametrization (39) of form factors $g$, $f$, $\mathcal{F}_1$, and $\mathcal{F}_2$. The parameters $a_{\mathcal{F}_1,0}$ and $a_{\mathcal{F}_2,0}$ are fixed from other parameters to satisfy kinematical constraints (33) and (34), respectively.

| | | $a_{g,0}$ | $a_{g,1}$ | $a_{g,2}$ | $a_{f,0}$ | $a_{f,1}$ | $a_{f,2}$ | $a_{\mathcal{F}_1,0}$ | $a_{\mathcal{F}_1,1}$ | $a_{\mathcal{F}_1,2}$ | $a_{\mathcal{F}_2,0}$ | $a_{\mathcal{F}_2,1}$ | $a_{\mathcal{F}_2,2}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $a_{g,0}$ | 0.0291(18) | 1.00000 | 0.46208 | 0.00157 | 0.23345 | 0.01973 | -0.00342 | 0.23345 | -0.00136 | 0.01362 | 0.08685 | -0.00553 | 0.01628 |
| $a_{g,1}$ | -0.045(35) | 0.46208 | 1.00000 | 0.01024 | 0.06133 | 0.21056 | 0.06890 | 0.06133 | -0.01804 | 0.02956 | -0.03921 | 0.00394 | 0.03000 |
| $a_{g,2}$ | -1.0(1.7) | 0.00157 | 0.01024 | 1.00000 | 0.01137 | 0.02396 | -0.01746 | 0.01137 | -0.00789 | 0.01312 | 0.00205 | 0.01056 | 0.00708 |
| $a_{f,0}$ | 0.01198(19) | 0.23345 | 0.06133 | 0.01137 | 1.00000 | 0.06216 | 0.12057 | 1.00000 | 0.07462 | -0.09380 | 0.39328 | 0.19208 | -0.13641 |
| $a_{f,1}$ | 0.018(11) | 0.01973 | 0.21056 | 0.02396 | 0.06216 | 1.00000 | -0.49732 | 0.06216 | 0.49162 | -0.24351 | 0.15088 | 0.45046 | -0.25008 |
| $a_{f,2}$ | -0.10(45) | -0.00342 | 0.06890 | -0.01746 | 0.12057 | -0.49732 | 1.00000 | 0.12057 | -0.33668 | 0.20338 | -0.18421 | -0.23045 | 0.19828 |
| $a_{\mathcal{F}_1,0}$ | 0.002006(31) | 0.23345 | 0.06133 | 0.01137 | 1.00000 | 0.06216 | 0.12057 | 1.00000 | 0.07462 | -0.09380 | 0.39328 | 0.19208 | -0.13641 |
| $a_{\mathcal{F}_1,1}$ | 0.0013(41) | -0.00136 | -0.01804 | -0.00789 | 0.07462 | 0.49162 | -0.33668 | 0.07462 | 1.00000 | -0.51671 | 0.17631 | 0.84705 | -0.49244 |
| $a_{\mathcal{F}_1,2}$ | -0.03(21) | 0.01362 | 0.02956 | 0.01312 | -0.09380 | -0.24351 | 0.20338 | -0.09380 | -0.51671 | 1.00000 | 0.19200 | -0.48706 | 0.98112 |
| $a_{\mathcal{F}_2,0}$ | 0.0484(16) | 0.08685 | -0.03921 | 0.00205 | 0.39328 | 0.15088 | -0.18421 | 0.39328 | 0.17631 | 0.19200 | 1.00000 | 0.07882 | 0.12820 |
| $a_{\mathcal{F}_2,1}$ | -0.059(87) | -0.00553 | 0.00394 | 0.01056 | 0.19208 | 0.45046 | -0.23045 | 0.19208 | 0.84705 | -0.48706 | 0.07882 | 1.00000 | -0.55376 |
| $a_{\mathcal{F}_2,2}$ | -0.9(1.1) | 0.01628 | 0.03000 | 0.00708 | -0.13641 | -0.25008 | 0.19828 | -0.13641 | -0.49244 | 0.98112 | 0.12820 | -0.55376 | 1.00000 |

TABLE VIII. Stability of BGL parametrization coefficients $a_{\{g,f,\mathcal{F}_1,\mathcal{F}_2\},\{0,1\}}$ given in Table VII. We list their shift in units of their uncertainty due to different choices of input and reference values $w_{\text{ref}}$. We omit $a_{\mathcal{F}_1,0}$, which is proportional to $a_{f,0}$ due to the kinematical constraint (33), as well as $a_{\{g,f,\mathcal{F}_1,\mathcal{F}_2\},2}$ for quadratic terms, which show tiny shift due to their large relative uncertainty. Note that $a_{g,n}$ is independent of the pseudoscalar resonance mass input. The maximum (minimum) value of $w_{\text{ref}}$ is denoted as $w_{\text{ref,max}}$ ($w_{\text{ref,min}}$). The quadratic fit ($N_{\{f,g,\mathcal{F}_1,\mathcal{F}_2\}} = 2$) is impossible with two values of $w_{\text{ref}}$, with which we compare coefficients of the linear fit ($N_{\{f,g,\mathcal{F}_1,\mathcal{F}_2\}} = 1$).

