Eur. Phys. J. C (2022) 82:1141
https://doi.org/10.1140/epjc/s10052-022-10984-9

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Semileptonic form factors for $B \to D^*\ell\nu$ at nonzero recoil from 2 + 1-flavor lattice QCD

Fermilab Lattice and MILC Collaborations

A. Bazavov[1], C. E. DeTar[2], D. Du[3], A. X. El-Khadra[4,5], E. Gámiz[6], Z. Gelzer[4], S. Gottlieb[7], U. M. Heller[8], A. S. Kronfeld[9], J. Laiho[3], P. B. Mackenzie[9], J. N. Simone[9], R. Sugar[10], D. Toussaint[11], R. S. Van de Water[9], A. Vaquero[2,a],

[1] Department of Computational Mathematics, Science and Engineering, and Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA
[2] Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA
[3] Department of Physics, Syracuse University, Syracuse, NY 13244, USA
[4] Department of Physics, University of Illinois, Urbana, IL 61801, USA
[5] Illinois Center for Advanced Studies of the Universe, University of Illinois, Urbana, IL 61801, USA
[6] Departamento de Física Teórica y del Cosmos, Universidad de Granada, 18071 Granada, Spain
[7] Department of Physics, Indiana University, Bloomington, IN 47405, USA
[8] American Physical Society, Ridge, NY 11961, USA
[9] Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
[10] Department of Physics, University of California, Santa Barbara, CA 93106, USA
[11] Department of Physics, University of Arizona, Tucson, AZ 85721, USA

Received: 13 May 2022 / Accepted: 31 October 2022 / Published online: 16 December 2022

Abstract We present the first unquenched lattice-QCD calculation of the form factors for the decay $B \to D^*\ell\nu$ at nonzero recoil. Our analysis includes 15 MILC ensembles with $N_f = 2+1$ flavors of asqtad sea quarks, with a strange quark mass close to its physical mass. The lattice spacings range from $a \approx 0.15$ fm down to 0.045 fm, while the ratio between the light- and the strange-quark masses ranges from 0.05 to 0.4. The valence b and c quarks are treated using the Wilson-clover action with the Fermilab interpretation, whereas the light sector employs asqtad staggered fermions. We extrapolate our results to the physical point in the continuum limit using rooted staggered heavy-light meson chiral perturbation theory. Then we apply a model-independent parametrization to extend the form factors to the full kinematic range. With this parametrization we perform a joint lattice-QCD/experiment fit using several experimental datasets to determine the CKM matrix element $|V_{cb}|$. We obtain $|V_{cb}| = (38.40 \pm 0.68_{\text{th}} \pm 0.34_{\text{exp}} \pm 0.18_{\text{EM}}) \times 10^{-3}$. The first error is theoretical, the second comes from experiment and the last one includes electromagnetic and electroweak uncertainties, with an overall $\chi^2/\text{dof} = 126/84$, which illustrates the tensions between the experimental data sets, and between theory and experiment. This result is in agreement with previous exclusive determinations, but the tension with the inclusive determination remains. Finally, we integrate the differential decay rate obtained solely from lattice data to predict $R(D^*) = 0.265 \pm 0.013$, which confirms the current tension between theory and experiment.

1 Introduction

High precision tests of the standard model (SM) offer exciting possibilities for discovering new physics. In particular, the flavor sector of the SM is very rich in phenomena that can be used to explore physics beyond the standard model (BSM). Most flavor physics revolves around the Cabibbo–Kobayashi–Maskawa (CKM) matrix, which relates the mass and flavor eigenstates of the quarks. Since it is a basis transformation, the CKM matrix is constrained by unitarity, so violations of this rule could indicate the influence of new physics. Weak processes that are loop-suppressed in the SM may also expose new physics. To determine CKM matrix elements to high precision and to perform precision tests of the SM in measurements of rare decay processes, it is essential

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1140/epjc/s10052-022-10984-9.

[a] e-mail: alexv@unizar.es (corresponding author)

to know the strong-interaction environment in which these processes occur.

Among the CKM matrix elements, $|V_{cb}|$ has arguably been one of the most perplexing. There is a long-standing tension between the determination of this element via exclusive and inclusive decays. The operator product expansion (OPE) is used to analyze inclusive decay experiments measuring semileptonic decays $B \to X_c \ell \nu$, where X_c represents any charmed hadron or combination of hadrons with a single c quark. On the other hand, exclusive decay experiments focus on decays with a specific charmed hadron in the final state, for example, $B \to D\ell\nu$ or $B \to D^*\ell\nu$. We expect both types of experiments to yield consistent results for $|V_{cb}|$; however, there is a $\sim 3\sigma$ discrepancy between the inclusive and exclusive determinations [1,2]. Since contributions from new physics are unlikely to explain these differences [3,4], the disagreement presents an obstacle to higher precision tests of the SM.

We turn now to the determination of $|V_{cb}|$ from the exclusive decay $B \to D^*\ell\nu$, which has an interesting history. As is detailed below in Eq. (7), to determine $|V_{cb}|$, a measurement of the differential decay rate $d\Gamma/dw$ is needed and the form factors must be computed by theory. Using lattice QCD, it has been possible to determine the key form factor at zero recoil. However, the differential decay rate vanishes at that point because of kinematic factors, so it is necessary to use the differential decay rate at nonzero recoil to extrapolate the form factors to zero recoil. Since the 1990s, there have been two parametrizations of the form factors, one by Boyd, Grinstein, and Lebed (BGL) [5–7], and the other by Caprini, Lellouch, and Neubert (CLN) [8]. The CLEO and BaBar experiments [9–11], for example, relied on the CLN parametrization to analyze the dependence of the recoil parameter of their data. This was also true for earlier reviews from the Heavy Flavor Averaging Group (HFLAV) [12].

The situation changed in 2017 with the publication of unfolded data from a $B \to D^*\ell\nu$ experiment by the Belle Collaboration [13]. This data release was quickly followed by several theoretical analyses [14–17] comparing the effect of the choice of parametrization on $|V_{cb}|$. They found that the CLN parametrization [8], at least as it is usually employed to extrapolate the experimental data to the zero-recoil point (see for instance, Sec.V.A of Ref. [18]), does not provide a good description of the experimental data, whereas the BGL parametrization [5–7] describes the data properly and yielded exclusive determinations of $|V_{cb}|$ that were compatible with the inclusive ones [14–17]. However, more recent analyses by the Belle Collaboration using the much larger untagged dataset [18] and by the BaBar Collaboration performing a new analysis of their old data [19] contradicted this picture and reinforced the long-standing tension between the inclusive and the exclusive determinations. Newer theoretical analyses using Belle's untagged dataset also found agreement between CLN and BGL results [20,21]. These analyses also conclude that CLN is still a useful parametrization given the current errors in the experimental measurements. Unfortunately, previous unquenched lattice-QCD calculations of the $B \to D^*$ form factor [22,23] cannot provide constraints on the shape, because they are limited to zero recoil. Hence, a precise calculation from first principles of the form factors involved in the exclusive process performed for a range of nonzero-recoil momenta, could be extremely helpful.

Another motivation to study this process is the existing tension between experimental measurements and SM predictions of several lepton-flavor-universality-violating (LFUV) observables in B-meson semileptonic decays [24]. The ratios of branching fractions of the semitauonic and other semileptonic $B \to D^{(*)}$ transitions

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

disagree with the SM at the $\sim 3\sigma$ level when $R(D)$ and $R(D^*)$ are taken together [1]. Although the last HFLAV average [1] shows a large discrepancy between theory and experiment, the most recent measurements from the BaBar, Belle, and LHCb Collaborations find $R(D^*)$ to be closer to SM expectations [25–29]. However, a complete lattice-QCD calculation of $R(D^*)$ is still lacking. In view of current tensions in these observables, the limitations of the available theoretical predictions, and the future improvements on the experimental side expected from LHCb and Belle II forthcoming data, an independent theoretical calculation with a tight control of systematic errors that could help to either confirm or reduce the tension is urgently needed.

In this work, we use lattice QCD to address these two points, i.e., tensions between exclusive and inclusive determinations of $|V_{cb}|$, and between the SM theoretical prediction and experiment for $R(D^*)$. Although lattice QCD has previously been used to extract $|V_{cb}|$ from experimental data for $B \to D^*\ell\nu$, the relevant decay amplitude has always been computed at zero recoil [22,30–32], except for an early study in the quenched approximation [33]. Here we compute the form factors that contribute to the $B \to D^*\ell\nu$ decay for nonzero values of the recoil parameter in full QCD with $2+1$ flavors of dynamical sea quarks and extrapolate their behavior to the large recoil region. Instead of using the standard procedure of extrapolating experimental results to zero recoil and then extracting $|V_{cb}|$ using the calculated value of the form factor at zero recoil, we do a joint fit of lattice and experimental data where $|V_{cb}|$ is one of the free parameters. Once the decay amplitude is determined, we integrate over the whole kinematic space to find the branching ratios with and without a τ in the decay. We calculate $R(D^*)$ and compare our results with existing experimental determinations. Preliminary reports of this analysis were presented in Refs. [34–38]. In keeping with previous work on the same

ensembles [39–42], this analysis was blinded until a systematic error budget was finalized. The final results were then frozen, apart from unblinding.

This article is organized as follows. In Sect. 2, we introduce the formalism and present the form factors that take part in our calculation. Section 3 gives details on the ensembles available and also describes the analysis of the lattice data up to the chiral-continuum extrapolation. In Sect. 4, we discuss the systematic errors, and Sect. 5 shows the z expansion and the joint fit with experimental data that leads to our final results for $|V_{cb}|$ and $R(D^*)$. Section 6 presents our conclusions. Appendix A includes details on the fit function, employed in the chiral-continuum extrapolation. Appendix B outlines the calculation of the matching factors and its errors. Appendix C explains in detail how the κ tuning correction for the heavy quarks was calculated. Finally, Appendix D provides a guide to ancillary files containing complete results for the form factors, with a full correlation matrix.

2 Form factor definitions

In this section, we set the definitions and notation used in the following sections for the form factors, ratio of correlators, currents, and renormalization factors, among others.

2.1 Form factors in the continuum

The $B \to D^* \ell \nu$ process is mediated by the axial $\mathcal{A}^\mu = \bar{c}\gamma_\mu\gamma_5 b$ and the vector $\mathcal{V}^\mu = \bar{c}\gamma_\mu b$ electroweak currents. The transition matrix for this process is usually decomposed into form factors inspired by the heavy quark effective theory (HQET) [43]:

$$\frac{\langle D^*(p_{D^*},\varepsilon)|\mathcal{A}^\mu|B(p_B)\rangle}{\sqrt{M_{D^*}M_B}} = i\varepsilon_\nu^*\Big[g^{\mu\nu}(w+1)\,h_{A_1}(w) - v_B^\nu(v_B^\mu h_{A_2}(w) + v_{D^*}^\mu h_{A_3}(w))\Big], \tag{2}$$

$$\frac{\langle D^*(p_{D^*},\varepsilon)|\mathcal{V}^\mu|B(p_B)\rangle}{\sqrt{M_{D^*}M_B}} = \epsilon^{\mu\nu}_{\rho\sigma}\varepsilon_\nu^* v_B^\rho v_{D^*}^\sigma\, h_V(w), \tag{3}$$

where ε is the polarization vector of the D^* meson, M_Y is the mass of the $Y = B$, D^* meson and p_Y its respective momentum. From the four-velocities $v_Y = p_Y/M_Y$, one can define the recoil parameter $w = v_B \cdot v_{D^*}$.

To express the differential decay rate, it is convenient to introduce helicity amplitudes according to the polarization of the off-shell W boson [44]:

$$H_\pm(w) = (w+1)\left[h_{A_1}(w) \mp \sqrt{\frac{w-1}{w+1}}h_V(w)\right], \tag{4}$$

$$H_0(w) = y(w+1)\Big\{(w-r)h_{A_1}(w) - (w-1)\left[rh_{A_2}(w) + h_{A_3}(w)\right]\Big\}, \tag{5}$$

$$H_S(w) = y\sqrt{w^2-1}\Big[(w+1)h_{A_1}(w) - (1-wr)h_{A_2}(w) - (w-r)h_{A_3}(w)\Big], \tag{6}$$

where $r = M_{D^*}/M_B$ and $y^2(1-2wr+r^2) = 1$. The differential decay rate for $B^- \to D^{0*}\ell^-\bar{\nu}$ is then

$$\frac{d\Gamma}{dw} = |V_{cb}|^2|\eta_{\rm EW}|^2\frac{G_F^2 M_B^5}{16\pi^3}\left(1-\frac{m_\ell^2}{q^2}\right)^2 r^3(w^2-1)^{1/2} \times\left\{\frac{1}{3y^2}\left(1+\frac{m_\ell^2}{2q^2}\right)\left[|H_+|^2+|H_-|^2+|H_0|^2\right] + \frac{m_\ell^2}{2M_B^2}|H_S|^2\right\}, \tag{7}$$

where $\eta_{\rm EW}$ is a short-distance electroweak correction [45], G_F is the Fermi constant determined from muon decay, and m_ℓ is the charged lepton mass. Note that the scalar helicity amplitude's contribution is suppressed by $(m_\ell/M_B)^2$. In practice, it is neglected for semielectronic and semimuonic decays. The process $\bar{B}^0 \to D^{+*}\ell^-\bar{\nu}$ needs an extra factor $(1+\alpha\pi)$ on the right-hand side of Eq. (7) in order to account for the Coulomb attraction among the charged decay products [46–48]. Other electromagnetic effects, including structure-dependent corrections, are smaller, of order α/π instead of $\alpha\pi$ [46–48]. For the determination of $|V_{cb}|$, the full angular information of the decay chain $B \to D^*\ell\nu$ ($D^* \to D\pi$) is used, as discussed in Sect. 5.2.

As noted in the introduction, experimental measurements of the ratio of branching fractions in Eq. (1) are in tension with the SM. To date, the $B \to D^*$ form factors have been estimated with HQET, QCD sum rules, and input from experiment. With the results presented below, however, we can compute this ratio directly from (lattice) QCD:

$$\mathcal{B}(B \to D^*\ell\nu) = \tau_B\int_1^{w_{{\rm Max},\ell}} dw\,\frac{d\Gamma}{dw}, \tag{8}$$

where $w_{{\rm Max},\ell} = (1+r^2-m_\ell^2/M_B^2)/2r$. In $R(D^*)$, the B-meson lifetime τ_B drops out, so we form it from the ratio of partial widths.

Many papers in the literature introduce a decay amplitude $\mathcal{F}(w)$, defined by

$$|\mathcal{F}(w)|^2 = \frac{1-2wr+r^2}{w+1}\frac{|H_+|^2+|H_-|^2+|H_0|^2}{(5w+1)(1-r)^2-8rw(w-1)}, \tag{9}$$

and refer to $\mathcal{F}$ as a "form factor." For example, experimental results are often reported as $|V_{cb}|^2|\eta_{\rm EW}|^2|\mathcal{F}(w)|^2$. At zero recoil, $\mathcal{F}(1) = h_{A_1}(1)$. Thus, previous work in lattice QCD on this decay has focused on this single number, rather than

the four independent functions, $h_{A_1}(w)$, $h_V(w)$, $h_{A_2}(w)$, and $h_{A_3}(w)$, computed in this work.

2.2 Extracting the form factors from lattice matrix elements

Our heavy quarks (b, c) are simulated using the Fermilab action [49], as discussed in Sect. 3.1. In this framework, the lattice currents for the quark transition $y \to x$ are

$$V^{\mu}_{xy} = \bar{\Psi}_x \gamma^{\mu} \Psi_y, \tag{10}$$

$$A^{\mu}_{xy} = \bar{\Psi}_x \gamma^{\mu} \gamma_5 \Psi_y, \tag{11}$$

where x, y indicates the flavor c, b and Ψ is the Fermilab-improved field,

$$\Psi = (1 + d_1 \boldsymbol{\gamma} \cdot \boldsymbol{D}_{\text{lat}})\psi. \tag{12}$$

In this expression, the original heavy-quark field ψ is rotated in order to reduce the discretization errors. In particular, the coefficient d_1 must be calculated for each value of the quark mass in order to remove the $O(a)$ terms.

The lattice current J^{μ}_{xy} is related to its equivalent in the continuum $\mathcal{J}^{\mu}_{xy}$ through the renormalization factors,

$$\mathcal{J}^{\mu}_{xy} \doteq Z_{J^{\mu}_{xy}} J^{\mu}_{xy}, \tag{13}$$

where the $\doteq$ symbol means that both sides of the equation have the same matrix elements. In practice, the renormalization factors are only calculated approximately up to some order in a and α_s. In this work, we use a technique called *mostly nonperturbative renormalization* [50,51] that eliminates most of the nonperturbative dependence of the renormalization factors by defining the factors,

$$\rho^2_{J^{\mu}} = \frac{Z_{J^{\mu}_{cb}} Z_{J^{\mu}_{bc}}}{Z_{V^4_{cc}} Z_{V^4_{bb}}}. \tag{14}$$

When taking appropriate ratios of three-point correlators, the dominant, nonperburbative contribution to the renormalization of the currents, collected in the flavor-diagonal renormalization factors $Z_{V^4_{xx}}$, cancels. The remaining matching factors $\rho_{J^{\mu}}$ are amenable to a perturbative calculation [50]. We compute these matching factors to one-loop in perturbation theory, with the full $m_c a$ dependence at zero recoil, but $m_c a = 0$ at nonzero recoil. Of course, these simplifications introduce a variety of errors that must be kept under control. The truncation in the perturbative expansion is expected to be small, because α_s ranges in our case from 0.20 to 0.35. In addition, the coefficients of the expansion are small due to several cancelations [50]. The errors coming from the other two approximations are estimated in Appendix B and taken into account accordingly.

Not only does the use of ratios reduce the error in the calculation of the matching factors, but it also reduces the statistical fluctuations from the correlators. We set up the calculation in the rest frame of the B meson while the D^* meson carries a momentum $\boldsymbol{p}$, which determines the recoil w. The first ratio is the nonzero-recoil version of the double ratio [30],

$$R^2_{A_1} = \frac{\langle D^*(\boldsymbol{p}_\perp)|A_j|B(\boldsymbol{0})\rangle \langle B(\boldsymbol{0})|A_j|D^*(\boldsymbol{p}_\perp)\rangle}{\langle D^*(\boldsymbol{0})|V^4|D^*(\boldsymbol{0})\rangle \langle B(\boldsymbol{0})|V^4|B(\boldsymbol{0})\rangle}, \tag{15}$$

where the $\perp$ symbol in the momentum $\boldsymbol{p}_\perp$ indicates that the polarization of the D^* is aligned with the current and perpendicular to the momentum (i.e., transverse polarization). A parallel symbol $\parallel$ is used for longitudinal polarization. This double ratio yields h_{A_1}, which is the only form factor that survives at zero recoil.

The following single ratios

$$X_V = \frac{\langle D^*(\boldsymbol{p}_\perp)|V_j|B(\boldsymbol{0})\rangle}{\langle D^*(\boldsymbol{p}_\perp)|A_j|B(\boldsymbol{0})\rangle}, \tag{16}$$

$$X_0 = \frac{\langle D^*(\boldsymbol{p}_\parallel)|A^4|B(\boldsymbol{0})\rangle}{\langle D^*(\boldsymbol{p}_\perp)|A_j|B(\boldsymbol{0})\rangle}, \tag{17}$$

$$X_1 = \frac{\langle D^*(\boldsymbol{p}_\parallel)|A_j|B(\boldsymbol{0})\rangle}{\langle D^*(\boldsymbol{p}_\perp)|A_j|B(\boldsymbol{0})\rangle}, \tag{18}$$

yield the remaining form factors. Last, the ratio [52,53]

$$x_f = \frac{\langle D^*(\boldsymbol{p}_{(\alpha)})|V_j|D^*(\boldsymbol{0})\rangle}{\langle D^*(\boldsymbol{p}_{(\alpha)})|V^4|D^*(\boldsymbol{0})\rangle} \tag{19}$$

yields the recoil parameter w and involves only the flavor-diagonal transition $D^* \to D^*$. Here $(\alpha) = \perp, \parallel$ is the polarization, and must be the same in numerator and denominator to achieve the right cancelation of form factors.