| Choice of input-B, $\chi_{J^P}^{T,L}$ or $w_{\text{ref}}$ | $a_{g,0}$ | $a_{g,1}$ | $a_{f,0}$ | $a_{f,1}$ | $a_{\mathcal{F}_1,1}$ | $a_{\mathcal{F}_2,0}$ | $a_{\mathcal{F}_2,1}$ |
|---|---|---|---|---|---|---|---|
| (i) $M_{\text{pole},n}$ for $J^P = 1^+, 1^-$ | $-1.62$ | $0.32$ | $-2.22$ | $0.04$ | $0.03$ | $0.01$ | $0.00$ |
| (ii) $M_{\text{pole},\{0,1\}}$ for $J^P = 0^-$ | $\cdots$ | $\cdots$ | $0.00$ | $0.00$ | $0.00$ | $0.57$ | $-0.04$ |
| (iii) $M_{\text{pole},2}$ for $J^P = 0^-$ | $\cdots$ | $\cdots$ | $0.00$ | $0.00$ | $0.00$ | $-1.59$ | $0.10$ |
| (iv) $\chi_{J^P}^{T,L}$ from lattice | $-0.74$ | $0.08$ | $-4.18$ | $-0.11$ | $0.02$ | $-1.24$ | $0.03$ |
| (v) $w_{\text{ref,max}} = 1.080$ | $-0.00$ | $-0.01$ | $-0.02$ | $-0.05$ | $-0.10$ | $0.12$ | $0.31$ |
| (vi) $w_{\text{ref,min}} = 1.040$ | $-0.00$ | $-0.01$ | $-0.03$ | $-0.02$ | $-0.04$ | $0.30$ | $-0.30$ |
| (vii) 2 $w_{\text{ref}}$'s = 1.025, 1.100 | $-0.01$ | $-0.09$ | $-0.04$ | $-0.08$ | $0.01$ | $0.11$ | $-0.04$ |

and $\sum_{k=0}^{N_{\mathcal{F}_2}} |a_{\mathcal{F}_2,k}|^2 = 0.9(2.1)$. The large uncertainty comes from the poorly determined quadratic coefficients. The sums up to the linear term, $\sum_{k=0}^{1} |a_{g,k}|^2 = 0.0029(31)$ and $\sum_{k=0}^{1} |a_{\mathcal{F}_2,k}|^2 = 0.006(10)$, suggest that the unitarity bound is rather weak as mentioned above.

The BGL parametrization and synthetic data of the form factors are plotted in Fig. 17. We observe reasonable consistency with the Fermilab/MILC results [19] except the vector form factor $g(w)$ near zero recoil and the axial form factor $\mathcal{F}_2$ extrapolated to large recoils. This can be seen in a more detailed comparison of the expansion
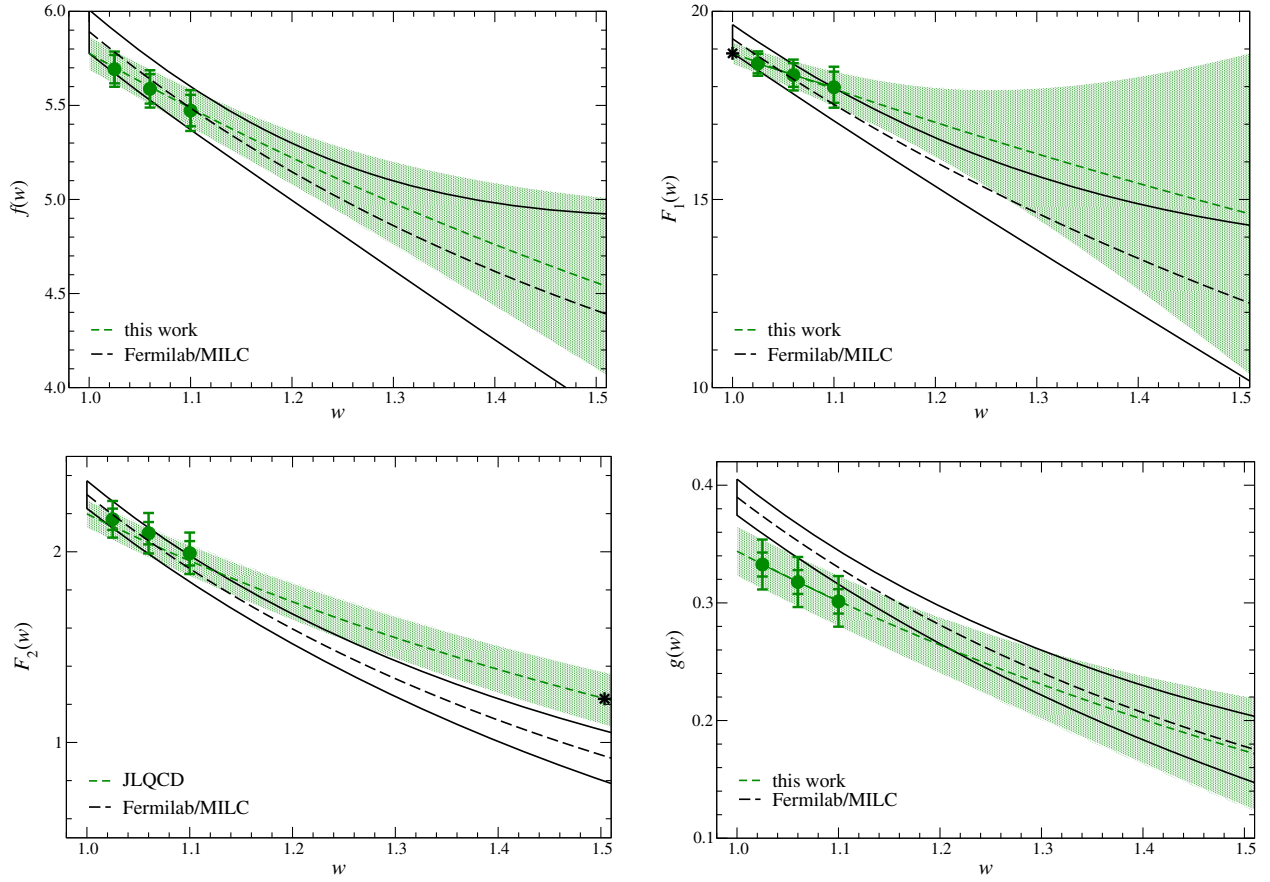