The ratio in Eq. (19) yields the three-velocity,

$$x_f = \frac{v_{D^*}}{w + 1}, \tag{20}$$

from which it is straightforward to calculate w. In the B-meson rest frame ($\boldsymbol{v}_B = \boldsymbol{0}$),

$$w = \frac{1 + x_f^2}{1 - x_f^2}. \tag{21}$$

The ratio X_1 defined in Eq. (18) can be used to extract $h_{A_3}(w)$ as

$$X_1(w) = w - \frac{(w^2 - 1)h_{A_3}(w)}{(w + 1)h_{A_1}(w)}. \tag{22}$$

The matching factors for these two ratios, x_f and X_1, are $\rho = 1 + O(\alpha_s^2)$, and thus no renormalization is required at LO. In contrast, the remaining ratios require several nontrivial matching factors,

$$X_0(w) = \frac{\rho_{A_j}}{\rho_{A^4}} \sqrt{w^2 - 1} \left(1 - \frac{h_{A_2}(w) + w h_{A_3}(w)}{(w + 1)h_{A_1}(w)}\right), \tag{23}$$

Table 1 List of ensembles used in this work. The columns, from left to right, list the approximate lattice spacing, the scale-setting parameter r_1/a in lattice units, the ratio am_l/am_s between the light- and the strange-quark masses, the spatial length of the lattice in fm, the mass of the lightest pseudoscalar meson M_π^P in MeV, the dimensionless factor $M_\pi^P L$, the dimensions of the lattice in lattice units, the total sample size expressed as the number of sources × the number of configurations, and the tadpole-improvement factor u_0 obtained from the average plaquette

a (fm)	r_1/a	am_l/am_s	L (fm)	M_π^P (MeV)	$M_\pi^P L$	Lattice size	Sources		Configs	u_0
≈ 0.15	2.2215(57)	0.0097/0.0484	2.4	340	3.9	$16^3 \times 48$	24	×	631	0.8604
≈ 0.12	2.8211(28)	0.02/0.05	2.4	560	6.2	$20^3 \times 64$	4	×	2052	0.8688
	2.7386(33)	0.01/0.05	2.4	390	4.5	$20^3 \times 64$	4	×	2259	0.8677
	2.7386(33)	0.007/0.05	2.4	320	3.7	$20^3 \times 64$	4	×	2110	0.8678
	2.7386(33)	0.005/0.05	2.9	270	3.8	$24^3 \times 64$	4	×	2099	0.8678
≈ 0.09	3.8577(32)	0.0124/0.031	2.4	500	5.8	$28^3 \times 96$	4	×	1996	0.8788
	3.7887(34)	0.0062/0.031	2.4	350	4.1	$28^3 \times 96$	4	×	1931	0.8782
	3.7716(34)	0.00465/0.031	2.7	310	4.1	$32^3 \times 96$	4	×	984	0.8781
	3.7546(34)	0.0031/0.031	3.4	250	4.2	$40^3 \times 96$	4	×	1015	0.8779
	3.7376(34)	0.00155/0.031	5.5	180	4.8	$64^3 \times 96$	4	×	791	0.877805
≈ 0.06	5.399(17)	0.0072/0.018	2.9	450	6.3	$48^3 \times 144$	4	×	593	0.8881
	5.353(17)	0.0036/0.018	2.9	320	4.5	$48^3 \times 144$	4	×	673	0.88788
	5.330(16)	0.0025/0.018	3.4	260	4.4	$56^3 \times 144$	4	×	801	0.88776
	5.307(16)	0.0018/0.018	3.8	220	4.3	$64^3 \times 144$	4	×	827	0.88764
≈ 0.045	7.208(54)	0.0028/0.014	2.9	320	4.6	$64^3 \times 192$	4	×	801	0.89511

$$X_V(w) = \frac{\rho_{A_j}}{\rho_{V_j}} \frac{\sqrt{w-1}}{\sqrt{w+1}} \frac{h_V(w)}{h_{A_1}(w)}, \tag{24}$$

$$R_{A_1}(w) = \frac{w+1}{2} \frac{h_{A_1}(w)}{\rho_{A_j}}. \tag{25}$$

From these equations, it is quite easy to extract all form factors as a function of the ratios defined in Eqs. (15)–(19),

$$h_{A_1}(w) = \rho_{A_j} \frac{2R_{A_1}}{w+1}, \tag{26}$$

$$h_{A_2}(w) = \rho_{A_j} \frac{2R_{A_1}}{w^2-1} \left(wX_1 - \sqrt{w^2-1} \frac{\rho_{A^4}}{\rho_{A_j}} X_0 - 1 \right), \tag{27}$$

$$h_{A_3}(w) = \rho_{A_j} \frac{2R_{A_1}}{w^2-1} (w - X_1), \tag{28}$$

$$h_V(w) = \rho_{A_j} \frac{2R_{A_1}}{\sqrt{w^2-1}} \frac{\rho_{V_j}}{\rho_{A_j}} X_V. \tag{29}$$

These expressions determine the four form factors up to discretization and matching errors.

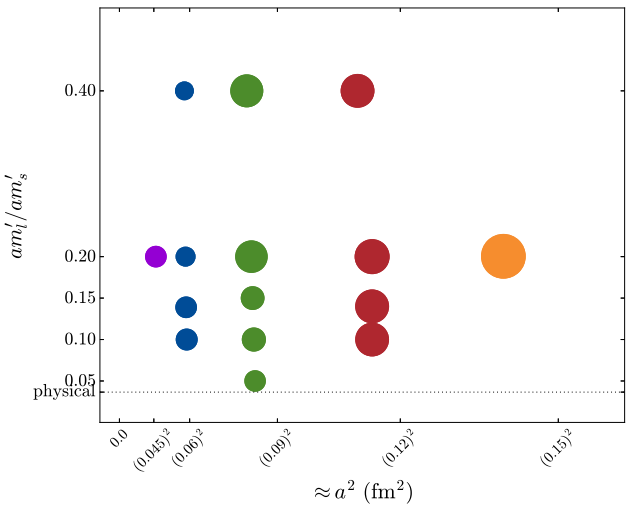

Fig. 1 Ensembles used in this work. The vertical (horizontal) axis shows the ratio between the light- and strange-quark masses (the lattice spacing). The area of each circle is proportional to the total sample size available on each ensemble. The horizontal line marks the value of the m_l/m_s ratio that results in pions with physical mass

3 Analysis

3.1 Lattice setup

In this analysis, we use 15 ensembles of gauge-field configurations, generated by the MILC Collaboration [54–56]. These ensembles include three flavors of asqtad-improved staggered sea quarks at five different lattice spacings, ranging from 0.15 fm in the coarsest case to 0.045 fm in the finest case. The mass of the strange sea quark is tuned to be close to its physical value, while the two light sea-quark masses are set equal, and cover a range of values that correspond to pion masses from $M_\pi \approx 560$ MeV to $M_\pi \approx 180$ MeV.

Table 2 Parameters used on each ensemble to generate the propagators for the valence, heavy quarks c and b. The approximate lattice spacing and the masses of the sea quarks (light and strange) in the first two columns identify the ensemble. The remaining columns show the clover term coefficient c_{SW}, the bare hopping parameter κ, the rotation parameter d_1 of the Fermilab action, and the values of the available source/sink Euclidean-time separations T in lattice units for the computed three-point correlators. The primes on the κ indicates that this is the value used in the simulation, as opposed to the physical (tuned) value, see Appendix C

a (fm)	am_l/am_s	c_{SW}	κ'_b	d_{1b}	κ'_c	d_{1c}	Sink times
≈ 0.15	0.0097/0.0484	1.567	0.0781	0.08354	0.1218	0.08825	10,11
≈ 0.12	0.02/0.05	1.525	0.0918	0.09439	0.1259	0.07539	12,13
	0.01/0.05	1.531	0.0901	0.09334	0.1254	0.07724	12,13
	0.007/0.05	1.530	0.0901	0.09332	0.1254	0.07731	12,13
	0.005/0.05	1.530	0.0901	0.09332	0.1254	0.07733	12,13
≈ 0.09	0.0124/0.031	1.473	0.0982	0.09681	0.1277	0.06420	17,18
	0.0062/0.031	1.476	0.0979	0.09677	0.1276	0.06482	17,18
	0.00465/0.031	1.477	0.0977	0.09671	0.1275	0.06523	17,18
	0.0031/0.031	1.478	0.0976	0.09669	0.1275	0.06537	17,18
	0.00155/0.031	1.4784	0.0976	0.09669	0.1275	0.06543	17,18
≈ 0.06	0.0072/0.018	1.4276	0.1048	0.09636	0.1295	0.05078	24,25
	0.0036/0.018	1.4287	0.1052	0.09631	0.1296	0.05055	24,25
	0.0025/0.018	1.4293	0.1052	0.09633	0.1296	0.05070	24,25
	0.0018/0.018	1.4298	0.1052	0.09635	0.1296	0.05076	24,25
≈ 0.045	0.0028/0.014	1.3943	0.1143	0.08864	0.1310	0.03842	32,33

The simulation parameters of all ensembles employed in this analysis are given in Table 1, while Fig. 1 provides a visual summary of the range of lattice spacings, sea-quark light-to-strange-mass ratios, and number of statistical samples.

In the light sector, we use the same value for the masses of the valence and sea quarks. The heavy quarks employ the clover action with the Fermilab interpretation, and since the regularization used for the light quarks has a different Dirac structure, we promote the staggered propagators to "naive" ones, so we can apply the standard Dirac spin algebra and combine them with the Wilson-like heavy quark propagators to construct heavy-light mesons [57]. The heavy quark masses are tuned so that the kinetic masses of the D_s and the B_s mesons are equal to their physical values (see Appendix C of Ref. [22]). In Table 2, we gather the parameters we used to calculate the heavy quark propagators for each ensemble. The simulation values chosen for the heavy quark masses to generate the meson correlators are close to, but not exactly the same as, our best-tuned values, which were determined *a posteriori*. Hence we apply a correction to the form factors to account for this slight mistuning which is described in Appendix C.

3.2 Correlation functions

The two- and three-point correlation functions are calculated using four sources, equally-spaced in time, except for the case of the coarsest ensemble that employs 24 sources. The sources are randomly shifted in space and time from one configuration to another in order to reduce correlations between successive gauge-field configurations within the same ensemble. A standard blocking analysis of the correlator data, ranging from block size 1 to block size 8, reveals that the autocorrelations in our ensembles are negligible and that the errors in the correlator points stay approximately constant as we increase the block size, in line with our previous analyses that employed the same gauge configurations, fermion formulations, and source set-up [22,39–42,53,58,59]. Therefore, we do not block the data in this work, and the correlators are processed through a single-elimination jackknife.

Two previous analyses with the asqtad ensembles [31,60] found that blocking the configurations by 4 or 8 was necessary in order to suppress autocorrelations. However, these analyses refer either to global observables (the topological susceptibility), or did not use the randomization procedure for the sources, which greatly reduces the autocorrelations in our data.

Given that one of our ensembles has a very fine lattice spacing $a \approx 0.045$ fm, one might be worried about the topology freezing and its effect in the final results of the form factors. We did not perform a topology freezing analysis in the asqtad ensembles, but we expect the behavior to be similar to that of the HISQ ensembles. Based on Refs. [61,62], we expect topology freezing to introduce a negligible bias in the chiral-continuum limit of the form factors.

The correlation functions described in the following two subsections contain the desired ground-state matrix elements, energies and form factors, but they also include contributions from excited states, which we must remove. For this purpose, we use two different kinds of interpolating operators per source: a local operator d and a smeared operator based on the Richardson $1S$ wave function [63,64], and therefore we fix the configuration to Coulomb gauge. For each meson, the

radius of the smearing operator is the same in physical units for all ensembles. We refer the reader to Ref. [64] for further details. The smeared operator increases the overlap with the ground state, allowing for a more precise determination of the lowest energy level and its overlap factors. The inclusion of a local operator gives us a useful handle on the excited states. Further, to quantify excited-state contributions and obtain robust estimates of the associated uncertainties, we use Bayesian constraints with Gaussian priors and fit functions that include varying numbers of excited states [65].

To implement the Bayesian constraints, we follow the procedure of Appendix B of Ref. [58]. We minimize the augmented χ^2_{aug}, as defined in Eq. (B3) of Ref. [58], but for the goodness of fit use the data-only χ^2 (evaluated at the minimum of χ^2_{aug}) and subtract the number of parameters from the number of data. Below we refer to this χ^2 and counting of degrees of freedom as the *deaugmented* χ^2/dof. In our experience, the p value calculated with a χ^2 and a number of dof as defined in Eq. (B5) of Ref. [58] is a good indicator of goodness of fit, but it has not been proven rigorously to follow a uniform distribution. When calculating p values, we further process the χ^2/dof ratio to take into account finite sample size [66].

3.3 Two-point functions

The D^* and B-meson two-point functions are needed to extract the overlap factors and the energy states, for these are required inputs for the ratio fits. The two-point functions are constructed using interpolating operators $\mathcal{O}_{Y_a}(\boldsymbol{p}, t)$, where $Y = \{B, D^*\}$ is the meson of interest, $a = \{d, 1S\}$ represents the smearing (point and Richardson), t is the time and $\boldsymbol{p}$ is the spatial momentum. These operators are constructed with the same quantum numbers as a pseudoscalar for $Y = B$ and a vector for $Y = D^*$. In terms of the interpolating operators, the two-point correlators are then

$$C^{\text{2pt}}_{Y_a \to Y_b}(\boldsymbol{p}, t) = \left\langle \mathcal{O}_{Y_b}(\boldsymbol{p}, t)\, \mathcal{O}^{\dagger}_{Y_a}(\boldsymbol{p}, 0) \right\rangle. \tag{30}$$

Inserting a complete set of states between the interpolating operators, we obtain the spectral decomposition:

$$C^{\text{2pt}}_{Y_a \to Y_b}(\boldsymbol{p}, t) = \sum_n s_n(t) \frac{\sqrt{Z_{Y_a,n}(\boldsymbol{p})\, Z_{Y_b,n}(\boldsymbol{p})}}{2E_n(\boldsymbol{p})} \times \left(e^{-E_n(\boldsymbol{p})t} + e^{-E_n(\boldsymbol{p})(L_t - t)} \right), \tag{31}$$

with $\sqrt{Z_{Y_{a,b}}(\boldsymbol{p})}$ the overlap factors, L_t the temporal extent of our lattice and $s_n(t)$ the extra sign that arises due to the presence of particles with the opposite parity in the staggered regularization for the fermions,

$$s_n(t) = \begin{cases} 1 & \text{Correct parity} \\ -(-1)^t & \text{Opposite parity} \end{cases}. \tag{32}$$

Most correlators are available in four different configurations according to the smearing of the source and the sink: d-d, d-$1S$, $1S$-d and $1S$-$1S$. In the case of the D^* meson, eight different momenta are available, namely, (0, 0, 0), (1, 0, 0), (1, 1, 0), (1, 1, 1), (2, 0, 0),(2, 1, 0), (2, 2, 0), (2, 2, 1), (3, 0, 0), and (4, 0, 0) in $2\pi/L$ units. Of these, only (0, 0, 0), (1, 0, 0) and (2, 0, 0) are used to calculate the form factors; the rest allow us to calculate the dispersion relation of the D^*. For (1, 0, 0) and (2, 0, 0), two different orientations of the momenta are considered, namely, parallel and perpendicular to the D^* polarization.

The outline of the analysis of the two-point functions, explained in detail in the following subsections, is as follows. First the zero-momentum correlators are fit using phenomenological guidance for the prior central values. For the ground state, we set a prior similar to the physical mass of the mesons, and the excited states differ by $\Delta E = 0.5$ GeV. The prior widths are large enough to accommodate significant departures from these assumptions. These choices for the central value and widths of the priors are such that they have no influence on the fit result for the ground states. In fact, we consider several variations of the energy priors to verify that their only function is to guarantee the stability of the fits without influencing the ground-state fit parameters. The results of the zero-momentum fits are used to construct priors for the dispersion-relation fits. In particular, the ground-state energies are expected to follow the continuum dispersion relation, and the overlap factors for the local operators should be approximately constant, barring, in both cases, discretization effects. Using data for a variety of momenta, we fit the ground-state energies to a dispersion-relation expression that includes discretization terms, see Eq. (36). The resulting fit is used to calculate a prior for the energy of the ground state of the two nonzero-momentum correlators.

3.3.1 Two-point function fits

For the two-point function fits, we employ the form

$$C^{\text{2pt}}_{Y_a \to Y_b}(t) = \sum_{i=0,1} (-1)^{i(t+1)} \mathcal{Z}_{i,a} \mathcal{Z}_{i,b} \left(e^{-E_i t} + e^{-E_i (L_t - t)} \right) + \sum_{i=2}^{2N-1} (-1)^{i(t+1)} \mathcal{Z}^2_{i,ab} \left(e^{-E_i t} + e^{-E_i (L_t - t)} \right), \tag{33}$$

where the state oscillates in time for odd i, but not for even i. This kind of fit is denoted as $N + N$, meaning we include N nonoscillating and N oscillating states. Both oscillating and nonoscillating excited states are fitted as the logarithm of the energy difference $\Delta E_i = E_i - E_{i-2}$ in order to avoid the collapse of two energy levels. In the higher states, we never interrelate energies of oscillating and nonoscillating states,

i.e., the fitted ΔE_i always refer to the difference between two states of the same type. The overlap factors in Eq. (31) are included in the fit function via

$$\mathcal{Z}_j = \sqrt{Z_j/2E_j}. \tag{34}$$

For the $\mathcal{Z}$ factors of the ground states, we also use a logarithm, forbidding the possibility $\mathcal{Z} \leq 0$.

We perform joint fits of all available correlators for a given combination of meson and momentum. That gives us three correlators corresponding to the d-d, $1S$-$1S$, and the crossed average between the d-$1S$ and the $1S$-d operators. In the cases where we distinguish between different orientations of the polarization of the D^* meson with respect to its momentum, the total number of correlators increases to six. The fitter uses the covariance matrix of the whole set of data, where the fit parameters are constrained with Gaussian priors. The prior central values for the energy levels in the fit functions for the zero-momentum correlators are guided by the experimental values for the meson mass in question and an empirical analysis of the data. The prior width of the physical (oscillating) ground state is chosen to be 140 MeV (520 MeV). In the fit functions for the nonzero-momentum correlators used for the dispersion relation, the ground-state energy prior central values are set equal to $\sqrt{M^2 + \boldsymbol{p}^2}$, where M is the posterior ground state energy from the fit to the corresponding zero-momentum correlator. The width of the prior is enlarged to encompass the expected discretization errors $O(\alpha_s a^2 p^2)$. The prior central value for the energy difference between two neighboring oscillating or nonoscillating states is taken to be 0.5 GeV. Their widths vary with the ensemble, but they are always larger than 0.2 GeV. The fit functions for the nonzero-momentum correlators employed in the three-point function analysis, namely momenta $2\pi(1, 0, 0)/L$ and $2\pi(2, 0, 0)/L$, use the dispersion-relation results as priors for the ground-state energy.