FIG. 17. BGL parametrization (green band) of the synthetic data (green circles) in the whole semileptonic region $1 \le w \lesssim 1.5$. The top-left, top-right, bottom-left, and bottom-right panels show $f(w)$, $\mathcal{F}_1(w)$, $\mathcal{F}_2(w)$, and $g(w)$, respectively. We also plot the parametrization by Fermilab/MILC [19] by the open black bands. The stars in the panels for $\mathcal{F}_1$ and $\mathcal{F}_2$ represent the values fixed from the kinematical constraints (33) and (34), respectively.
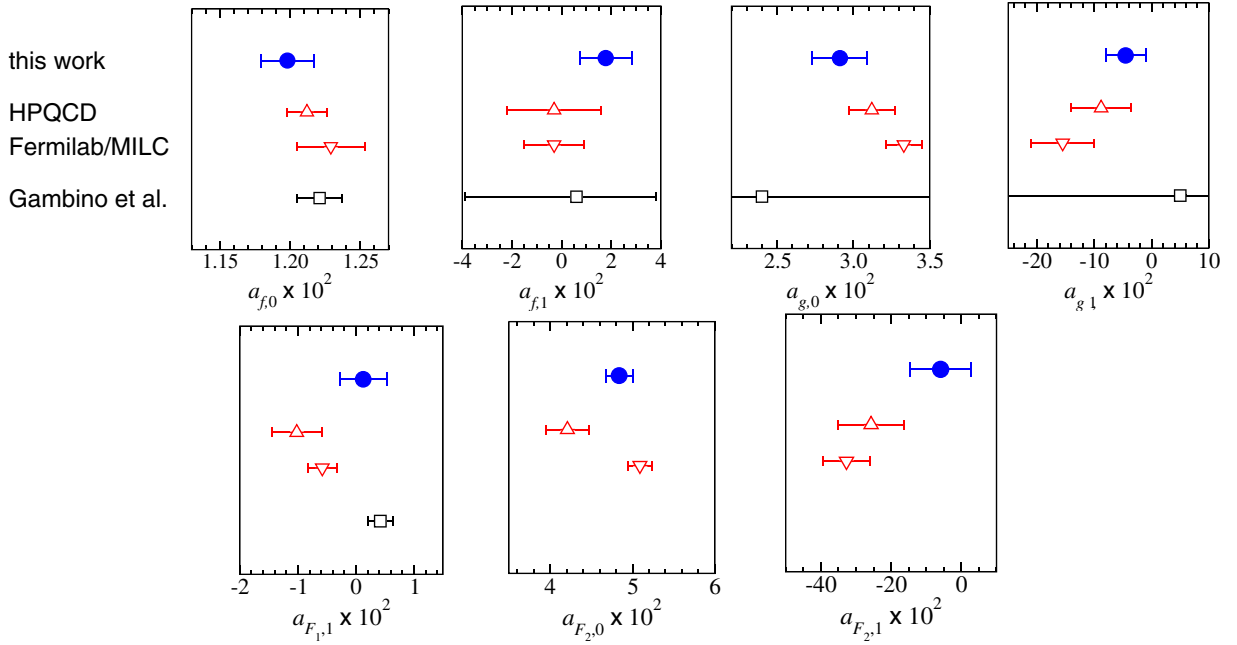
FIG. 18.   Comparison of expansion coefficients for BGL parametrization (39). Our results are plotted as filled blue circles. Open red triangles are obtained by HPQCD (triangles up) [21] and Fermilab/MILC (triangles down) [19]. We also plot results from a recent phenomenological analysis of the Belle data [16] as open black squares [14]. Since all these studies impose the kinematical constraint (33) leading to $a_{\mathcal{F}_1,0} \propto a_{f,0}$, we omit a panel for $a_{\mathcal{F}_1,0}$.

coefficients in Fig. 18, where there is slight tension in $a_{g,0}$, $a_{g,1}$, and $a_{\mathcal{F}_1,1}$. Recent HPQCD results show better consistency with ours, except for $a_{\mathcal{F}_1,1}$ and $a_{\mathcal{F}_2,0}$.

In the same figure, we also plot a phenomenological estimate [14] obtained from the Belle data [16] combined with a lattice input of the FLAG average of $h_{A_1}(1) \propto f(1)$ [71]. Note that this analysis employs the same BGL parametrization with the input ($M_{\mathrm{pole},n}$ and $\chi_{J^P}^{\{T,L\}}$) and orders ($N_f$, $N_g$, $N_{\mathcal{F}_1}$), which are the same as our main analysis. The phenomenological result of $a_{f,0}$ is well determined and shows good consistency with lattice results, simply because it is mainly fixed from the lattice input. Other poorly determined coefficients suggest the importance of lattice calculations to obtain controlled BGL parametrization.

On the other hand, the slope $a_{\mathcal{F}_1,1}$ is well determined from the phenomenological analysis of the experimental data, which are precise at large recoils. This study observes a reasonable agreement in $a_{\mathcal{F}_1,1}$ and, hence, in the differential decay rate as discussed in the next subsection (see Fig. 19). We note, however, that the Fermilab/MILC and HPQCD studies reported tension in the $w$ dependence of the differential decay rate with experiment.

Since $\mathcal{F}_2$ describes the contribution suppressed by $m_\ell^2$ to the decay rate (38), its expansion coefficients are poorly constrained by the experimental data for the light lepton channels $B \to D^* \ell \nu_\ell$ ($\ell = e, \mu$). The lattice determination

of $\mathcal{F}_2$ is, therefore, helpful towards a precision new physics search using the $\tau$ channel. Through numerical integration of Eq. (38) for $\ell = e$, $\mu$, and $\tau$ with fit results in Table VII, we obtain a pure SM value

$$R(D^*) = 0.252(22)(4), \tag{49}$$

which does not resort to phenomenological models nor experimental data. The second error comes from the test of the stability of the BGL parametrization summarized in Table VIII. It mainly comes from the shift in $w_{\mathrm{ref\,min}}$ and $w_{\mathrm{ref\,max}}$, but smaller than statistical and other systematic errors of the BGL parametrization. This estimate can be improved as $R(D^*) = 0.252(+0.009/-0.016)(4)$ by imposing the unitarity bound (44) on the expansion coefficients. These are consistent with Fermilab/MILC and HPQCD estimates, 0.265(13) and 0.279(13), respectively, and to be compared with the experimental average 0.295(14) [1]. For a more precise and reliable estimate of $R(D^*)$, it is helpful to simulate larger recoils as well as resolving the tensions in $a_{\mathcal{F}_2,\{0,1\}}$ shown in Fig. 18.