The energy levels are constrained to be the same across smearings, but the overlap factors are different, and they are represented with different parameters. For the ground states of the crossed average, we do not fit the $\mathcal{Z}_j$ amplitudes, but we impose the exact constraint

$$\mathcal{Z}^2_{d,1S} = \mathcal{Z}_{1S}\mathcal{Z}_d. \tag{35}$$

The $\mathcal{Z}_j$ amplitudes of the excited states of the crossed average are treated as separate fit parameters as they may describe a mix of excited states. Our $\mathcal{Z}_j$ amplitudes are also allowed to depend on the orientation of the momentum, when applicable, and Eq. (35) applies independently to each orientation. The priors for the $\mathcal{Z}_j$ factors of the ground states follow a log-normal distribution. Their central values are estimated following an empirical examination of the data, and their widths are large enough to accommodate significant departures from those original choices, roughly within one order of magnitude. In particular, the width of the physical ground state amplitude prior is set to 0.5 for all ensembles, and the width of the posterior is usually 20 times smaller. The prior is enlarged for the nonzero momentum correlators by a factor $\approx (1+2\alpha_s p^2)$. In the case of the oscillating ground state, the width of the amplitude is set to 1.2 for the zero momentum correlators and 2.0 for the nonzero momentum ones, with a typical posterior width of 0.5. In contrast, we use a Gaussian distribution for the excited-state priors. The width is fixed to be 3.5 for all ensembles, whereas the width of the resulting posteriors is typically an order of magnitude smaller. We test thoroughly that the fit results for the ground states are largely unaffected by the choice of priors and prior widths, as long as the fit remains stable.

The fit ranges are chosen following a systematic procedure: t_{Max} is chosen such that the correlator points for $t < t_{\text{Max}}$ have fractional errors smaller than $\approx$ 20–30%. In this way, the covariance matrix is not contaminated by excessive noise, but the value of t_{Max} becomes ensemble dependent. However, the correlator fits are generally insensitive to variations of t_{Max} within this constraint. In contrast, t_{Min} is chosen to have the same value in physical units for all ensembles and momenta. We do this because we expect the degree to which excited states influence the fit depends on their physical separation from the ground state. We apply the following four criteria to select the best t_{Min} value: (1) when including all ensembles and momenta, the p value must follow a sufficiently flat distribution for 15 ensembles, (2) the $2+2$ and the $3+3$ fits must agree on the nonoscillating ground-state energy and overlap factors, (3) the fit result must be stable under small variations of the fit range, and (4) the product $\mathcal{Z}_d\sqrt{2E} = Z_d$ for the overlap factor of the ground state should be approximately independent of the momentum, barring discretization effects. When these conditions are all fulfilled, we consider that the systematic errors due to the omission of still further excited states have been included in the statistical fit error. This usually leaves us with a small range of possible values for t_{Min}, we chose among those the one that complies best with all these conditions. The selected values are listed in Table 3. An example of the level of agreement that is reached between our $2+2$ and $3+3$ two-point correlator fits is shown in Table 4.

Previous experience [22,39,41,42,53,58], which also applies to this study, has shown that it is better to impose the four criteria introduced above on a set of fits, rather than choosing, on a case-by-case basis, the fit with the smallest χ^2/dof, the smallest error, or some other notion of "best" fit. The case-by-case approach amplifies meaningless statistical fluctuations, which can introduce problems in subsequent steps of the analysis (here, the chiral-continuum extrapolation). Figure 2 shows the stability of $1+1$, $2+2$, and $3+3$ fits on ensembles at four lattice spacings, denoting the common t_{Min}. It illustrates that we could have chosen smaller values of t_{Min} on some ensembles, if we had adopted ensemble-by-

Table 3 Fit ranges in physical units for the two-point function fits used in the analysis of the form factors. As the number of included states increases, t_{Min} is reduced to include information from the rapidly decaying excited states in the fit. For the coarsest ensembles and momentum $2\pi(4,0,0)/L$, the fit ranges do not yield enough points to perform a fit. In those cases the $2\pi(4,0,0)/L$ point is simply dropped

p $(2\pi/L)$	Two-point		
	$2+2$ t_{Min} (fm)	$3+3$ t_{Min} (fm)	t_{Max} (fm)
(0, 0, 0)	1.021	0.631	3.301
(1, 0, 0)			3.241
(2, 0, 0)			2.431
(1, 1, 0)	1.021	0.631	2.836
(1, 1, 1)			2.551
(2, 1, 0)			2.251
(2, 1, 1)			2.161
(2, 2, 0)			1.921
(2, 2, 1)			1.801
(3, 0, 0)			1.681
(4, 0, 0)			1.201

ensemble criteria. Hence, our common t_{Min} value is conservatively chosen.

3.3.2 The dispersion relation

The calculation of the dispersion relation serves two purposes: first, we can estimate a good prior for the two-point functions that enter in the analysis of the form factors; second, by checking the size of the deviations from the continuum dispersion relation, we can test whether the discretization errors due to the heavy quarks are under control. The dispersion relation which includes discretization effects can be written as

$$a^2E^2(\boldsymbol{p}) = (aM_1)^2 + \frac{M_1}{M_2}(a\boldsymbol{p})^2 + \frac{1}{4}\left[\frac{1}{(aM_2)^2} - \frac{aM_1}{(aM_4)^3}\right](a^2\boldsymbol{p}^2)^2 - \frac{aM_1w_4}{3}\sum_{i=1}^{3}(ap_i)^4 + O(p_i^6), \tag{36}$$

where M_1 is the rest mass, M_2 is the kinetic mass, and M_4 is a further mass-like quantity. A key observation of Ref. [49] is that the matching of the relativistic Wilson action via HQET or NRQCD to continuum QCD removes discretization effects that grow uncontrollably with aM. In Eq. (36) discretization effects are described by the coefficients of the $(a\boldsymbol{p})^n$ terms, parameterized by w_4, M_1, M_2 and M_4, for which explicit expressions are given in Ref. [49]. We tune the kinetic mass

Table 4 Results for the D^*-meson two-point correlator fits on the $a \approx 0.12$ fm, $m_l = 0.14m_s$ ensemble. We compare the nonoscillating ground-state energy and overlap factors for the $2+2$ and $3+3$ state fits. We include here all fits that distinguish between the different orientations of the momentum. In the analysis we use the $3+3$ state fit result

p $(2\pi/L)$	E_{D^*}		$\mathcal{Z}_{1S}^{\perp}$		$\mathcal{Z}_{d}^{\perp}$		$\mathcal{Z}_{1S}^{\parallel}$		$\mathcal{Z}_{d}^{\parallel}$		χ^2/dof	
	$2+2$	$3+3$	$2+2$	$3+3$	$2+2$	$3+3$	$2+2$	$3+3$	$2+2$	$3+3$	$2+2$	$3+3$
(1, 0, 0)	1.0837(21)	1.0841(19)	2.527(37)	2.531(39)	0.2777(43)	0.2767(46)	2.682(39)	2.674(42)	0.2912(45)	0.2897(48)	101.8/92	120.5/95
(2, 0, 0)	1.1854(49)	1.1890(50)	1.555(61)	1.579(63)	0.253(10)	0.259(10)	1.842(76)	1.881(83)	0.296(12)	0.302(12)	47.3/50	62.7/60

M_2 to match the experimentally observed mass, according to the nonrelativistic interpretation of the clover action [49].[1]

The leading $O(a^2)$ discretization effects are due to $M_1/M_2 \sim 1$. Expectations for this ratio can be inferred from perturbation theory for the quark masses [68] and by tracing contributions to the binding energy [69,70]. On this basis, we expect M_1/M_2 to be $1 + O(\alpha_s, (am_{0c})^2)$, and we would like to test whether the leading deviation from the continuum dispersion relation, $E^2 = \boldsymbol{p}^2 + M^2$, grows as $O(\alpha_s a^2 p^2)$. We can check whether our nonzero momentum fits show deviations of order $O(\alpha_s a^2 p^2)$ from the continuum dispersion relation, and we can also fit the energies from our correlator fits to Eq. (36), considering the coefficients in front of the powers of momenta as fit parameters. These results are used to guide the prior central values for the ground-state energies of the two-point correlators with nonzero momentum that are part of the three-point analysis which yields the form factors. In order to make this prior independent of the form-factor data, we exclude the $p = 2\pi(1,0,0)/L$, $2\pi(2,0,0)/L$ momenta from the dispersion-relation calculation. As explained in Sect. 3.3, data for different polarizations of the D^* meson are available only for $p = 2\pi(1,0,0)/L$ and $2\pi(2,0,0)/L$. Therefore, these are the only momenta for which we obtaine the form factors at nonzero recoil.

In Table 5, we show the results of our two-point correlator fits that enter in the dispersion-relation fit for a particular ensemble. There is good agreement between the $2+2$ and the $3+3$ state fits, indicating that the systematic error from the omission of higher states is negligible. Results for other ensembles show a similar behavior. Figure 3 compares the continuum dispersion relation with our data. The data points show small discretization errors, which tells us that, indeed, these errors are under control.

3.4 Three-point functions

With our previously defined interpolating operators, we can also construct three-point correlators by sandwiching a current between two meson states,

$$C^{J^\mu}_{X_a\to Y_b}(\boldsymbol{p},t) = \left\langle \mathcal{O}_{Y_b}(\boldsymbol{0},T)\, J^\mu(\boldsymbol{p},t)\, \mathcal{O}^\dagger_{X_a}(-\boldsymbol{p},0)\right\rangle. \tag{37}$$

Using the same notation as in Eq. (31), we can write the spectral decomposition of the three-point correlators for a particular source-sink separation T as

$$C^{J^\mu}_{X_a\to Y_b}(\boldsymbol{p},t) = \sum_n s_n(t)\, s_m(T-t)\sqrt{Z_{Y_b,n}(\boldsymbol{p})}\frac{e^{-E_n(\boldsymbol{p})t}}{2E_n(\boldsymbol{p})} \times \left\langle Y_b,n,\boldsymbol{p}|J^\mu|X_a,m,\boldsymbol{0}\right\rangle \times \sqrt{Z_{X_a,m}(\boldsymbol{0})}\frac{e^{-M_m(T-t)}}{2M_m}, \tag{38}$$

where we choose $t < T \ll L_t$, such that wraparound terms with $t \to L_t - t$ and $T - t \to L_t - (T-t)$ in the exponent are completely negligible, at most $\sim 10^{-16}$,

In our three-point functions, we always use a Richardson $1S$ smearing for the B meson, but the D^* meson operator is either $1S$-smeared or point d. This gives a variety of possibilities for constructing ratios of correlators. For x_f we use

$$x_f(\boldsymbol{p},t,T) = \frac{C^{V_j}_{D^*_{1S}\to D^*_a}(\boldsymbol{p}_{\perp,\parallel},t,T)}{C^{V_4}_{D^*_{1S}\to D^*_a}(\boldsymbol{p}_{\perp,\parallel},t,T)}, \tag{39}$$

where $a = d, 1S$ and the orientation of the momentum can be arbitrary, as long as it is the same for the correlator in the numerator and denominator. These combinations cancel the leading overlap factors and exponentials. The same cancelation can be achieved in X_V,

$$X_V(\boldsymbol{p},t,T) = \frac{C^{V_j}_{B_{1S}\to D^*_a}(\boldsymbol{p}_\perp,t,T)}{C^{A_j}_{B_{1S}\to D^*_a}(\boldsymbol{p}_\perp,t,T)}. \tag{40}$$

In these two ratios we can find the desired matrix element in the limit $t \gg 0$ and $T - t \gg 0$, with $t < T$. The double ratio and the other two single ratios can be expressed in the same way

$$X_0(\boldsymbol{p},t,T) = \frac{C^{A_4}_{B_{1S}\to D^*_a}(\boldsymbol{p}_\parallel,t,T)}{C^{A_j}_{B_{1S}\to D^*_a}(\boldsymbol{p}_\perp,t,T)}\sqrt{\frac{Z_{D^*,a}(p_\perp)}{Z_{D^*,a}(p_\parallel)}}, \tag{41}$$

$$X_1(\boldsymbol{p},t,T) = \frac{C^{A_j}_{B_{1S}\to D^*_a}(\boldsymbol{p}_\parallel,t,T)}{C^{A_j}_{B_{1S}\to D^*_a}(\boldsymbol{p}_\perp,t,T)}\sqrt{\frac{Z_{D^*,a}(p_\perp)}{Z_{D^*,a}(p_\parallel)}}, \tag{42}$$

$$R_{A_1}(\boldsymbol{p},t,T) = \frac{C^{A_j}_{B_{1S}\to D^*_a}(\boldsymbol{p}_\perp,t,T)\, C^{A_j}_{D^*_a\to B_{1S}}(\boldsymbol{p}_\perp,t,T)}{C^{V_4}_{D^*_a\to D^*_{1S}}(\boldsymbol{0},t,T)\, C^{V_4}_{B_{1S}\to B_{1S}}(\boldsymbol{0},t,T)} \times \frac{Z_{D^*,a}(p_\perp)}{\sqrt{Z_{D^*,a}(0)\, Z_{D^*,1S}(0)}} \times \frac{M^2_{D^*}}{E^2_{D^*}(\boldsymbol{p})}e^{-(E_{D^*}(\boldsymbol{p})-M_{D^*})T}, \tag{43}$$

but in this case the computed ratios depend on extra factors that must be removed before extracting the matrix elements. The overlap factors are removed per jackknife bin using the results of the two-point correlator fits. In this way we can propagate correlations from one fit to the other. The $M_{D^*}/E_{D^*} = 1/w$ factor in Eq. (43) is removed using the value of the recoil parameter, as extracted from Eq. (21) per jackknife bin.

[1] In the Fermilab formulation, one can adjust $M_1 = M_2$ by introducing an asymmetry parameter into the lattice action, but this is not necessary for valence quarks. Adjusting $M_4 = M_2$ requires a more improved action [67].

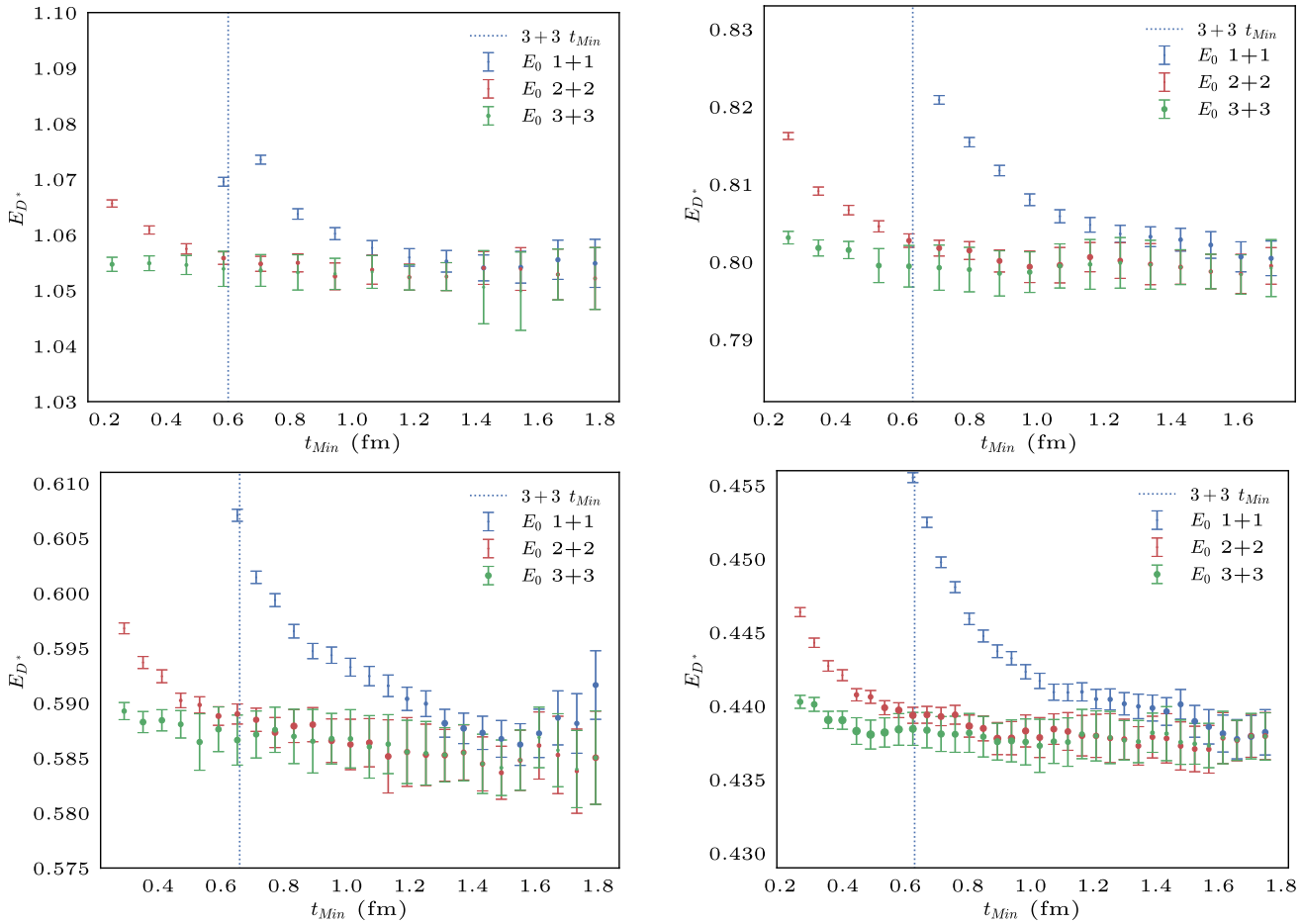

Fig. 2 Stability plots for the D^*-meson energy in lattice units with zero momentum (i.e., the mass) for four different ensembles with $m_l/m_s = 0.2$ and approximate lattice spacings 0.12 fm (upper left), 0.09 fm (upper right), 0.06 fm (lower left) and 0.045 fm (lower right). The size of each point is proportional to the p value, and visible point markers represent reasonable fits. The y axis represents the energy of the ground state in lattice units on each ensemble, with a range that roughly correspond to the same interval in physical units. The t_{Min} chosen for the analysis, which corresponds to $3+3$ correlator fits, is marked by a vertical dotted line. The point chosen is the closest one to the line. It is evident that on some ensembles smaller choices of t_{Min} are still in the stability region, which means that our common t_{Min} is conservatively chosen

Table 5 Results for the D^*-meson two-point correlator fits on the $a = 0.12$ fm, $m_l = 0.14m_s$ ensemble. We compare the nonoscillating ground-state energy and overlap factors for the $2+2$ and $3+3$ state fits. We include here all fits that did not distinguish between the different orientations of the momentum. Momentum $2\pi(4,0,0)/L$ did not have enough degrees of freedom left for the $3+3$ state fit, and hence it is not shown. In the analysis we use the $3+3$ state fit result

p $(2\pi/L)$	aE_{D^*}		$\mathcal{Z}_{1S}$		$\mathcal{Z}_d$		χ^2/dof	
	$2+2$	$3+3$	$2+2$	$3+3$	$2+2$	$3+3$	$2+2$	$3+3$
(0, 0, 0)	1.0419(17)	1.0405(17)	3.079(25)	3.052(24)	0.2865(26)	0.2835(30)	54.5/55	56.2/59
(1, 1, 0)	1.1205(31)	1.1215(26)	2.217(44)	2.215(44)	0.2786(56)	0.2782(51)	36.8/36	47.2/40
(1, 1, 1)	1.1569(36)	1.1576(30)	1.939(45)	1.928(43)	0.2769(61)	0.2760(53)	30.3/27	35.5/31
(2, 1, 0)	1.2191(43)	1.2202(35)	1.478(40)	1.477(41)	0.2682(64)	0.2683(55)	10.5/21	19.5/25
(2, 1, 1)	1.2513(47)	1.2536(38)	1.324(40)	1.323(39)	0.2666(67)	0.2663(56)	16.5/18	23.5/22
(2, 2, 0)	1.3089(65)	1.3100(49)	1.052(49)	1.050(44)	0.2611(68)	0.2607(57)	13.8/12	15.4/16
(2, 2, 1)	1.3378(72)	1.3402(58)	0.989(53)	0.959(44)	0.2598(68)	0.2596(58)	12.0/9	15.5/13
(3, 0, 0)	1.3168(79)	1.3168(79)	0.822(73)	0.83(12)	0.2523(68)	0.2513(59)	3.6/6	9.0/10