## B. Fit including Belle data

In this subsection, we carry out a simultaneous fit to our lattice and Belle data to estimate $|V_{cb}|$. The differential decay rate with respect to $w$ and three decay angles $\theta_l$, $\theta_v$, and $\chi$ is given as
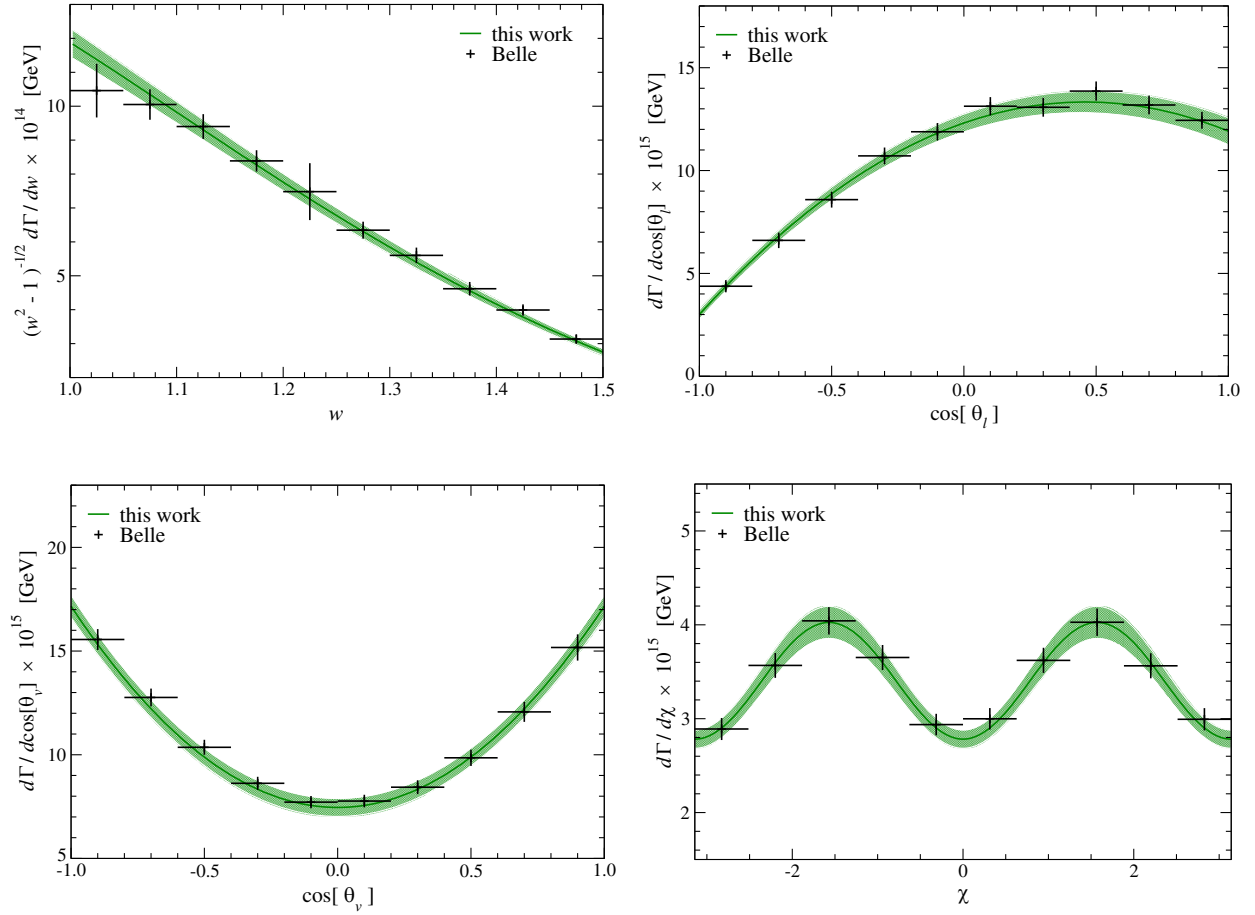
FIG. 19. Differential decay rate with respect to recoil parameter, $d\Gamma/dw$ (top-left panel), and three decay angles, $d\Gamma/d\cos[\theta_\ell]$ (top-right panel), $d\Gamma/d\cos[\theta_v]$ (bottom-left panel), and $d\Gamma/d\chi$ (bottom-right panel). We divide $d\Gamma/dw$ by a phase space factor $\sqrt{w^2-1}$ to make its $w$ dependence milder and monotonous. The symbols show Belle results, whereas the green band is reproduced from our simultaneous fit to our lattice and Belle data.

$$\frac{d\Gamma}{dw\,d\cos[\theta_\ell]\,d\cos[\theta_v]\,d\chi} = \frac{3G_F^2}{1024\pi^4}|V_{cb}|^2\eta_{\mathrm{EW}}^2 M_B r^2 \sqrt{w^2-1}\,q^2\Big\{(1-\cos[\theta_\ell])^2\sin^2[\theta_v]H_+^2(w)+(1+\cos[\theta_\ell])^2\sin^2[\theta_v]H_-^2(w)$$

$$+4\sin^2[\theta_\ell]\cos^2[\theta_v]H_0^2(w)-2\sin^2[\theta_\ell]\sin^2[\theta_v]\cos[2\chi]H_+(w)H_-(w)$$

$$-4\sin[\theta_\ell](1-\cos[\theta_\ell])\sin[\theta_v]\cos[\theta_v]\cos[\chi]H_+(w)H_0(w)$$

$$+4\sin[\theta_\ell](1+\cos[\theta_\ell])\sin[\theta_v]\cos[\theta_v]\cos[\chi]H_-(w)H_0(w)\Big\} \tag{50}$$

with the helicity amplitudes given in (35) and (36). For a simultaneous fit with the Belle data for $B^0 \to D^{*-}\ell^+\nu_\ell$ ($\ell = e, \mu$) in Ref. [16], we set $m_\ell = 0$, and include the Coulomb factor $(1+\alpha\pi)$ in the right-hand side. The decay angles are chosen the same as for the Belle paper, where the full kinematical distribution is described by four differential decay rates $d\Gamma/dw$, $d\Gamma/d\cos[\theta_\ell]$, $d\Gamma/d\cos[\theta_v]$, and $d\Gamma/d\chi$, which are integrated over the other three variables. The differential decay rates are provided at ten equal-size bins in the range of each variable, $w \in [1.0, 1.5]$, $\cos[\theta_{\{\ell,v\}}] \in [-1, 1]$, and $\chi \in [-\pi, \pi]$.

To estimate $|V_{cb}|$, we fit our synthetic data of the form factors to the BGL parametrization (39) alongside the Belle data of the calculated differential decay rates where the expansion coefficients $a_{\{g,f,\mathcal{F}_1,\mathcal{F}_2\},\{0,1,2\}}$ are shared parameters and $|V_{cb}|$ is an additional parameter determined from the fit. The expression (38) is used for $d\Gamma/dw$. We calculate $d\Gamma/d\cos[\theta_{\{l,v\}}]$ and $d\Gamma/d\chi$ by analytically integrating over two other decay angles and by numerically integrating over $w$ via Simpson's rule with 800 steps within each bin as in the Belle paper [16].

TABLE IX. Fit results of simultaneous fit to our lattice and Belle data.

| $F$ | $g$ | $f$ | $\mathcal{F}_1$ | $\mathcal{F}_2$ |
|---|---|---|---|---|
| $a_{F,0}$ | 0.02896(84) | 0.01196(18) | 0.002003(30) | 0.0489(14) |
| $a_{F,1}$ | −0.050(23) | 0.0227(92) | 0.0040(14) | −0.018(24) |
| $a_{F,2}$ | −1.0(1.9) | −0.46(24) | −0.040(27) | −1.0(2.0) |

Our results for the expansion coefficients are summarized in Table IX. These are consistent with those in Table VII, which are obtained without the experimental data, with slightly smaller uncertainty. We obtain

$$|V_{cb}| = 39.19(90)(12) \times 10^{-3}, \qquad (51)$$

where the first error comes from our main fit with input-A in Table V, perturbative $\chi_{J^P}^{\{T,L\}}$ in Table VI, and $w_{\text{ref}} = 1.025$, 1.060, 1.100. The second error is obtained by testing the stability of the fit against the different choices of the input and $w_{\text{ref}}$ as in Table VIII. This is in good agreement with the previous determination from the exclusive decay (3). In contrast to previous lattice studies, as shown in Fig. 19, we observe consistency between our lattice and Belle data leading to an acceptable value of $\chi^2/\text{d.o.f.} \sim 0.90$. Towards the resolution of the $|V_{cb}|$ tension, we need a more detailed analysis taking care of, for instance, bias in $|V_{cb}|$ discussed in Refs. [14,15] due to D'Agostini effects [72] and the strong systematic correlation of experimental data. That analysis is left for future work.

## V. CONCLUSION

In this article, we present our calculation of the $B \to D^* \ell \nu_\ell$ semileptonic form factors in $2 + 1$-flavor lattice QCD. The Möbius domain-wall action is employed for all relevant quark flavors to simulate lattices of cutoffs up to 4.5 GeV with chiral symmetry preserved to good accuracy. This removes the leading $O(a)$ discretization errors and explicit renormalization is not necessary to calculate the SM form factors through the correlator ratios. The ground state saturation of the correlator ratios is carefully studied by simulating four source-sink separations. The bottom quark mass is limited to $m_b \leq 0.7a^{-1}$ to suppress $O(a^2)$ and higher discretization error. Note also that the chiral logarithm is suppressed by the small $D^*$–$D$ mass splitting squared. As a result, the lattice spacing and quark mass dependence turns out to be mild in our simulation region leading to good control of the continuum and chiral extrapolation. We find finite size effects to be small by direct simulations of two volumes at our smallest pion mass.

One of the main results is the synthetic data of the form factors $g$, $f$, $\mathcal{F}_1$, and $\mathcal{F}_2$ in the relativistic convention at three reference values of the recoil parameter $w = 1.025$, 1.060, and 1.100, which can be used in future analyses to determine $|V_{cb}|$ and to search for new physics. Through the BGL parametrization, we obtain $R(D^*) = 0.252(22)$ in the SM. A combined analysis with the Belle data yields $|V_{cb}| = 39.19(91) \times 10^{-3}$. These are consistent with previous lattice studies. However, the expansion coefficients for $g$, $\mathcal{F}_1$, and $\mathcal{F}_2$ show significant tension among lattice studies. In particular, our slope for $\mathcal{F}_1$ is slightly larger and consistent with a phenomenological analysis of the Belle data. As a result, our data show better consistency with the Belle data in the differential decay rates.

High statistics simulations on finer lattices would be helpful in resolving this tension. Simulating larger recoils can lead to a better estimate of $R(D^*)$ as well as more detailed comparison with experiment in a wide range of the recoil parameter to search for new physics. In order to resolve the $|V_{cb}|$ tension, it would also be very helpful to study the inclusive decay on the lattice for more direct comparison of the exclusive and inclusive analyses in the same simulations [73]. While it is an ill-posed problem to calculate the relevant hadronic tensor from discrete lattice data on a finite volume, a joint project of the JLQCD and RBC/UKQCD collaborations along a strategy proposed in Refs. [74,75] is in progress [76–78] towards a realistic study of $B \to X_c \ell \nu_\ell$.

[1] Y. S. Amhis *et al.* (Heavy Flavor Averaging Group), Phys. Rev. D **107**, 052008 (2023).

[2] A. Crivellin and S. Pokorski, Phys. Rev. Lett. **114**, 011802 (2015).

[3] A. Vaquero, Proc. Sci. LATTICE2022 (**2023**) 250.

[4] T. Kaneko, Proc. Sci. LATTICE2022 (**2023**) 238.

[5] C. Bernard *et al.* (Fermilab/MILC Collaborations), Phys. Rev. D **79**, 014506 (2009).