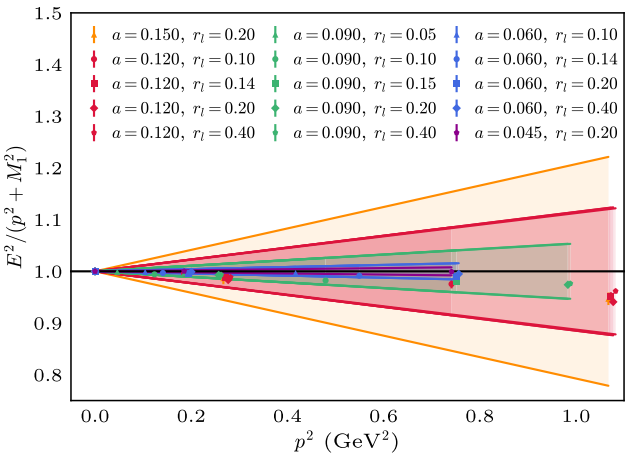

Fig. 3 Ratio between the fitted energies and the expected energies from the continuum dispersion relation. The color encodes the lattice spacing, whereas the different symbols encode the light-to-strange mass ratio. The cones show the expected size of the discretization errors per lattice spacing for our lattice formulation. Hence, each cone follows the equation $\alpha_s a^2 p^2$, where α_s is different for each lattice spacing, and it is calculated following Refs. [71,72]. The cones are color coded to match points with the same lattice spacing. All plotted points for each particular lattice spacing lie within their respective cones, which in turn means that the discretization errors are well within the expected size

The double ratio R_{A_1} deserves further comment. First, we reanalyze the zero-momentum correlators [22] using the criteria given above. That also implies the double ratio $R_{A_1}(\boldsymbol{p} = \boldsymbol{0})$ is constructed only for the $a = 1S$ smearing. Second, we did not generate three-point functions at nonzero momentum of the form $C^{A_j}_{D_a^* \to B_{1S}}(\boldsymbol{p}_\perp, t, T)$, so we use the time reversal operation $\mathcal{T}$ to obtain the missing correlator,

$$C^{A_j}_{B_{1S} \to D_a^*}(\boldsymbol{p}_\perp, t, T) \xrightarrow{\mathcal{T}} C^{A_j}_{D_a^* \to B_{1S}}(\boldsymbol{p}_\perp, T - t, T). \tag{44}$$

3.4.1 Three-point function fits

The three-point functions are also affected by the oscillating states introduced by the staggered regularization. So are the ratios constructed with such three-point correlators, but the dependence on the oscillating states is not as clean as in the case of the two-point functions. Our ratios do not show any noticeable oscillatory behavior in source and sink, but states that oscillate at both ends introduce a nonnegligible overall shift on the ratio central value that depends on the sink time T as $(-1)^T$. In order to remove this contribution, we smooth the data following Refs. [22,31,53,73], namely, we calculate the three-point correlators at two different values of the sink time T, and then we compute the following weighted average to suppress this unwanted shift in most ratios:

$$\bar{R}(t, T) \equiv \frac{1}{2}R(t, T) + \frac{1}{4}R(t, T+1) + \frac{1}{4}R(t+1, T+1),$$
$$\text{with} \quad R = X_0, X_1, X_V, x_f \text{ and } R_{A_1}(\boldsymbol{p} = 0). \tag{45}$$

The contribution of the oscillating shift is then greatly suppressed.

The double ratio at nonzero momentum, $R_{A_1}(\boldsymbol{p} \neq 0)$, requires the explicit removal of the sink-dependent exponentials in order to avoid bias,

$$\begin{aligned}\bar{R}_{A_1}(\boldsymbol{p} \neq 0, t, T) &\equiv \frac{1}{2}R_{A_1}(\boldsymbol{p}, t, T)e^{(E_{D^*}(\boldsymbol{p}) - M_{D^*})T} \\ &+ \frac{1}{4}R_{A_1}(\boldsymbol{p}, t, T+1)e^{(E_{D^*}(\boldsymbol{p}) - M_{D^*})(T+1)} \\ &+ \frac{1}{4}R_{A_1}(\boldsymbol{p}, t+1, T+1)e^{(E_{D^*}(\boldsymbol{p}) - M_{D^*})(T+1)}.\end{aligned} \tag{46}$$

These exponentials are removed using the energy and mass values coming from the two-point correlator fits per jackknife bin. The ratio averages defined in Eqs. (45) and (46) suppress the contributions from the unwanted oscillations to a fraction of the statistical errors. Therefore, we henceforth employ only the averaged ratios in our analysis and omit the bar for simplicity. The data are then processed through a single elimination jackknife, and the extra overlap factors and exponentials are removed by using the values obtained in the two-point correlator fits per jackknife bin. Then the ratios are fitted to the functional form:

$$\begin{aligned}R(\boldsymbol{p}, t, T) = K\Big(1 &+ A_1 e^{-\Delta E^1_X t} + A_2 e^{-\Delta E^2_X t} \\ &+ B_1 e^{-\Delta E^1_Y (T-t)} + B_2 e^{-\Delta E^2_Y (T-t)}\Big),\end{aligned} \tag{47}$$

where K is the matrix element we want to extract, and the extra terms take into account the presence of excited states, assuming their contribution is small. The labels X, Y represent mesons at source and sink, respectively, and $\Delta E^j_{X,Y}$ represents the energy difference between the ground state and the j^{th} excited state. The second excited states at source and sink (included in the A_2 and B_2 terms) are necessary to remove systematic errors due to unaccounted excited states. In order to check this point, we computed the ratio x_f with different polarizations of the D^* meson. We expect the extracted matrix elements from different polarizations to agree, except for discretization effects that should be reduced as the lattice spacing decreases. Nonetheless, our results show a difference between the analysis with a single excited state at source and sink and the analysis with two excited states at each end. The addition of extra excited states not only increases the error, as expected, but also brings the central values calculated with different polarizations closer. Overall there is a large reduction in the difference between the cases with polarization parallel and perpendicular to the momentum. This behavior depends only mildly on the lattice spacing, as can be checked in Fig. 4.

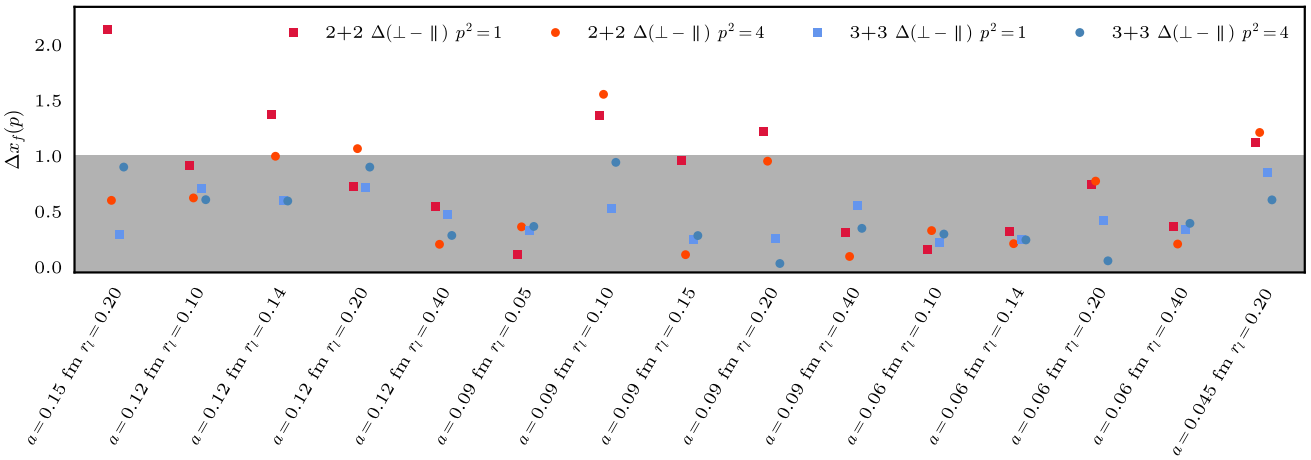

Fig. 4 Difference in sigmas between x_f calculated with parallel and perpendicular polarizations, for momenta $p^2 = 1, 4$ and one $(2+2)$ and two $(3+3)$ excited states at source and sink in the ratio. The gray band corresponds to differences $\leq 1\sigma$. It is clear that an increase in the number of excited states reduces the difference between polarizations to 1σ or less, whereas with a single excited state the difference can reach as high as 2σ in some cases. Since we do not have any other way to keep in check the systematics due to excited states for other ratios, we take a conservative approach and use the $3+3$ state fits for the rest of the analysis

The two available sets of correlators with smearing operators $a = d$, $1S$ are fit simultaneously, so that they share ΔE and K, but each smearing has its own $A_{1,2}$ and $B_{1,2}$. We employ a loose prior for K with a central value roughly set by the ratios at $\approx T/2$, where the excited states are more suppressed, and with a width large enough to accommodate significant variations: in the case of the double ratio R_{A_1} the width of the prior is set to 0.1, whereas the other ratios use 0.05. In all cases, the width of the prior encompasses all available correlator points for the single ratios. Except for the double ratio, the excited states at source and sink carry different signs, hence we ensure the prior covers the central value of the matrix element. Typically, the posterior is almost an order of magnitude narrower that the prior, although in the least precise cases th the posterior width is $\approx 60\%$ of the prior. The priors for the ΔE of the first (second) excited states are taken from the two-point function fits with $3+3$ states, but we increase the error by a factor of three (eight) to allow ΔE to differ from the two-point correlator values. This increase takes into account the fact that the t_{Min} in the ratio fits is much smaller than in the two-point function fits, and the excited-state pattern might be different as well. The priors for $A_{1,2}$ and $B_{1,2}$ are taken to be 0(2) and 0(1) in the smeared and point source cases respectively.[2] As stated above, our priors are conservative enough that significant variations of their central values and/or widths do not result in a relevant change in the posterior for the matrix element.

[2] We observe that the local operators have a smaller overlap with the excited states, and we reduce the width of their corresponding priors for $A_{1,2}$ and $B_{1,2}$ accordingly.

The fit ranges are chosen following criteria similar to the two-point function case. We use the same value of t_{Min} and t_{Max} in physical units for all ensembles and all ratios, except for the double ratio, where we use $t_{\text{Max}} = T - t_{\text{Min}}$ to account for the fact that the states on source and sink are exactly the same. In this case, we take the same t_{Min} as for the rest of the ratios. We choose the fits that show stability over small variations of the fit range and result in a reasonably flat distribution for the p values.

3.5 Calculation of the recoil parameter

As in previous work [52,53], we use the ratio x_f to define the recoil parameter, following Eq. (21). The disadvantage of this method is that it introduces systematic errors due to the renormalization of the currents. One could also use the continuum dispersion relation to define the recoil parameter:

$$w = \sqrt{\frac{M_{D^*}^2 + \boldsymbol{p}^2}{M_{D^*}^2}}, \tag{48}$$

where $\boldsymbol{p}$ is the three-momentum of the D^* meson, and the mass is either M_1 or M_2. The different choices for the different definition of the mass are expected to result in slightly different discretization errors that are resolved in our chiral-continuum extrapolation, so the choice should not affect the final results. As shown in Fig. 5, the error in the x_f method encompasses the differences in the rest- and the kinetic-mass versions of Eq. (48). In this work, we take a conservative approach and define w via Eq. (21), but we note that all choices lead to results for the form factors, $|V_{cb}|$, and $R(D^*)$ that are compatible within errors.

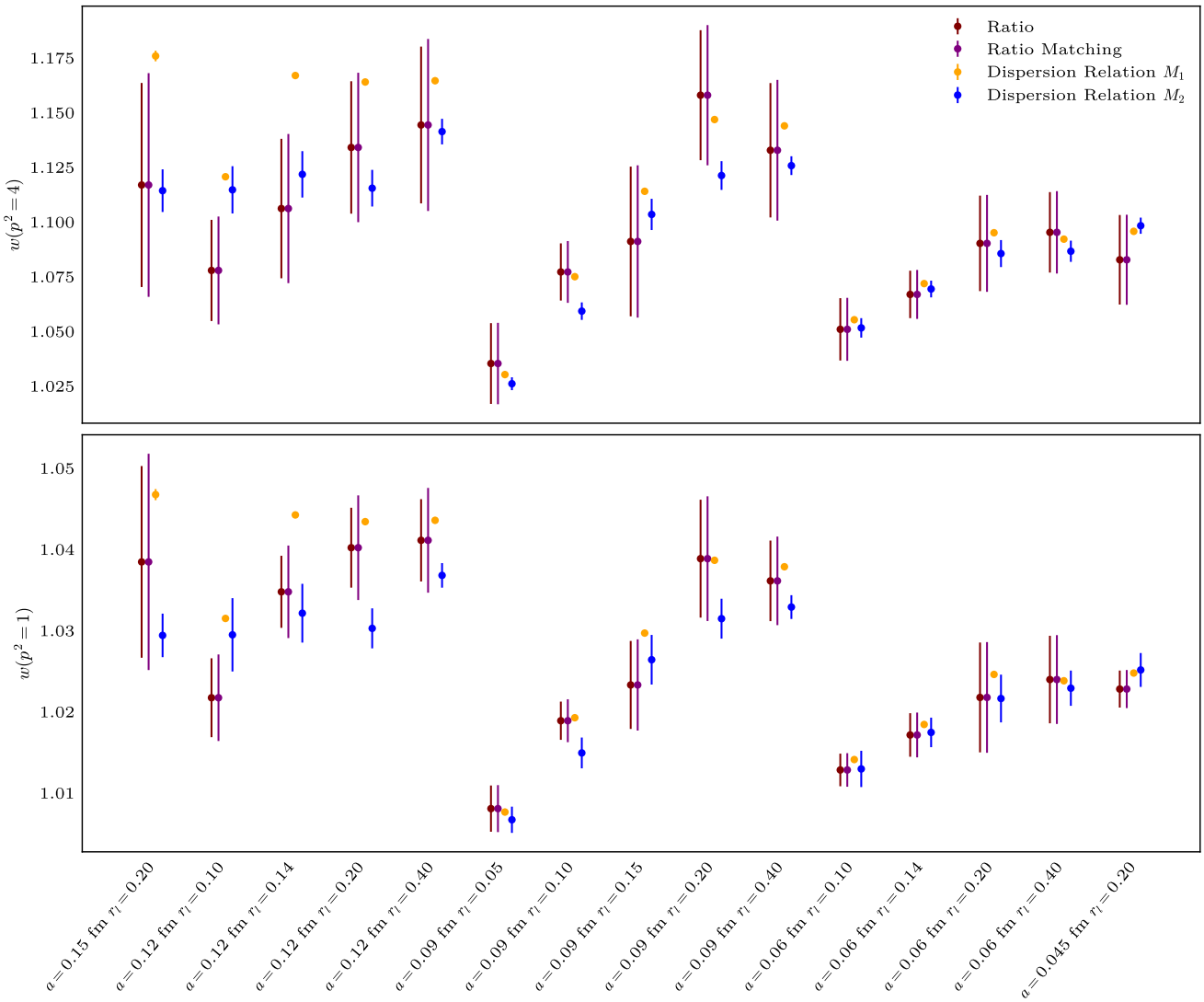

Fig. 5 Comparison of different definitions of the recoil parameter. The horizontal axis labels the ensemble, whereas the vertical axis shows the recoil value w, calculated from the different definitions. The ratio method is shown before ("Ratio") and after ("Ratio matching") including the matching errors, whereas the points labeled as "Dispersion relation $M_{1,2}$" use Eq. (48) and the fitted rest (M_1) or kinetic (M_2) mass of the D^* meson. The three-momentum is set to $p = 2\pi(1, 0, 0)/L$ [$p = 2\pi(2, 0, 0)/L$] in the lower (upper) panel

3.6 Current renormalization and blinding

As outlined in Sect. 2.2, the ratios described in Eqs. (39)–(43) are constructed in such a way that the flavor-diagonal renormalization factors $Z_{V_{cc}^4} Z_{V_{bb}^4}$ from Eq. (14) cancel out. Hence, what remains is only the computation of the different matching factors ρ_X that enter in the ratios. These factors can be calculated using perturbation theory, but the calculation becomes cumbersome for $w > 1$. In this article, we use the approximation $am_{2c} \to 0$, where am_{2c} is the charm kinetic mass, which removes the dependence on w, because a light quark cannot modify the dynamics of a heavy quark in the heavy quark limit (see Appendix C and Refs. [50,74]). Then we incorporate errors coming from these approximations. The w dependence introduces an error proportional to $w - 1$, and the $am_{2c} \to 0$ approximation increases the error by $O(\alpha_s am_{2c})$.

The calculation of the axial matching factor $\rho_{A_j}(1)$ then follows exactly the procedure of Ref. [22]. The other ratios require further calculations that are detailed in Appendix B. The resulting values for all matching factors are gathered in Table 6, where the errors shown are from the VEGAS integration. The errors associated with the approximations we make to obtain those factors are discussed in Appendix B. They are added to those in Table 6 before carrying out the chiral-continuum extrapolation.

We also calculate the matching factors for the ensembles involved in the heavy-quark (HQ) mistuning corrections.

Table 6 Matching factors calculated at one loop in perturbation theory for all ensembles and form factors. The errors shown in the table do not include the systematic errors coming from the approximations employed in the calculation of the matching factors. They are nonetheless included in the chiral-continuum extrapolation. These errors and more details on the calculation are described in Appendix B. The last two columns include the values of α_V and $m_{2c}a$ used in Appendix B to estimate the matching errors

a (fm)	am_l/am_s	ρ_{A_j}	ρ_{A_4}/ρ_{A_j}	ρ_{V_j}/ρ_{A_j}	α_V	$m_{2c}a$
≈ 0.15	0.0097/0.0484	0.990864(48)	0.945800(10)	0.8835513(23)	0.3589	0.6000
≈ 0.12	0.02/0.05	0.994199(25)	0.961710(9)	0.9133327(20)	0.3047	0.4259
	0.01/0.05	0.988811(22)	0.957865(9)	0.9096509(22)	0.3108	0.4460
	0.007/0.05	0.993665(41)	0.960041(9)	0.9101966(21)	0.3102	0.4469
	0.005/0.05	0.993640(23)	0.960038(9)	0.9101935(21)	0.3102	0.4471
≈ 0.09	0.0124/0.031	0.998292(13)	0.971160(5)	0.9320012(15)	0.2582	0.3269
	0.0062/0.031	0.998077(13)	0.970701(6)	0.9310817(16)	0.2607	0.3317
	0.00465/0.031	0.997956(18)	0.970559(6)	0.9308184(16)	0.2611	0.3350
	0.0031/0.031	0.997901(12)	0.970405(7)	0.9305130(16)	0.2619	0.3361
	0.00155/0.031	0.997889(16)	0.970353(7)	0.9303983(16)	0.2623	0.3365
≈ 0.06	0.0072/0.018	1.001973(7)	0.978328(4)	0.9460531(12)	0.2238	0.2343
	0.0036/0.018	1.002050(11)	0.978408(4)	0.9461304(11)	0.2245	0.2327
	0.0025/0.018	1.002013(13)	0.978356(5)	0.9460189(12)	0.2249	0.2336
	0.0018/0.018	1.002015(6)	0.978309(4)	0.9459167(12)	0.2253	0.2339
≈ 0.045	0.0028/0.014	1.005588(8)	0.984902(2)	0.9578549(7)	0.2013	0.1647

This is a departure from Ref. [53], where the matching factors are applied after the HQ-mistuning correction. In this way, we completely separate the HQ-mistuning corrections from the matching errors.

During the analysis, we blinded the form-factor data by multiplying the matching factor ρ_{A_j} by an unknown, random number close to 1. The ratios ρ_{A_4}/ρ_{A_j} and ρ_{V_j}/ρ_{A_j} are left unchanged. Via the correlator ratios and Eqs. (26)–(29), all form-factor values on all ensembles are thus multiplied by a common factor. All stages of the analysis were tested either through independent fits performed by two co-authors or by using independent methods or codes. Only after the full analysis was complete, including construction of the systematic error budget (Sect. 4) and z expansion (Sect. 5), the blinding factor was removed from ρ_{A_j}, and the analysis scripts were rerun to extract the unblinded results for the form factors.

3.7 Heavy-quark-mass adjustment

The bare masses for the b and c quarks are tuned such that the kinetic masses of the B_s and D_s meson on each ensemble are equal to their physical values. Nonetheless, the tuning procedure has errors that must be taken into account. The procedure is explained in Appendix C and outlined here. As we generate configurations, values for the heavy-quark masses that give approximately the correct meson masses can be estimated. These initial values are employed to compute the two- and three-point functions that we analyze in this work. At the end of the data generation, the much larger statistical sample allows for a more precise determination of the b and c quark masses by using the procedure detailed in Ref. [22]. We must then correct for this mismatch.

The correction is calculated nonperturbatively by studying the effect of a varying heavy-quark mass in a single ensemble on all correlation functions. Since the heavy-quark mismatch is small, a linear fit in $1/m_Q$ for each flavor $Q = c, b$ and form factor is usually sufficient. Once the functions that describe the evolution of the form factors with the quark masses are known, we can apply the correction to all other ensembles. Table 20 in Appendix C gathers all our adjustments. The error in the correction comes from the error in the heavy-quark-mass tuning procedure and from the error of the linear fit to find the heavy-quark mass dependence.

3.8 Chiral-continuum extrapolation

After applying the renormalization factors and the heavy-quark-mistuning corrections as described in previous subsections, the resulting form factors still need to be corrected for the fact that they are calculated at nonzero lattice spacing and nonphysical values for the light-quark masses. An extrapolation to both the continuum limit and the physical value of the light-quark masses is thus necessary to extract values that can be used in a physical calculation. The extrapolation should be based on an appropriate effective field theory (EFT) description of lattice QCD. The relevant EFT for the calculation at hand is rooted staggered chiral perturbation theory (rSχPT), which describes how the form factors behave as the lattice spacing and the light-quark masses approach the desired limits, extended to include heavy-light observables [75]. The unquenched MILC configurations generated with $2+1$ flavors of improved staggered fermions make use of the fourth-root procedure for eliminating the unwanted four-fold degeneracy of staggered quarks. At nonzero lattice spacing, this procedure has small violations of unitarity [76–80] and

locality [81]. Nevertheless, a careful treatment of the continuum limit, in which all assumptions are made explicit, argues that lattice QCD with rooted staggered quarks reproduces the desired local theory of QCD as $a \to 0$ [82,83]. When coupled with other analytical and numerical evidence (see Refs. [84–86] for reviews), this gives us confidence that the rooting procedure is indeed correct in the continuum limit. We then use the following functions obtained in SU(3) rSχPT to lowest nontrivial order in the heavy-quark expansion to fit the different form factors:

$$h_Y(a,m,m_s,w) = \left(K_Y + \frac{\chi_Y(\Lambda_\chi)}{m_c^{k_Y}} + f_Y^w + f_Y^{\mathrm{NLO}} + f_Y^{\mathrm{NNLO}}\right) \times \left(1 + f_Y^{\mathrm{HQ}}\right), \tag{49}$$

where $Y = A_1, A_2, A_3$, and V; $K_{A_2} = 0$ but $K_Y = 1$ otherwise; $k_{A_1} = 2$ but $k_Y = 1$ otherwise. These expressions contain the correct dependence in χPT on the light- and strange-quark masses, the lattice spacing, and the recoil parameter w at next-to-leading order (NLO). This result expands on the one in Ref. [87] by adding the missing recoil dependence in the relevant places. The terms

$$f_Y^{\mathrm{NLO}} = \frac{g_{D^*D\pi}^2}{48\pi^2 f_\pi^2 r_1^2}\,\mathrm{logs}^Y_{\mathrm{SU(3)}}(a,m,m_s,w,\Lambda_\chi) + c_{m_1,Y}\,x_l + c_{a_1,Y}\,x_{a^2}, \tag{50}$$

introduce nonanalytic dependence on the light and the strange-quark masses through the chiral logarithms $\mathrm{logs}^Y_{\mathrm{SU(3)}}$. Those terms also include the leading taste-breaking discretization effects from the light-quark sector. The explicit expression of the logarithms for each form factor is given in Appendix A and includes a dependence on the recoil parameter w. The coefficient of the chiral logarithms comes from χPT and it is known, but the current determinations of the coupling $g_{D^*D\pi}$ are not very accurate, hence we fit the coupling with a Gaussian prior 0.53 ± 0.08, compatible with experimental data [88–90] and lattice-QCD results [91–96]. We fix the pion decay constant appearing in the chiral logs in Eq. (50), and elsewhere in the fit function to the three-flavor FLAG 2019 average with the error increased by the estimated 0.7% charm sea-quark contribution $f_\pi = 130.2 \pm 1.2$ MeV [97].

The other terms in Eq. (50) introduce analytic NLO corrections in the light-quark masses through $x_l = 2B_0 m/(8\pi^2 f_\pi^2)$ and in the lattice spacing through $x_{a^2} = [a/(4\pi f_\pi r_1^2)]^2$, where B_0 is the low-energy-constant (LEC) of χPT that relates the light- and strange-quark masses with the meson masses. The value of B_0 for each lattice spacing is the same as in the earlier analysis at zero recoil [22], which uses exactly the same ensembles, and is given in Appendix A. We take into account truncation errors by including the term

$$f_Y^{\mathrm{NNLO}} = c_{c,Y}\,x_l x_{a^2} + c_{m_2,Y}\,x_l^2 + c_{a_2,Y}\,x_{a^2}^2, \tag{51}$$

which describes the dependence on the light-quark masses and the lattice spacing a at next-to-next-to-leading order (NNLO), not including logarithmic terms. According to χPT power-counting, these analytical terms are expected to have coefficients of $O(1)$, so we take them as fit parameters with priors 0 ± 1. We don't include analytical terms in the strange-quark mass because we do not have data at different values of m_s and nonzero recoil, and our $h_{A_1}(1)$ result using only zero-recoil data agrees within errors with our previous result from Ref. [22]. Also, χPT predicts a much milder dependence on m_s than on the light-quark masses.

We allow a simple NLO analytical dependence on w to describe the behavior in our small-recoil range through the term

$$f_Y^w = -\rho_Y^2(w-1) + \kappa_Y(w-1)^2, \tag{52}$$

where the fit parameters ρ_Y and κ_Y are related to the slope and curvature of the form factor h_Y respectively. We can reasonably expect the slope of the form factors to be roughly 1, but in order to accommodate substantial deviations from this value, we set the prior of ρ_Y to 1 ± 2. The priors of κ_Y are chosen to be 0 ± 3, and the posteriors are compatible with zero within a fraction of a sigma.

The constant term $\chi_Y(\Lambda_\chi)$ in Eq. (49) is a LEC of the chiral effective theory, and it is suppressed for h_{A_1} by a factor of $1/m_c^2$ due to Luke's theorem [98], whereas the other form factors receive contributions of order $O(1/m_c)$ in the heavy-quark power counting. The dependence of this LEC on the chiral scale Λ_χ cancels against the dependence of the nonanalytical terms in Eq. (50). We set the prior of this LEC to 0 ± 1, except for h_{A_1} where we use 0.0 ± 0.2 to reflect the suppression due to Luke's theorem.

The last term in Eq. (49) accounts for the heavy quark discretization errors,

$$f_Y^{\mathrm{HQ}} = \beta_Y^{\alpha_s a}\alpha_s a\Lambda_{\mathrm{QCD}} + \beta_Y^{a^2} a^2\Lambda_{\mathrm{QCD}}^2 + \beta_Y^{a^3} a^3\Lambda_{\mathrm{QCD}}^3, \tag{53}$$

with β_Y^p the coefficient of the term of order $O(p)$ corresponding to the form factor h_Y, and $\Lambda_{\mathrm{QCD}} = 0.6$ GeV for normalization purposes. In previous articles,[3] we have employed the universal functions described in Refs. [49,67]. But in those cases there was only one heavy meson. Here we need to deal with the B and the D^*, and our data are not accurate enough to distinguish the different terms described by the universal functions. In order to avoid terms that mimic the effect of others, we consider it a better strategy to implement generic discretization error terms, which would account for the same dependence as the universal functions. A side effect of this

[3] See, for instance, Ref. [64].

approach is that the heavy- and NLO light-quark discretization effects become mixed together through terms with the same dependence on the lattice spacing. To avoid this, we drop the $O(a^2)$ term from Eq. (53), which has the same dependence as the $O(a^2)$ term already present in Eq. (50), and we enlarge the prior of the latter assuming that both corrections are independent, i.e., with a quadrature sum. The priors for the β_Y^p coefficients are set to 0 ± 1, but the $O(a^2)$ coefficients of Eqs. (50) and (53) have different normalizations. For this reason, the final prior for $c_{a_2,Y}$ becomes 0.0 ± 6.1. This approach does not allow us to distinguish cleanly the origin of the discretization errors, but it accounts for the correct dependence and size of discretization errors. However, absorbing the $O(a^2)$ term from Eq. (53) in Eq. (50) may have further effects, since the correction in Eq. (53) is applied to the chiral-continuum fit function in Eq. (49). It is possible that this procedure does not account for discretization effects in the shapes of the form factors, which would give rise to higher order terms of the form $a^2(w-1)$ and $a^2(w-1)^2$. We test for such effects by performing two alternate chiral-continuum fits. In the first, we add the terms $x_{a^2}(w-1)$ and $x_{a^2}(w-1)^2$ to Eq. (52), where the priors for the coefficients of these terms are chosen as 0(1). We find that the results of this fit differ by at most 0.1σ in their central values from our base fit, while the statistical fit uncertainty is unchanged. In the second variation, we keep the $O(a^2)$ term in Eq. (53). In this case, the central values are consistent with those of our base fit within 0.25σ, again with an unchanged uncertainty. Hence, we conclude that such discretization effects are already accounted for in our base fit.

Since each ensemble is statistically independent from the others, there are no correlations among them. On the other hand, we keep track of correlations both between different form factors within the same ensemble and within the same form factor calculated at different momenta, by combining the jackknife data of all form factors into a large, block-diagonal dataset. Our large statistics allow us to resolve the full covariance matrix without resorting to thinning procedures or singular-value-decomposition cuts on its eigenvalues. Nonetheless, we use the shrinkage procedure described in Refs. [99–102] to ensure the small eigenvalues of the covariance matrix have the correct behavior, and we find that our results do not change with respect to the analysis without shrinkage.

The systematic errors coming from the heavy-quark mistuning corrections and those introduced by the matching factors are built into our chiral-continuum extrapolation by constructing the combined covariance matrix,

$$C_{ij} = C_{ij}^{\text{stat}} + \delta_i^{(\rho)}\delta_j^{(\rho)} + \delta_i^{(\kappa)}\delta_j^{(\kappa)}, \tag{54}$$

where the first term includes the statistical covariance, and the second and the third ones account for the matching factor and heavy-quark-mass mistuning correction errors respectively. The i, j indices run over all form factors, ensembles, and momenta. With $\delta_i^{(\rho,\kappa)}$ we represent either the shift in the i^{th} datum due to a correction the heavy-quark mass (κ) or the propagated error of the form factor from the errors in the matching factors (ρ) as calculated in Eqs. (B.49a)–(B.49d). As a result, the systematic errors introduce new correlations between all data points. In fact, Eq. (54) assumes the worst case scenario that the matching systematic errors and the errors coming from the heavy-quark mistuning are 100% anticorrelated.

The extrapolation results for the four form factors are shown in Fig. 6. As one can see, h_{A_1}, which is protected by Luke's theorem, receives small corrections from 1 at $w = 1$. The other form factors do not enjoy this privilege, and for them the plots show large corrections from the HQET limit. Figure 6 also shows the result of the previous Fermilab-MILC calculation at zero recoil $w = 1$ for comparison [22]. The agreement is good, although the errors have increased, mainly due to more conservative choices in this work, which stem from the data at nonzero recoil requiring an extra excited state at source and sink in the ratio calculations, which resulted in larger errors. For consistency, we employed the same approach at zero recoil as well. The deaugmented χ^2/dof of the chiral-continuum extrapolation is 85.2/95.

4 Systematic errors

This section provides specific information on our estimates of every source of systematic error in the determination of the form factors h_X. Even though only h_{A_1} contributes to the decay amplitude at zero recoil, all form factors are nonzero at $w = 1$, and their errors need not be suppressed at small recoil. Even if the errors of h_V, h_{A_2}, and h_{A_3} become large, however, their contribution to the decay amplitude and, hence, the resulting uncertainty in the decay amplitude is still suppressed at small recoil.

Some general features of the uncertainties in the form factors can be understood via HQET. The form factor h_{A_1} is protected by Luke's theorem [98] and, indeed, we find HQET corrections of a few percent. The form factors h_V and h_{A_3}, which are not protected by Luke's theorem, receive HQET corrections at the $\sim 30\%$ level. The form factor h_{A_2} starts in HQET with terms of order α_s and $1/m_c$, which is roughly consistent with our data, $h_{A_2} \sim -\frac{1}{2}$. Figure 7 shows the error budget for the different form factors in the continuum as a function of the recoil parameter. The relative uncertainty in each form factor follows the same pattern as the HQET corrections: small for h_{A_1}, moderate for h_V and h_{A_3}, and large for h_{A_2}. In the last case, the relative uncertainty is large, because the overall value of h_{A_2} is smaller than the others.

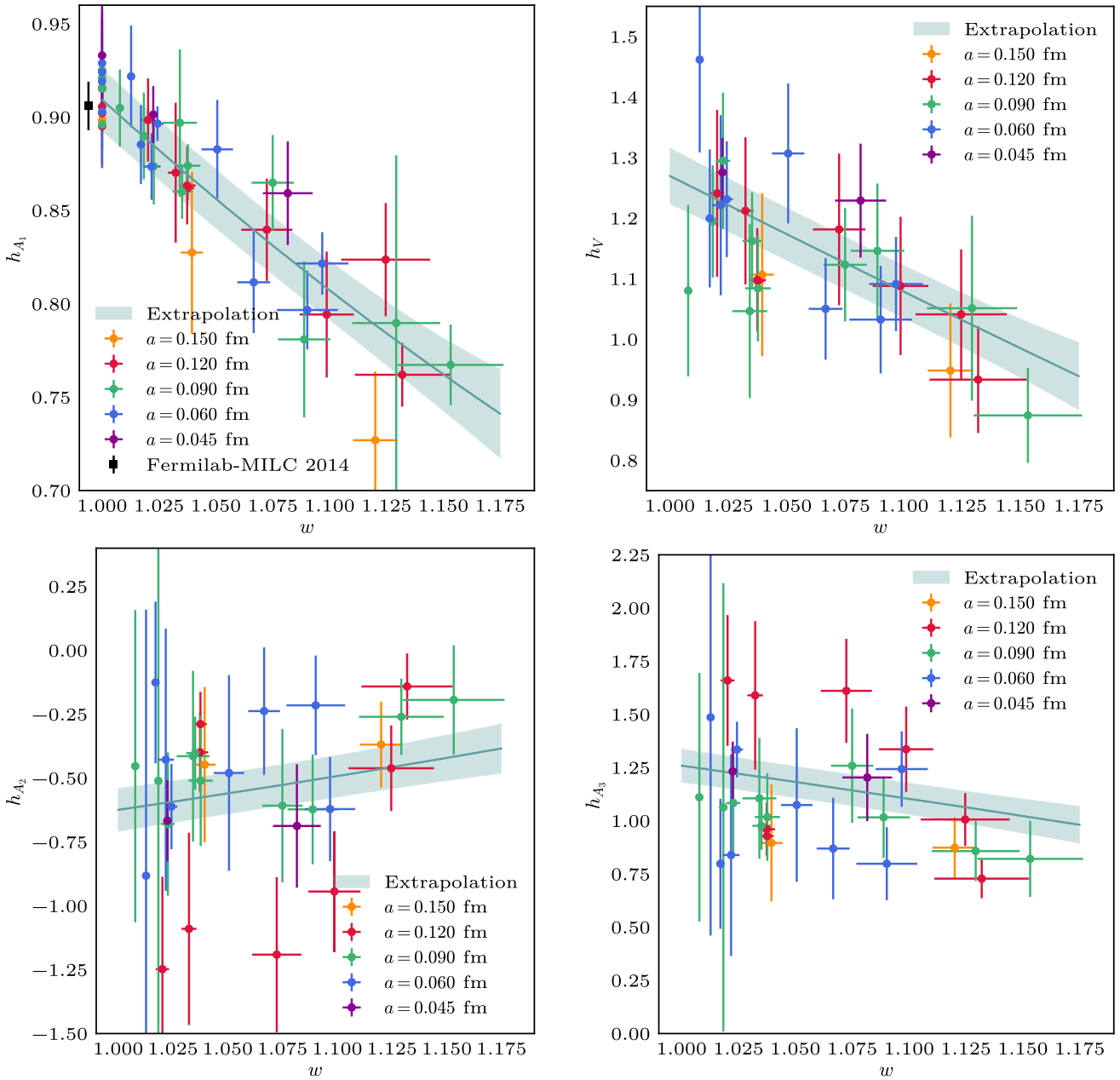

Fig. 6 Chiral-continuum extrapolation for the form factors h_{A_1} (top left), h_V (top right), h_{A_2} (bottom left), and h_{A_3} (bottom right). The color encodes different lattice spacings, whereas the band shows the result of the fit. The upper-left plot for h_{A_1} also shows the zero-recoil result from Ref. [22]. Note that correlations in w and between form factors play an important role in controlling the final uncertainty

Our chiral-continuum extrapolation ansatz to NNLO incorporates errors from statistics, choices in the chiral-continuum extrapolation, discretization effects, $O(am_c\alpha_s)$ matching errors, and heavy-quark parameter mistuning. Thus, they are all entangled in the fit, and it is not straightforward to extract each particular contribution. In addition, our treatment of the heavy-quark discretization errors includes a term identical to one of gluon and light-quark discretization errors. We can, however, roughly estimate each contribution by making modifications to the fit. In this spirit, we define the statistical contribution to the error as the error obtained in a NLO fit without mistuning correction or matching-factor errors included. We have a specific way to deal with the matching factors, which is explained below. The contribution coming from the chiral-continuum extrapolations is estimated by comparing the fit errors with and without NNLO terms.