[6] J. A. Bailey *et al.* (Fermilab/MILC Collaborations), Phys. Rev. D **89**, 114504 (2014).

[7] J. Harrison, C. Davies, and M. Wingate (HPQCD Collaboration), Phys. Rev. D **97**, 054502 (2018).

[8] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, Phys. Rev. D **95**, 115008 (2017); **97**, 059902 (E) (2018).

[9] D. Bigi, P. Gambino, and S. Schacht, Phys. Lett. B **769**, 441 (2017).

[10] B. Grinstein and A. Kobach, Phys. Lett. B **771**, 359 (2017).

[11] C. G. Boyd, B. Grinstein, and R. F. Lebed, Phys. Rev. D **56**, 6895 (1997).

[12] D. Bigi, P. Gambino, and S. Schacht, J. High Energy Phys. 11 (2017) 061.

[13] F. U. Bernlochner, Z. Ligeti, and D. J. Robinson, Phys. Rev. D **100**, 013005 (2019).

[14] P. Gambino, M. Jung, and S. Schacht, Phys. Lett. B **795**, 386 (2019).

[15] D. Ferlewicz, P. Urquijo, and E. Waheed, Phys. Rev. D **103**, 073005 (2021).

[16] E. Waheed *et al.* (Belle Collaboration), Phys. Rev. D **100**, 052007 (2019); **103**, 079901(E) (2021).

[17] A. Abdesselam *et al.* (Belle Collaboration), arXiv:1702.01521.

[18] We note that more recent data on the normalized distribution from the full Belle dataset [79] as well as a recent report from the succeeding SuperKEKB/Belle II experiment [80] are available in a different event-reconstruction approach.

[19] A. Bazavov *et al.* (Fermilab/MILC Collaborations), Eur. Phys. J. C **82**, 1141 (2022); **83**, 21(E) (2023).

[20] A. X. El-Khadra, A. S. Kronfeld, and P. B. Mackenzie, Phys. Rev. D **55**, 3933 (1997).

[21] J. Harrison and C. T. H. Davies (HPQCD Collaboration), arXiv:2304.03137.

[22] R. C. Brower, H. Neff, and K. Orginos, Comput. Phys. Commun. **220**, 1 (2017).

[23] S. Hashimoto, A. X. El-Khadra, A. S. Kronfeld, P. B. Mackenzie, S. M. Ryan, and J. N. Simone, Phys. Rev. D **61**, 014502 (1999).

[24] T. Kaneko, Y. Aoki, B. Colquhoun, H. Fukaya, and S. Hashimoto (JLQCD Collaboration), Proc. Sci. LATTICE2018 (**2019**) 311.

[25] T. Kaneko, Y. Aoki, G. Bailas, B. Colquhoun, H. Fukaya, S. Hashimoto, and J. Koponen (JLQCD Collaboration), Proc. Sci. LATTICE2019 (**2020**) 139.

[26] T. Kaneko, Y. Aoki, B. Colquhoun, M. Faur, H. Fukaya, S. Hashimoto, J. Koponen, and E. Kou (JLQCD Collaboration), Proc. Sci. LATTICE2021 (**2022**) 561.

[27] B. Colquhoun, S. Hashimoto, T. Kaneko, and J. Koponen (JLQCD Collaboration), Phys. Rev. D **106**, 054502 (2022).

[28] S. Borsanyi *et al.*, J. High Energy Phys. 09 (2012) 010.

[29] P. Weisz, Nucl. Phys. **B212**, 1 (1983).

[30] P. Weisz and R. Wohlert, Nucl. Phys. **B236**, 397 (1984); **B247**, 544(E) (1984).

[31] M. Luscher and P. Weisz, Commun. Math. Phys. **97**, 59 (1985); **98**, 433(E) (1985).

[32] P. A. Boyle (RBC/UKQCD Collaborations), Proc. Sci. LATTICE2014 (**2015**) 087.

[33] C. Morningstar and M. J. Peardon, Phys. Rev. D **69**, 054501 (2004).

[34] S. Aoki *et al.* (JLQCD Collaboration), Phys. Rev. D **78**, 014508 (2008).

[35] S. Aoki *et al.* (JLQCD and TWQCD Collaborations), Prog. Theor. Exp. Phys. **2012**, 01A106 (2012).

[36] T. Kaneko, S. Aoki, G. Cossu, H. Fukaya, S. Hashimoto, and J. Noaki (JLQCD Collaboration), Proc. Sci. LATTICE2013 (**2014**) 125.

[37] L. Randall and M. B. Wise, Phys. Lett. B **303**, 135 (1993).

[38] M. J. Savage, Phys. Rev. D **65**, 034014 (2002).

[39] K. Nakayama, B. Fahy, and S. Hashimoto (JLQCD Collaboration), Phys. Rev. D **94**, 054507 (2016).

[40] S. Aoki *et al.* (JLQCD, TWQCD Collaborations), Phys. Lett. B **665**, 294 (2008).

[41] H. Fukaya *et al.* (JLQCD Collaboration), Phys. Rev. Lett. **98**, 172001 (2007).