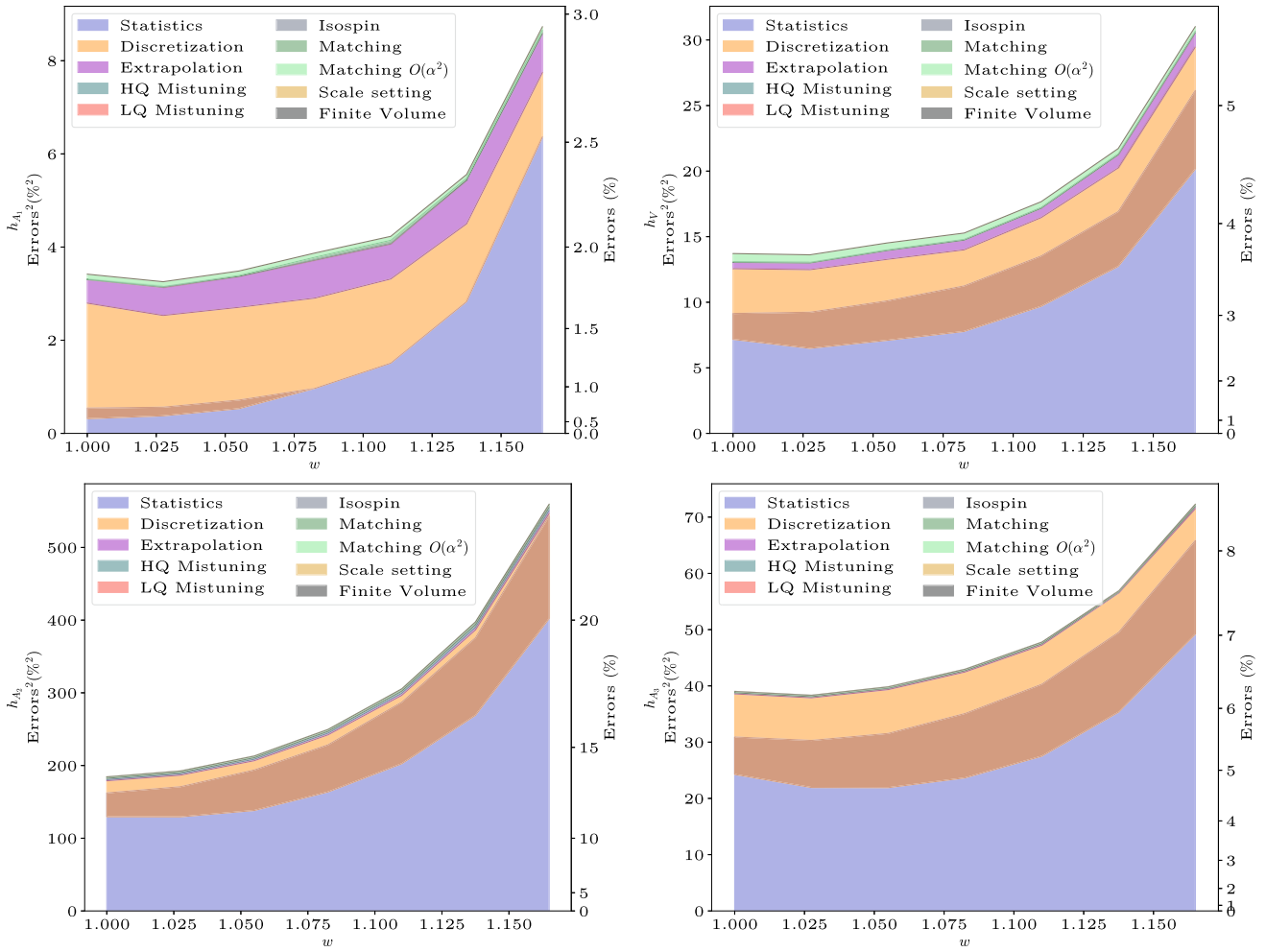

Fig. 7 Contributions to the total error of the form factors h_{A_1} (top left), h_V (top right), h_{A_2} (bottom left), and h_{A_3} (bottom right) as a function of the recoil parameter w. The two largest contribution come from statistics, in blue, and quark discretization effects, in orange. These two contributions overlap in the brown band in the plot, because of the common term of order a^2 in Eqs. (50) and (53). The remaining contributions do not overlap. Note the differences in vertical scales

There are more contributions to the final error that have been taken into account: light-quark mass mistuning, scale setting, isospin effects, and finite-volume effects. The final error is taken to be the quadrature sum of these uncertainties with that of the chiral-continuum extrapolation error, which (again) includes statistical, chiral-continuum extrapolation, discretization, heavy-quark mistuning, and matching errors, as shown in Table 7. In the rest of this section, we discuss each source of uncertainty one by one, explaining how they enter this error budget.

4.1 Statistics and stability of the correlator fits

In principle, the determination of masses, energies, and form factors depends on choices made in fitting the two- and three-point correlation functions, but we argue that the associated uncertainties are encompassed in the statistical component of the first line of Table 7. We have analyzed the two-point functions with both 2 + 2 and 3 + 3 states in the fit. Only when the two results agree within statistical errors do we select a particular fitting range. In this way, the influence of excited states is reduced below the statistical uncertainty of the B masses and D^* energies. For the three-point-function ratios, we find that excited states play a more important role, as can be seen for the example of x_f in Fig. 4. When fitting the three-point functions, we therefore include extra states at the source and sink in order to control this potential source of systematic error.

Bias can arise from the choice of fitting ranges. To avoid the problems that can come from choosing different fitting ranges for different ensembles, we impose the same t_{Min} in physical units for all two-point correlator fits. Our t_{Max} is chosen differently and varies from ensemble to ensemble, but the impact of a different t_{Max} is much smaller, because

Table 7 Error budget for all form factors at $w = 1.11$. The first row shows the combined error coming from our chiral-continuum fit, which encompasses the statistical errors, the matching errors, the systematics due to our chiral-continuum extrapolation, errors coming from HQ-mistuning corrections, and discretization errors. The next several rows show estimates of the individual contributions (in parentheses as a reminder that they are contained in the first row). Since the terms that describe the discretization errors come from both the (N)NLO terms in the extrapolation and the HQ discretization terms, there is an overlap between the statistical errors (determined as those of a chiral-continuum extrapolation at NLO without any matching or heavy-quark mistuning errors taken into account) and the discretization errors, and the sum in quadrature of the numbers in parenthesis does not equal the first row. The remaining rows show other contributions, which are added to the first row in quadrature to obtain the total error in the last row. Dashes represent terms so small that were not included in the final computation of the error

Source	$h_V(\%)$	$h_{A_1}(\%)$	$h_{A_2}(\%)$	$h_{A_3}(\%)$
Chiral-continuum fit error	4.2	2.0	17.4	6.9
(Statistics)	(3.7)	(1.2)	(16.9)	(6.3)
(Chiral-continuum extr.)	(0.8)	(0.9)	(1.7)	(0.5)
(LQ and HQ discretization)	(2.6)	(1.3)	(9.7)	(4.4)
(HQ mistuning)	(0.0)	(0.0)	(1.7)	(0.0)
(Matching $O(am_c\alpha_s)$)	(0.3)	(0.2)	(1.7)	(0.5)
LQ mistuning	0.0	0.0	0.1	0.0
Matching $O(\alpha_s^2)$	0.7	0.3	0.5	0.3
Scale setting	0.0	0.0	0.3	0.1
Isospin effects	0.1	0.1	0.4	0.2
Finite volume	–	–	–	–
Total error	4.3	2.0	17.4	6.9

these points have much larger errors. We refer to the reader to Sect. 3.3.1, where all the details are explained. For the form-factor ratio fits, we employ the same range in physical units for all ensembles and most ratios. For the double ratio R_{A_1} and x_f, where the same pattern of states is expected at source and sink, we use a symmetric fit range with $t_{\text{Max}} = T - t_{\text{Min}}$. Our fits also take into account all correlations between the data points fitted, and, as pointed out in Sect. 3.2, we find that autocorrelations in our data are negligible.

For each correlator fit, we compute a p value from the deaugmented χ^2 and number of degrees of freedom, as explained in 3.2. We then verify that these p values follow an approximately uniform distribution.

4.2 Stability of the chiral-continuum extrapolation

To assess the stability of the chiral-continuum extrapolation, we repeat the fit for several different functional forms. Compared with the base fit, we omit, in turn, the NNLO terms, data at $w > 1.10$, the coarsest ensemble, the finest ensemble, and heavy-quark discretization terms of order a^3. The results are very stable under modifications, as shown in Fig. 8. The most dramatic changes occur when we remove the $w > 1.10$ data, leading to shifts of around one standard deviation in the worst case. Obviously, the large-recoil extrapolation is affected when removing data at larger recoil, but we find that the z expansion, discussed below in Sect. 5.1, stabilizes the final results for $|V_{cb}|$ and $R(D^*)$ in this respect. Table 8 shows that the quality of fit remains good, with $\chi^2/\text{dof} \lesssim 1$, in all cases.

4.3 Discretization errors

The improved action in the ensemble simulations has light-quark and gluon discretization errors of order $\alpha_s a^2$. Simple power-counting arguments suggest that the discretization errors range from $\sim 0.5\%$ in the finest ensemble to $\sim 10\%$ in the coarsest one. The chiral-continuum extrapolation describes, however, such terms via the lattice-spacing dependence of the chiral logarithms and the analytical terms proportional to a^2. Moreover, the form factors do not seem to be sensitive to the lattice spacing. Hence, we expect the chiral-continuum extrapolation to take all these errors into account, and no further systematic uncertainties are added to the final result.

In order to take into account the discretization errors coming from the lattice treatment of the heavy quarks, we include extra terms in the chiral-continuum extrapolation as shown in Eq. (53). These terms are motivated by the HQET description of cutoff effects [50,103], which uses HQET to derive the mismatch between the lattice gauge theory at hand and continuum QCD. The result is a set of functions that depend on the heavy-quark mass, and that can account for discretization effects of different sizes (in our case, order $\alpha_s a$, a^2 and a^3).

In our analysis, we would like to introduce these functions for both the B and the D^* mesons. Our data do not, however, distinguish between the contributions of the two mesons, so

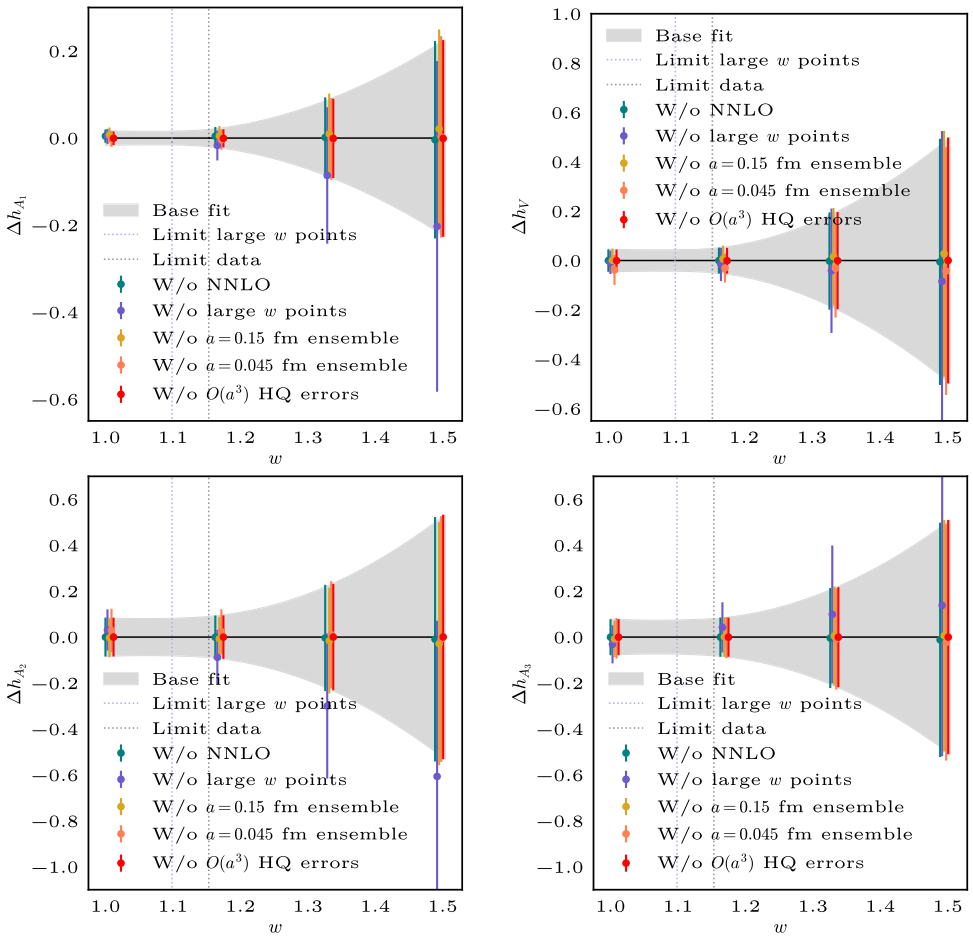

Fig. 8 Changes in chiral-continuum-extrapolated values as a result of selected variations in the analysis for h_{A_1} (top left), h_V (top right), h_{A_2} (bottom left), and h_{A_3} (bottom right) as a function of the recoil parameter. The gray band shows the error band of the base fit. The points at $w = 1.00$, 1.17, 1.34, and 1.50 show the deviation between the form factors calculated from the base fit and each variation of the chiral-continuum fit. The dotted vertical lines show the maximum w for which lattice data are available and the limit $w = 1.10$ used in one of our stability tests

Table 8 Values of χ^2/dof for several variations in the chiral-continuum extrapolation

	Base	W/o NNLO	W/o large w	W/o $a = 0.15$ fm	W/o $a = 0.045$ fm	W/o HQ $O(a^3)$
χ^2/dof	85.2/95	86.0/107	71.1/75	79.4/86	81.6/86	85.3/99

we instead use a single generic term for both mesons. Equation (53) shows the terms used in the end: $\alpha_s a$, a^2, and a^3. Since the a^2 term is already included in the light-quark discretization errors, it would be superfluous to include it again here. The downside of this approach is that it is impossible for us to disentangle light- and heavy-quark discretization errors in the error budget, and as such, we report them together.

One can estimate the size of these individual effects from variations of the chiral-continuum extrapolation with and without the terms in Eq. (53), and also removing the $O(a^2)$ term coming from NLO corrections in Eq. (50). Heavy- and light-quark discretization errors turn out to be the largest contribution to the total error in our analysis, and the inclusion of the terms listed in Eq. (53) is key in order to account for the heavy-quark systematic errors.

4.4 Matching errors

The matching factors are calculated at one-loop order in perturbation theory. Appendix B explains how we estimate the uncertainties, listed in Eqs. (B.49a)–(B.49d). They are included in the chiral-continuum extrapolation through Eq. (54). We can estimate the error introduced by the uncertainty in the matching up to order $am_c\alpha_s$ by removing the contribution of the matching factors to Eq. (54). The effect of higher order contributions is estimated by including an overall factor $(1 + r_2^{h_X}\alpha_s^2 + r_3^{h_X}\alpha_s^3)$ multiplying Eq. (49) in the fit and checking the shift in the central value of the form factor. The priors for the $r_{2,3}^{h_X}$ coefficients are set to 0(1); the posteriors have central values and widths close to those of the priors. We see no impact in including $O(\alpha_s^3)$ terms, but the $O(\alpha_s^2)$ contribute to the final error at the subpercent level. We collect all observed differences in the corresponding line in Table 7.

4.5 Heavy quark mistuning

The form factors are adjusted for the differences between the simulated masses of the heavy quarks and the physical ones before the chiral-continuum extrapolation. The correction procedure is detailed in Appendix C. The largest correction is about 1σ, but in general the correction is negligible. Equation (54) includes the contribution of the mistuning in the chiral-continuum extrapolation, therefore, we do not need to add any further error. Switching off these corrections gives small variations in the results in our chiral-continuum extrapolation, as shown in Table 7 and Fig. 7.

4.6 Light quark mistuning

The endpoint for the light quark masses in the chiral-continuum extrapolation is set to $r_1 m_l = 0.003612(126)$ [104]. We can determine the uncertainty in the form factors coming from a mistuning in the light quark mass by varying $r_1 m_l$ within 1σ and monitoring its effect on the form factors. The resulting uncertainty is shown in Table 7 and Fig. 7.

4.7 Scale setting

In order to determine the relative lattice spacing, we use the distance scale r_1/a defined from the force between static quarks [105,106], which has been extensively computed [54]. Absolute scale setting is taken from the chiral f_π analysis of the MILC Collaboration [64], leading to $r_1 = 0.3117(22)$ fm. The form factors are dimensionless, so uncertainties from scale setting appear indirectly through the tuning of the heavy-quark masses, the setting of the light-meson masses in the chiral logarithms, and in the approach to the continuum limit.

We estimate the systematic error associated with r_1/a and r_1 by propagating their errors to the final result. We find that the form factors change only slightly when we vary r_1 or r_1/a by $\pm 1\sigma$, and we include an extra error associated to this variation as shown in Table 7 and Fig. 7.

4.8 Isospin effects

The whole calculation of the form factors has been done assuming isospin symmetry. The main effect of isospin breaking is to modify the endpoint of the chiral extrapolation through a change in the pion mass. This effect could bring the endpoint of the extrapolation closer to the $D\pi$-threshold cusp described by the chiral logs. We estimate the errors introduced by this approximation by varying the endpoint of the extrapolation in the pion mass from m_{π^0} to m_{π^+}, by modifying the value of $r_1 m_l$ from 0.003612 to 0.004065. While the pion-mass difference is mainly due to isospin breaking QED effects, here we are using it as proxy for the valence quark mass difference, to which we do not have direct access. Following the resulting difference, we assign an error ranging from 0.0% to 0.5%, depending on the form factor and the value of the recoil parameter, as shown in Fig. 7 and Table 7. This increase in the error has no impact in the final result for the form factors.

As an alternative way of estimating these effects, we have also tried to move the endpoint of the extrapolation to isospin symmetric points with $m_l = m_u$ and $m_l = m_d$. The difference of the values of the form factors between these two endpoints overestimates the isospin breaking errors, because it includes sea-quark effects that cancel out at first order. The estimate of the isospin breaking errors is larger with this method, but it is still negligible. Hence we can safely assume that isospin effects are insignificant at our current level of precision.

4.9 Finite-volume effects

To estimate finite-volume effects in our heavy-light χPT description of the form factors, we replace the loop integrals by discrete sums. Following Refs. [87,107], we estimate the correction to the integrals in the formulas appearing in $B \to D^*$ at zero recoil to be smaller than 0.01%. This comes from the fact that the contribution of the chiral logarithms to the form factors is quite small. We have not calculated the corrections at nonzero recoil, and one also expects an increase in the error close to the cusp of the chiral logs. Given that $M_\pi L > 4$ on most ensembles, and $M_\pi L \geq 3.7$ always, there is no reason to expect such a large increase in the error as to make the finite-volume corrections sizable. Hence, we do not assign any additional error due to them.

5 Determination of $|V_{cb}|$ and $R(D^*)$

After calculating the form factors, we can reconstruct the decay amplitude using Eq. (9) and use experimental data to extract $|V_{cb}|$. Similarly, the form factors lead directly to $R(D^*)$ via Eqs. (1), (7), and (8). There is a problem: the form factors are obtained only at small values of the recoil parameter, and an extrapolation to large w with the chiral--continuum fit formula would greatly increase the error. To bring the large w behavior under control we use a standard, model-independent parametrization based on unitarity and analyticity to extrapolate the form factors to the large recoil region.

Historically the CLN parametrization [8] has been widely used for this process. However, recent developments have called into question the reliability of CLN fits, given the high accuracy of the latest experiments and calculations [14,16]. Apart from using outdated data to derive the coefficients of the expansions, the main criticism of CLN in its most com-

mon usage is the lack of error estimates for theoretical ingredients. Even though the original CLN article [8] includes equations defining the covariance matrix of the slope and the curvature of the reference form factor, the final expressions omit this information. Finally, the strong unitarity constraints, based on heavy-quark symmetry, play an important role in the CLN parametrization. Instead of introducing these constraints as an additional assumption, we would rather perform a fit without imposing them at the outset and then use them after the fits as a consistency check.

To perform the z expansion we therefore use the completely general BGL parametrization [5–7]. Nonetheless, we compare our results with an updated version of CLN in Sect. 5.4.2. For a review on the status of the parametrizations for heavy-to-heavy decays, see Ref. [108].

5.1 z expansion with the BGL parametrization

The z expansion is based on a conformal map that takes w to a variable z, which remains small over the physical region for the decay process, namely,

$$z = \frac{\sqrt{w+1} - \sqrt{2N}}{\sqrt{w+1} + \sqrt{2N}}. \tag{55}$$

The value $N = 1$ is most commonly used, because it fixes the point $z = 0$ at zero recoil, but a symmetric range has been advocated, with the claim that it reduces errors in the expansion [7,8]. With $N = 1$, the maximum recoil point for massless leptons $w_{\rm Max} \approx 1.503$ becomes $z_{\rm Max} \approx 0.056$, so indeed $z \ll 1$ in any case, and any reasonable expansion in z with coefficients of order 1 converges with a few terms. The conformal map given in Eq. (55) also pushes the branch cut in the w plane onto the unit circle, $|z| = 1$, and subthreshold poles onto the real axis near $z = -1$. Since the nearest threshold is very far ($|z| \sim 1$) from the valid kinematic range ($0 \le z \lesssim 0.056$) and higher-energy thresholds even farther, there is no advantage in using an alternative parametrization, such as the one proposed in Ref. [109], that would take care of the behavior at such values of z.

The BGL parametrization does not apply directly to the h_X form factors, but to combinations with definite spin-parity [108]:

$$g = \frac{h_V}{M_B\sqrt{r}}, \tag{56}$$

$$f = M_B\sqrt{r}(1+w)h_{A_1}, \tag{57}$$

$$\mathcal{F}_1 = M_B^2\sqrt{r}(1+w)\Big[(w-r)h_{A_1} - (w-1)\left(rh_{A_2} + h_{A_3}\right)\Big], \tag{58}$$

$$\mathcal{F}_2 = \frac{1}{\sqrt{r}}\Big[(1+w)h_{A_1} + (rw-1)h_{A_2} + (r-w)h_{A_3}\Big], \tag{59}$$

where $r = M_{D^*}/M_B$. These form factors are proportional to the helicity amplitudes $H_- - H_+$, $H_+ + H_-$, H_0, and H_S, respectively [cf., Eqs. (4)-(6)]. Thus, the form factor $\mathcal{F}_2$ is important only with massive leptons, in particular in the determination of $R(D^*)$.

The BGL parametrization expresses the dependence of the form factors on z as

$$f_i(z) = \frac{1}{P_i(z)\,\phi_i(z)} \sum_{j=0}^{\infty} a_{i,j} z^j, \tag{60}$$

where the functions $P_i(z)$ are called *Blaschke factors*, and the ϕ_i are known as *outer functions*. As discussed below, wise choices of $\phi_i(z)$ make the coefficients of the expansion $a_{i,j}$ of order 1 and ensure rapid convergence of the series. The Blaschke factors are given by

$$P_i(z) = \prod_p \frac{z - z_p}{1 - z z_p}, \tag{61}$$

with

$$z_p(M_p, N) = \frac{\sqrt{(1+r)^2 - \frac{M_p^2}{M_B^2}} - \sqrt{4Nr}}{\sqrt{(1+r)^2 - \frac{M_p^2}{M_B^2}} + \sqrt{4Nr}}. \tag{62}$$

They include the explicit poles with mass M_p below the BD^* threshold and with the appropriate quantum numbers. Table 9 shows the poles we use for the BGL form factors. Although some analyses employ four 1^- resonances, the fourth one is very far from $z > 0$ and its value uncertain. We therefore follow Ref. [15] and use only three.

The z-expansion coefficients of the BGL form factors are then defined via

$$g = \frac{1}{P_{1^-}(z)\,\phi_g(z)} \sum_{j=0}^{\infty} a_j z^j, \tag{63}$$

$$f = \frac{1}{P_{1^+}(z)\,\phi_f(z)} \sum_{j=0}^{\infty} b_j z^j, \tag{64}$$

$$\mathcal{F}_1 = \frac{1}{P_{1^+}(z)\,\phi_{\mathcal{F}_1}(z)} \sum_{j=0}^{\infty} c_j z^j, \tag{65}$$

$$\mathcal{F}_2 = \frac{1}{P_{0^-}(z)\,\phi_{\mathcal{F}_2}(z)} \sum_{j=0}^{\infty} d_j z^j, \tag{66}$$

where the Blaschke factors' subscripts denote the J^P of the $l\nu$ final state.

Setting $N = 1$, we choose the outer functions to be [5–7]

$$\phi_g = 16r^2 \sqrt{\frac{n_I}{3\pi\,\tilde{\chi}_{1^-}^T(0)}} \frac{(1+z)^2(1-z)^{-\frac{1}{2}}}{\left[(1+r)(1-z) + 2\sqrt{r}(1+z)\right]^4}, \tag{67}$$

Table 9 Poles for the Blaschke factors, taken from Ref. [15] and references therein. For $J^P = 1^-, 0^-\,(1^+)$, the first two (first only) resonances are well determined from either a lattice calculation or experimental measurements. All other masses are based on model estimates

Form factor	J^P	Masses M_p (GeV)
g	1^-	6.329, 6.920, 7.020
$f, \mathcal{F}_1$	1^+	6.739, 6.750, 7.145, 7.150
$\mathcal{F}_2$	0^-	6.275, 6.842, 7.250

$$\phi_f = \frac{4r}{M_B^2}\sqrt{\frac{n_I}{3\pi\,\chi^T_{1^+}(0)}}\,\frac{(1+z)(1-z)^{\frac{3}{2}}}{\left[(1+r)(1-z)+2\sqrt{r}(1+z)\right]^4}, \tag{68}$$

$$\phi_{\mathcal{F}_1} = \frac{4r}{M_B^3}\sqrt{\frac{n_I}{6\pi\,\chi^T_{1^+}(0)}}\,\frac{(1+z)(1-z)^{\frac{5}{2}}}{\left[(1+r)(1-z)+2\sqrt{r}(1+z)\right]^5}, \tag{69}$$

$$\phi_{\mathcal{F}_2} = 8\sqrt{2}r^2\sqrt{\frac{n_I}{\pi\,\tilde{\chi}^L_{1^+}(0)}}\,\frac{(1+z)^2(1-z)^{-\frac{1}{2}}}{\left[(1+r)(1-z)+2\sqrt{r}(1+z)\right]^4}, \tag{70}$$

where the undefined symbols under the square root are given in Table 10, along with the numerical values we use for M_B and M_D^*. In testing the effect of this pole and various other choices for the pole positions, we find that although the numerical values of the z-expansion coefficients depend on the details, the final curves for the form factors are largely independent of such choices.

With these outer functions, the coefficients of the expansion satisfy the following unitarity constraints [5–7],

$$\sum_{j=0}^{\infty} a_j^2 \lesssim 1, \quad \sum_{j=0}^{\infty}\left(b_j^2 + c_j^2\right) \lesssim 1, \quad \sum_{j=0}^{\infty} d_j^2 \lesssim 1, \tag{71}$$

where the $\lesssim$ symbols reflect the fact that the values for the χ factors in Table 10 are not exact, so the bounds are not precisely known. These constraints provide information from unitarity and analyticity, which can be used together with the output of the chiral-continuum extrapolation.

In this analysis, we do not impose the unitarity constraints, Eq. (71), on the BGL coefficients, but we check that the final results comply with the constraints within errors. We note that uncertainties in the values of $\chi^{T,L}_{J^P}$ provided in Table 10 complicate strict unitarity constraints on the z expansion coefficients. Still, we find that an implementation of the unitarity constraints using hard cutoffs (following, for instance, Ref. [110]) leaves the fit results essentially unchanged.

There are two kinematic relations between the form factors, one at zero recoil and another at maximum recoil,

$$\mathcal{F}_1(1) = M_B(1-r)f(1), \tag{72}$$

Table 10 Inputs for the outer functions, taken from Ref. [15] and references therein. The χ parameters are calculated in perturbative QCD through $O(\alpha_s^2)$, and depend on charm and bottom quark mass inputs, see Ref. [111]

Input	Value
M_{D^*} (GeV)	2.010
M_B (GeV)	5.280
n_I	2.6
$\chi^T_{1^+}(0)$ (GeV^{-2})	3.894×10^{-4}
$\tilde{\chi}^T_{1^-}(0)$ (GeV^{-2})	5.131×10^{-4}
$\tilde{\chi}^L_{1^+}(0)$	1.9421×10^{-2}

$$\mathcal{F}_2(w_{\rm Max}) = \frac{1+r}{M_B^2(1+w_{\rm Max})(1-r)r}\mathcal{F}_1(w_{\rm Max}). \tag{73}$$

These constraints follow trivially from the HQET basis of form factors (the h_X), but the BGL parametrization does not automatically impose them.[4] Equation (72) is straightforward to implement for the BGL parametrization with $N = 1$ in Eq. (55), since it amounts to a relationship between the b_0 and the c_0 coefficients of the expansion,

$$\frac{1-r}{\sqrt{2}(1+\sqrt{r})^2}b_0 = c_0. \tag{74}$$

On the other hand, Eq. (73) can be imposed by adding an extra data point that enforces the constraint. Alternatively, we can remove d_0 and write it as a function of the remaining d_j and c_j. The two approaches give compatible results. In our final value, we choose not to impose the second constraint. The results of the chiral-continuum extrapolation trivially build in both constraints, so we would expect any well-behaved expansion to keep this property, even when extrapolating to the whole recoil range. Indeed, the zero-recoil constraint is satisfied to very high accuracy, even if we do not impose it. This happens because the z expansion is quite constrained by the lattice-QCD values. For the maximum-recoil constraint, we just check for compatibility within errors.

5.1.1 Synthetic data

The output of the chiral-continuum extrapolation is not a set of points, but a set of functions that express the form factors at any value of the recoil parameter. In order to make the results amenable to a BGL fit with experimental data, we use *synthetic data*, together with their covariance, based on selected central values of the chiral-continuum fit, evaluated at zero

[4] It is interesting to note that the CLN parametrization does include the constraint at zero recoil given in Eq. (72). The constraint at maximum recoil, Eq. (73), is not imposed in CLN, and indeed does not hold unless the original CLN expressions are modified.

lattice spacing and physical quark masses. The selected values with correlations are included in the ancillary files, as explained in Appendix D. We choose data points at three w values, {1.03, 1.10, 1.17}, as representative of the span of our lattice-QCD data, but we have checked that varying these values does not change significantly the curves generated for the form factors from the z expansion, as long as there are no w values too close to $w = 1$. This robust behavior is not surprising, since the full covariance matrix is available, and hence the same amount of information is provided. The covariance matrix of the lattice data is well defined for all recoil values, but at $w = 1$ the kinematic constraint given in Eq. (72) is exactly satisfied. Therefore, the form factors are no longer independent as $w \to 1$, and the covariance matrix becomes singular. The singularity can be avoided by choosing any value slightly larger than $w = 1$. Also, we cannot push w much higher than 1.17 without going outside the region where lattice-QCD data are available, because the uncertainty grows rapidly. We have selected three points per form factor because in the continuum limit there are only twelve free independent functions in our chiral-continuum extrapolation, three per form factor. Adding more points does not increase the accuracy of the z-expansion fits, because the new points are not independent.

We first carry out BGL fits to our lattice-QCD form factors. The constant and linear coefficients are well determined by the data, and no prior constraints are used for them. The quadratic and cubic coefficents are constrained with priors 0(1) to stabilize the fit. Keeping terms up to quadratic (linear) order in z and imposing the kinematic relation given in Eq. (72), leaves just one (three) degree(s) of freedom. Since unitarity constrains the size of the coefficients, we can include cubic terms with the unitary-inspired priors to check stability of the results against truncation effects. Table 11 gathers the coefficients for these three versions of the z expansion, with the χ^2/dof and unitarity sums. All fits satisfy the unitarity constraints within errors. The unitarity sums are computed by taking the median of the distributions obtained from the Gaussian posteriors, along with the confidence levels, $\pm 34.1\%$, for the uncertainties. We note that the distributions of the squares are not Gaussian when the errors on the posteriors are large. This is the case for coefficients that are not well determined by the underlying lattice data. The kinematic constraint at $z = 0$, Eq. (72), is satisfied to very high accuracy, even when it is not imposed. On the other hand, the constraint at $z_{\rm Max}$, Eq. (73), is satisfied only within approximately one standard deviation, unless, of course, it is imposed.

The cubic coefficients (a_3, b_3, c_3, d_3) shown in Table 11 are not well determined by our lattice data, resulting in unitarity sums, with central values > 1, while still consistent with unitarity within error. However, the cubic fit provides useful information on truncation effects. We see that the lower-order coefficients in the cubic fit are in very good agreement with the coefficients in the quadratic fit. Further, we find, that the cubic coefficients have little effect on the decay rate and the form factors, as expected, since $|z| \ll 1$. In contrast, and as Table 11 shows, the linear expansion leads to coefficients with nearly the same central values as in the quadratic and cubic fits, but with slightly smaller errors, suggesting an underestimation of the truncation error. We conclude that the errors coming from the truncation of the series in Eqs. (63)–(66) are already included in the uncertainties on the coefficients of the quadratic fit, which we choose for the main result of the z expansion. The full correlation matrix of the quadratic fit, which can be used to reconstruct the output of the z expansion, is included in the ancillary files in binary format, as explained in Appendix D. Form factors from this information can be used in phenomenology with no assumption about the presence of new physics. Below we discuss z fits incorporating shape information from experiment, which are more precise but possibly contaminated by new physics in the light semileptonic channel.

Experimentalists [18,19] usually adjust the order of the z expansion to allow unconstrained fits to the form factors without violating unitarity. Following this criterion we find that we can remove the a_2 and the d_2 coefficients, and obtain the same fit results as in our quadratic fit without including any 0(1) priors for the higher order coefficients. This ensures that our utilization of priors for some coefficients indeed does not influence the fit results, and that their only function is to stabilize the fit.

5.1.2 *Functional method*

A functional method can also be used to fit the result of the chiral-continuum extrapolation to the BGL parametrization [39]. The method exploits the fact that the chiral-continuum fit functions are linear in the fit parameters. The covariance in the fit parameters is then easily converted to the covariance of the values of the resulting fitted form factors (at zero lattice spacing and physical quark masses) at any pair of recoil parameters (w, w'). Through the BGL parametrization, this covariance is then converted to a covariance in the form factors values at any z pair, (z, z'). The method has the esthetic property that information from the best fit continuum form factors is spread over the entire physical region in z, rather than at a few arbitrarily chosen discrete points.

We have compared form factors from the functional approach with those from the synthetic data. They show no discernible difference in the form factors, implying that the systematic errors associated with the choice of synthetic data from the chiral-continuum extrapolation are very small. Since the functional fits do not provide any new insight, and they make it difficult to combine data from several sources, we focus on the synthetic-data results in the rest of the paper.

Table 11 Results of linear, quadratic, and unitarity-constrained cubic z expansions using only lattice-QCD data. The coefficient c_0 is fixed by the constraint given in Eq. (72), and it is shown for convenience

	Linear	Quadratic	Cubic
a_0	0.0330(12)	0.0330(12)	0.0330(12)
a_1	− 0.157(52)	− 0.156(55)	− 0.155(55)
a_2		− 0.12(98)	− 0.12(98)
a_3			− 0.004(1000)
b_0	0.01229(23)	0.01229(23)	0.01229(23)
b_1	− 0.002(10)	− 0.003(12)	− 0.003(12)
b_2		0.07(53)	0.05(55)
b_3			− 0.01(100)
c_0	0.002059(38)	0.002059(38)	0.002059(38)
c_1	− 0.0057(22)	− 0.0058(25)	− 0.0057(25)
c_2		− 0.013(91)	− 0.02(10)
c_3			0.10(95)
d_0	0.0508(15)	0.0509(15)	0.0509(15)
d_1	− 0.317(59)	− 0.328(67)	− 0.327(67)
d_2		− 0.02(96)	− 0.02(96)
d_3			− 0.0006(10000)
χ^2/dof	0.83/5	0.63/1	0.61/-3
$\sum_i^N a_i^2$	$0.026^{+0.019}_{-0.013}$	$0.47^{+1.51}_{-0.36}$	$1.4^{+2.3}_{-1.0}$
$\sum_i^N (b_i^2 + c_i^2)$	$2.4^{+1.8}_{-0.6} \times 10^{-4}$	$0.13^{+0.44}_{-0.10}$	$1.6^{+2.2}_{-1.1}$
$\sum_i^N d_i^2$	$0.103^{+0.041}_{-0.034}$	$0.54^{+1.40}_{-0.38}$	$1.4^{+2.2}_{-1.0}$

5.2 Determination of $|V_{cb}|$

The lattice-QCD form factors can be used in conjunction with experimental data to perform a joint fit to the BGL parametrization, with an additional fit parameter for the relative normalization, which is nothing but $|V_{cb}|$. In these fits, the low-recoil behavior is determined by lattice QCD, and the large-recoil behavior by experiment. As experimental input, we use the 2018 raw dataset from Belle [18] and the synthetic data generated from the 2019 BaBar analysis [19]. These data are combined with the lattice-QCD synthetic data. We do not use Belle's 2017 tagged dataset [13], because it is still unpublished.

Experiments extract the fully differential decay rate, not only with respect to the recoil parameter, but also to all angular variables in the decay chain $B \to D^* \ell \nu$, $D^* \to D\pi$ [14,18,112],

$$\begin{aligned}
&\frac{d\Gamma}{dw\, d\cos\theta_v\, d\cos\theta_\ell\, d\chi} \\
&= |V_{cb}|^2\, |\eta_{\rm EW}|^2\, \frac{3G_F^2 M_B^5}{1024\pi^4} r^3 \sqrt{w^2-1}(1-2wr+r^2) \\
&\times \Big[(1-\cos\theta_\ell)^2 \sin^2\theta_v H_+^2(w) + (1+\cos\theta_\ell)^2 \\
&\times \sin^2\theta_v H_-^2(w) + 4\sin^2\theta_\ell \cos^2\theta_v H_0^2(w) \\
&- 2\sin^2\theta_\ell \sin^2\theta_v \cos 2\chi H_+(w) H_-(w) - 4\sin\theta_\ell \\
&\times (1-\cos\theta_\ell) \sin\theta_v \cos\theta_v \cos\chi H_+(w) H_0(w) \\
&+ 4\sin\theta_\ell (1+\cos\theta_\ell) \sin\theta_v \cos\theta_v \cos\chi \\
&\times H_-(w) H_0(w) \Big] \mathcal{B}(D^* \to D\pi), \qquad (75)
\end{aligned}$$

where $\mathcal{B}(D^* \to D\pi)$ is the branching fraction of the daughter D^* decay; further, θ_v, θ_ℓ, and χ are the polar angle of the D in the D^* rest frame, the polar angle of the charged lepton in the rest frame of the virtual W meson, and the angle between the $\ell\nu$ and $D\pi$ planes, respectively. As with Eq. (7), for neutral B^0 decays the right-hand side of Eq. (75) should have an additional factor $(1 + \alpha\pi)$ for the Coulomb attraction in the final state (see for example, Refs. [113,114]). Other electromagnetic corrections are expected to be smaller, and we will neglect them, keeping only $|\eta_{\rm EW}|^2$ and the Coulomb factor.[5] Following our previous estimates of EM effects [22,53], as well as the HFLAV procedure [1,115], we use $\eta_{\rm EW} = 1.0066(50)$ in our calculation.

Belle marginalizes on one variable at a time, integrating (binning) the rest. BaBar's method consists of a full, four-dimensional analysis without integrating over any variable. The collaboration claims such an analysis is needed to achieve correct results [19]. Nonetheless, both the Belle and

[5] Both the Belle and BaBar experiments use the PHOTOS package to account for low-energy EM radiation; to the best of our knowledge, the EM interactions between charged particles in the final state, described by the Coulomb factor, are not included in PHOTOS.

BaBar Collaborations give compatible final values of $|V_{cb}|$ in their respective publications [18,19].