[42] H. Fukaya, S. Aoki, T. W. Chiu, S. Hashimoto, T. Kaneko, H. Matsufuru, J. Noaki, K. Ogawa, T. Onogi, and N. Yamada (TWQCD Collaboration), Phys. Rev. D **76**, 054503 (2007).

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

[44] S. Aoki, H. Fukaya, S. Hashimoto, and T. Onogi, Phys. Rev. D **76**, 054508 (2007).

[45] S. Aoki, G. Cossu, X. Feng, H. Fukaya, S. Hashimoto, T. Kaneko, J. Noaki, and T. Onogi (JLQCD Collaboration), Phys. Rev. D **96**, 034501 (2017).

[46] B. Fahy, G. Cossu, and S. Hashimoto (JLQCD Collaboration), Proc. Sci. LATTICE2016 (**2017**) 118.

[47] W. Detmold, C. J. D. Lin, and S. Meinel, Phys. Rev. D **85**, 114508 (2012).

[48] D. Becirevic and F. Sanfilippo, Phys. Lett. B **721**, 94 (2013).

[49] K. U. Can, G. Erkol, M. Oka, A. Ozpineci, and T. T. Takahashi, Phys. Lett. B **719**, 103 (2013).

[50] F. Bernardoni, J. Bulava, M. Donnellan, and R. Sommer (ALPHA Collaboration), Phys. Lett. B **740**, 278 (2015).

[51] J. M. Flynn, P. Fritzsch, T. Kawanai, C. Lehner, B. Samways, C. T. Sachrajda, R. S. Van de Water, and O. Witzel (RBC/UKQCD Collaborations), Phys. Rev. D **93**, 014510 (2016).

[52] J. Noaki *et al.* (JLQCD and TWQCD Collaborations), Phys. Rev. Lett. **101**, 202004 (2008).

[53] A. F. Falk, H. Georgi, B. Grinstein, and M. B. Wise, Nucl. Phys. **B357**, 185 (1990).

[54] A. F. Falk and B. Grinstein, Phys. Lett. B **249**, 314 (1990).

[55] M. Neubert, Nucl. Phys. **B371**, 149 (1992).

[56] M. E. Luke, Phys. Lett. B **252**, 447 (1990).

[57] C. Stein, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics* (University of California Press, Berkeley, CA, 1956), p. 197.

[58] O. Ledoit and M. Wolf, J. Multivariate Anal. **88**, 365 (2004).

[59] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. **B530**, 153 (1998).

[60] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. C **40**, 100001 (2016).

[61] R. J. Dowdall, C. T. H. Davies, T. C. Hammant, and R. R. Horgan (HPQCD Collaboration), Phys. Rev. D **86**, 094510 (2012).

[62] B. Colquhoun, C. T. H. Davies, R. J. Dowdall, J. Kettle, J. Koponen, G. P. Lepage, and A. T. Lytle (HPQCD Collaboration), Phys. Rev. D **91**, 114509 (2015).

[63] S. Godfrey, Phys. Rev. D **70**, 054017 (2004).

[64] A. K. Rai and N. Devlani, Proc. Sci. Hadron2013 (**2013**) 045.

[65] E. J. Eichten and C. Quigg, Phys. Rev. D **49**, 5845 (1994).

[66] R. L. Workman *et al.* (Particle Data Group), Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).

[67] X.-J. Li, Y.-S. Li, F.-L. Wang, and X. Liu, Eur. Phys. J. C **83**, 1080 (2023).

[68] G. Martinelli, S. Simula, and L. Vittorio, Phys. Rev. D **104**, 094512 (2021).

[69] J. Grigo, J. Hoff, P. Marquard, and M. Steinhauser, Nucl. Phys. **B864**, 580 (2012).

[70] D. Bigi and P. Gambino, Phys. Rev. D **94**, 094008 (2016).

[71] S. Aoki *et al.* (Flavour Lattice Averaging Group), Eur. Phys. J. C **80**, 113 (2020).

[72] G. D'Agostini, Nucl. Instrum. Methods Phys. Res., Sect. A **346**, 306 (1994).

[73] S. Hashimoto, Prog. Theor. Exp. Phys. **2017**, 053B03 (2017).

[74] M. T. Hansen, H. B. Meyer, and D. Robaina, Phys. Rev. D **96**, 094513 (2017).

[75] P. Gambino and S. Hashimoto, Phys. Rev. Lett. **125**, 032001 (2020).

[76] A. Barone, A. Jüttner, S. Hashimoto, T. Kaneko, and R. Kellermann, Proc. Sci. LATTICE2022 (**2023**) 403.

[77] R. Kellermann, A. Barone, S. Hashimoto, A. Jüttner, and T. Kaneko, Proc. Sci. LATTICE2022 (**2023**) 414.

[78] A. Barone, S. Hashimoto, A. Jüttner, T. Kaneko, and R. Kellermann, J. High Energy Phys. 07 (2023) 145.

[79] M. T. Prim *et al.* (Belle Collaboration), Phys. Rev. D **108**, 012002 (2023).

[80] F. Abudinén *et al.* (Belle II Collaboration), arXiv:2301.04716.