For Belle, we integrate Eq. (75) in the required bins, and the BGL expressions are introduced in the integrated results. Also, we multiply the right-hand side of Eq. (75) by the Coulomb factor $(1 + \alpha\pi)$, because these data are for neutral B mesons only. We perform a combined fit to both the electron and muon modes, instead of averaging them.[6] For BaBar, we fit the lattice-QCD synthetic data with those in Ref. [19]. BaBar publishes results for a BGL fit to their data that includes both neutral and charged B-meson decays. According to Ref. [117], 35.1% of the decays in this data set correspond to B^0 decays, and the remaining 64.9% correspond to $B^\pm$ decays. These fractions imply that a Coulomb factor of $(1 + 0.351\alpha\pi)$ should be applied to the $B^0 + B^\pm$ BaBar dataset.

BaBar uses the previous Fermilab-MILC result of h_{A_1} at zero recoil to extract $|V_{cb}|$ [19,22]. In order to minimize the influence of the lattice-QCD results for $h_{A_1}(1)$ from Ref. [22] in the joint fit, we create synthetic data from Ref. [19] for $|\eta_{EW}|^2|V_{cb}|^2|\mathcal{F}(w)|^2$ at five recoil values away from $w = 1$, as shown in Fig. 9. Since the BaBar Collaboration employs a linear fit for all form factors, we exhaust the number of degrees of freedom with just five points. Each one of these synthetic data points has a larger impact than each single data point from Belle's untagged dataset, because the Babar synthetic data inherits its precision from the precision of the underlying, full dataset. In the end, the green band for Belle in Fig. 9 is noticeably narrower than BaBar's, so it is expected that the Belle untagged dataset has a larger impact in the final results for $|V_{cb}|$ and $R(D^*)$.

The kinematic constraint in Eq. (72) is included in these fits, and although there are no direct experimental measurements that determine $\mathcal{F}_2$, experiments also have an impact on the d_j coefficients through the correlations between this and other form factors. Our preferred fits, coming from quadratic z expansions, are shown in Fig. 9 and Table 12. As in the lattice-QCD-only fits, the fits in Table 12 include 0(1) priors for the quadratic and higher (if applicable) coefficients of each form factor, while leaving the rest of the coefficients unconstrained. Fig. 9 shows that the mean of the lattice estimate falls below the experimental curves, but the errors are large enough to make the difference remain at $\approx 2\sigma$. The correlations between the different lattice, synthetic data points determine very precisely the slope of the decay amplitude, forcing it to be noticeably larger than what we obtain from our fits to experimental data. The full correlation matrix is provided in the ancillary files, as described in Appendix D. Form factors from this information can be used in phenomenology, under the assumption that only the τ couples to new physics.

Our final result for $|V_{cb}|$ is obtained from the quadratic BGL fit to the lattice-QCD form factors and both experimental datasets (see the column labeled "Lattice+both" in Table 12), which yields

$$|V_{cb}| = (38.40 \pm 0.78) \times 10^{-3}, \tag{76}$$

and a $\chi^2/\text{dof} = 126/84$. This relatively large χ^2/dof indicates tensions among the datasets: a combined fit of Belle and BaBar data, using lattice-QCD input only for normalization results in a large χ^2/dof of 104/76. It is therefore to be expected that the combined fit would result in a similarly large χ^2/dof. Further, we note that our fit to the lattice-QCD form factors only has $\chi^2/\text{dof} < 1$, as shown in Table 12, which also lists the results of joint fits of lattice-QCD form factors with each experimental dataset separately, as well as with the combined Belle and BaBar data. We find that all joint lattice-QCD with experimental data fits have $\chi^2/\text{dof} > 1$, including the one leading to Eq. (76), but the central values of $|V_{cb}|$ do not differ more than approximately one standard deviation among these fits, and the sizes of the errors are similar. We also see a general agreement in the coefficients of the expansion, particularly in the important low-order ones.

Because most previous inclusive and exclusive determinations of $|V_{cb}|$ omit the Coulomb factor, we also perform the BGL fits without it; the results are collected in Table 13. Compared to the results for $|V_{cb}|$ in Table 12, the central values are shifted by the respective Coulomb factors. They are consistent with previous exclusive determinations, for example $|V_{cb}|_{\text{excl}} = (39.9 \pm 0.9) \times 10^{-3}$ from the PDG [2]. The long-standing tension with inclusive determinations thus remains: $|V_{cb}|_{\text{incl}} = (42.2 \pm 0.8) \times 10^{-3}$ [2].

Since the Belle data are binned in different variables, there is a normalization constraint between the different bins, assuming that they contain the same underlying data. Then only 37 of the 40 bins are truly independent for each mode [116], because the sum of all bins for a particular variable should give the same total number of events. Such constraints should be reflected as zero eigenmodes, or – with rounding errors – very small eigenvalues in the 40×40 statistical correlation matrices. The correlation matrices provided in Ref. [18] are constructed using Monte Carlo simulations, and do not resolve these constraints due to the underlying approximations. We therefore investigate the effect of removing the last bin on each one of the angular variables data and reconstructing its value from the total normalization. We find that this procedure correctly introduces the anticipated constraints between the bins, while the values of the reconstructed last bins are compatible with those given Ref. [18]. Hence, the expected zero eigenvalues in the statistical correlation matrices are recovered. With this procedure our combined Belle + lattice-QCD BGL fit does not yield

[6] The systematic correlation matrices given in Ref. [18] do not include off-diagonal blocks for the correlated systematic errors between electron and muon modes. They can be reconstructed from the given data [116], but we do not attempt such a reconstruction in our analysis.

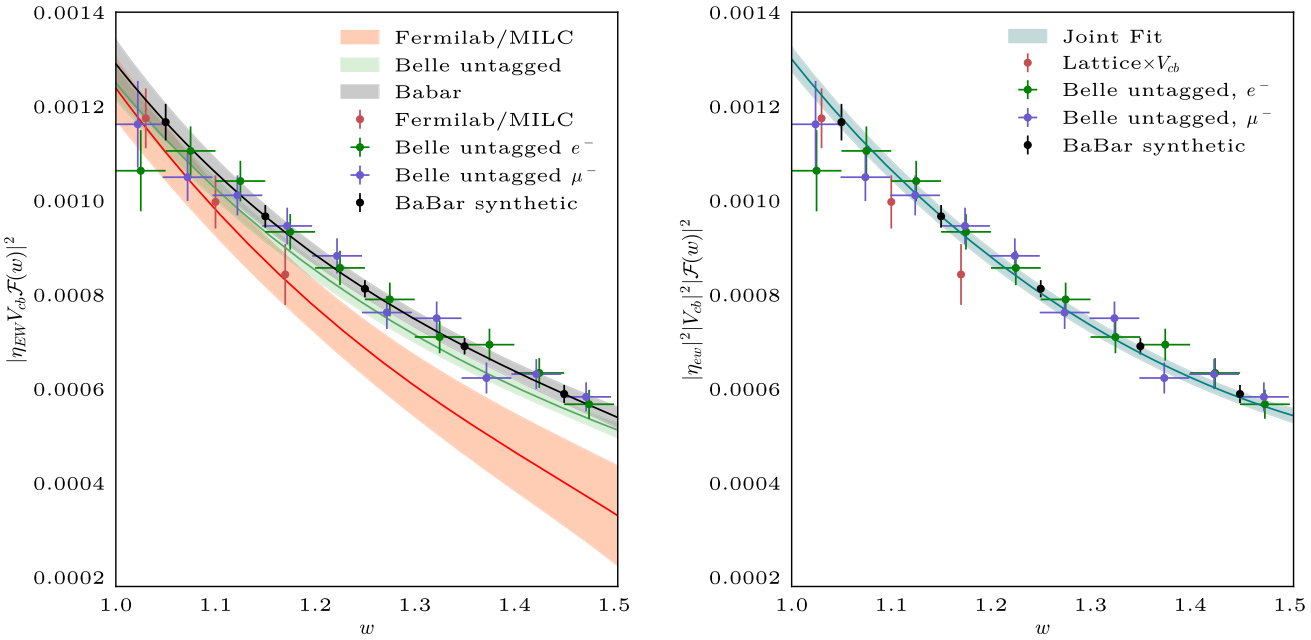

Fig. 9 Results for separate fits to each dataset (left) and joint fit of all data (right). On the left we compare the BaBar result (gray), the Belle result from the untagged dataset (green), and the lattice-QCD result coming from our synthetic data (red). To allow for an straightforward comparison of lattice and experimental data, the data points and bands have been normalized with the central value of $|V_{cb}|$ as obtained in our joint fit, and taking into account the Coulomb factor corresponding to each case. All results agree within $\approx 2\sigma$ over the whole kinematic range. There is tension between the slope predicted by the lattice calculation and that of the experimental data. Since the lattice-QCD slope is well determined, correlations in the joint fit cause the central lattice-QCD values to fall slightly below the experimental values

Table 12 Quadratic z expansion results. The second column shows results from a fit only to synthetic lattice-QCD data (the same as the "quadratic" column in Table 11), the third, from a joint fit to lattice QCD plus BaBar's synthetic data, the fourth, from lattice QCD plus Belle's untagged dataset, the fifth, lattice QCD plus both experiments, and the last, a combined fit of all experimental data using the value of $h_{A_1}(1)$ extracted from our chiral-continuum extrapolation as normalization. The coefficient c_0 is fixed by the constraint given in Eq. (72), and it is shown for convenience

	Lattice QCD	Lattice + BaBar	Lattice + Belle	Lattice + both	$h_{A_1}(1)$ + both
a_0	0.0330(12)	0.0331(12)	0.0325(11)	0.0321(10)	0.0262(42)
a_1	− 0.156(55)	− 0.091(41)	− 0.160(45)	− 0.147(31)	0.03(15)
a_2	− 0.12(98)	− 0.19(20)	− 0.69(94)	− 0.63(20)	− 0.12(12)
b_0	0.01229(23)	0.01229(22)	0.01238(22)	0.01249(22)	0.01228(23)
b_1	− 0.003(12)	0.0104(72)	0.015(10)	0.0021(43)	0.0046(48)
b_2	0.07(53)	0.44(17)	− 0.30(24)	0.07(11)	− 0.17(21)
c_0	0.002059(38)	0.002058(37)	0.002073(37)	0.002092(37)	0.002056(38)
c_1	− 0.0058(25)	− 0.0010(11)	0.0010(17)	0.00062(86)	0.00163(92)
c_2	− 0.013(91)	0.022(50)	0.035(57)	0.060(26)	0.017(37)
c_3		0.24(77)	− 0.34(76)	− 0.94(48)	− 0.66(62)
d_0	0.0509(15)	0.0517(15)	0.0522(15)	0.0531(14)	
d_1	− 0.328(67)	− 0.220(55)	− 0.180(49)	− 0.201(42)	
d_2	− 0.02(96)	0.20(92)	− 0.01(90)	0.0007(8980)	
χ^2/dof	0.63/1	8.50/4	111/79	126/84	104/76
$\sum_i^N a_i^2$	$0.47^{+1.51}_{-0.36}$	$0.05^{+0.12}_{-0.05}$	$0.7^{+2.0}_{-0.7}$	$0.43^{+0.30}_{-0.21}$	$0.04^{+0.05}_{-0.03}$
$\sum_i^N (b_i^2 + c_i^2)$	$0.13^{+0.44}_{-0.10}$	$0.56^{+0.98}_{-0.39}$	$0.54^{+1.00}_{-0.43}$	$0.90^{+1.17}_{-0.57}$	$0.57^{+1.13}_{-0.55}$
$\sum_i^N d_i^2$	$0.54^{+1.40}_{-0.38}$	$0.46^{+1.35}_{-0.37}$	$0.41^{+1.25}_{-0.34}$	$0.41^{+1.25}_{-0.33}$	
$\lvert V_{cb}\rvert \times 10^3$		39.1(1.0)	38.17(85)	38.40(78)	39.35(91)

Table 13 $|V_{cb}|$ results for our different BGL fits without including the Coulomb factor

	Lattice + BaBar	Lattice + Belle	Lattice + both	$h_{A_1}(1)$ + both
$\|V_{cb}\| \times 10^3$	39.3(1.1)	38.60(86)	38.74(78)	39.75(92)
χ^2/dof	8.48/4	111/79	124/84	103/76

any significant changes in the final values and uncertainties for $|V_{cb}|$ and $R(D^*)$, but we observe a substantial decrease in χ^2/dof from 111/79 to 96/73. In the case of our joint fit, which includes Belle, BaBar and lattice-QCD data, the χ^2/dof decreases from 126/84 to 109/78. Nevertheless, the results for $|V_{cb}|$ and $R(D^*)$ quoted in this work use the Belle data and correlation matrices as given in Ref. [18].

The BGL fit to the BaBar data [19] includes fewer coefficients than our BGL fit to the lattice-QCD form factors. We test for the presence of truncation errors by performing BGL fits to the BaBar data including higher coefficients with priors of 0.0(5). This increases the errors in the BaBar data points, most likely because the extra coefficients are completely uncorrelated with the rest of the BaBar data. Because the joint fit to all data is currently dominated by the Belle and lattice-QCD data, the addition of extra coefficients in the BaBar expansion does not change our final results for $|V_{cb}|$ and $R(D^*)$ in a meaningful way. Hence, for the our final results quoted in this work, the synthetic data points from Ref. [19] are generated without adding extra coefficients.

In the Belle and BaBar analyses the number of coefficients in the BGL z expansion is limited to exclude those that cannot be properly determined by their data, and thus avoiding apparent unitarity violations. This procedure, however, does not account for possible truncation errors. Repeating the fits in Table 12 without the c_3 and d_2 coefficients yields similar results with reduced errors and much smaller sums for the unitarity constraints.

5.3 Determination of $R(D^*)$

From the fit results in Table 12 we can calculate $R(D^*)$ through direct integration of the differential decay rate over the whole kinematic range. In Fig. 10, we show the differential decay rate as a function of the recoil parameter extracted using lattice-only data (red and brown curves), compared with that of our joint fit. The curves below (maroon and blue) show the differential decay rate for the τ case. Our final result for $R(D^*)$ from our purely lattice-QCD calculation is

$$R(D^*)_{\text{Lat}} = 0.265 \pm 0.013. \tag{77}$$

If we assume that new physics effects are visible only at large lepton masses (i.e., the τ), we can use our joint fit of the lattice and light-lepton experimental data to obtain a more precise SM value of $R(D^*)$. We note that in our joint fit, the curve corresponding to light leptons is determined mainly from experiment, and the one corresponding to the τ comes mainly from the lattice data. In that case, we obtain

$$R(D^*)_{\text{Lat+Exp}} = 0.2484(13), \tag{78}$$

where the Coulomb factor is included. Its removal does not change significantly neither the central value nor the error. We emphasize, however, that Eq. (77) is the SM prediction, relying only on lattice QCD, while Eq. (78) is also based on the shape information coming from experimental data. In any case, the correlated difference between the two results is 1.3σ. Our values also agree with previous theoretical determinations [20,21,118–120]. We note that more recent experimental measurements have found $R(D^*)$ to be consistently smaller than before, hence reducing the tension between theory and experiment [1]. The current status of the $R(D)$-$R(D^*)$ determinations is summarized in Fig. 11.

5.4 Tests

5.4.1 Imposing the constraint at maximum recoil

As we explained above, our preferred analysis does not impose the kinematic constraint in Eq. (73), is trivially satisfied in the HQET basis of form factors (the h_X) used in our chiral-continuum extrapolation. However, the BGL expansion does not naturally incorporate it. Maximum recoil is far from the region where lattice data are available, and there are no experimental data available for this decay with a heavy lepton $\ell = \tau$. Thus, to the extent that the BGL expansion does not match the HQET-basis form factors precisely, we expect small deviations from Eq. (73) in the BGL fit. Such deviations are tolerable because small violations of the constraint do not have any physical consequences, as long as they are within errors. Figure 10 shows that our fits, nonetheless, satisfy the maximum-recoil constraint to within approximately 1σ.

Imposing the constraint in the fit model, we find new values for $|V_{cb}| = 38.36(78) \times 10^{-3}$ and $R(D^*) = 0.274(10)$, which are compatible with the values obtained in our preferred analysis. It is not surprising that the constraint does not alter the value of $|V_{cb}|$. After all, the CKM matrix element is extracted mainly from the behavior of the form factors at small recoil and does not entail the form factor $\mathcal{F}_2$. The error on $R(D^*)$, on the other hand, is slightly reduced by the constraint.

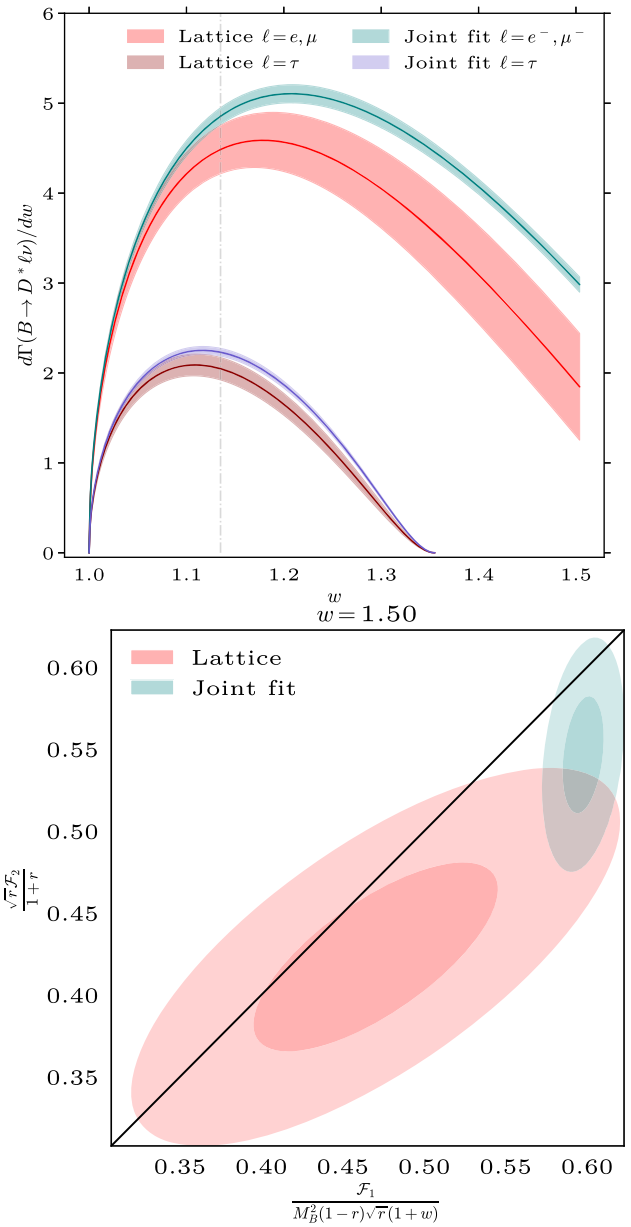

Fig. 10 Top: differential decay rate calculated using only lattice data (red and maroon) and lattice plus experimental data (green and blue). The higher curves are for a massless lepton, whereas the lower curves are for the τ. Although the pure-lattice curves are consistently below the experimental ones, especially at large recoil, both of them agree within 2σ. Bottom: test of the kinematic constraint at maximum recoil Eq. (73). Shown is a contour plot up to 2σ of the form factors $\mathcal{F}_1$ and $\mathcal{F}_2$ resulting from the lattice-data-only fit and the joint fit of lattice and experimental data. The constraint is satisfied along the diagonal. We see both fits satisfy the constraint within errors

A comparison of the BGL coefficients of the constrained analysis with those of our preferred one is shown in Table 14. We do not find significant changes, and the coefficients in both analyses are compatible with each other within 1σ, although the differences increase with the order of the coefficients. This behavior is expected, as the low-order coefficients are well determined by the data, and the higher-order coefficients become more relevant at maximum recoil. The new information does not improve the quality of the fit in a significant way. This is particularly clear in the pure lattice fit, where the χ^2 almost doubles, but the number of degrees of freedom increases just from three in the unconstrained fit to four in the constrained one. It appears that the constraint introduces small tensions with the BGL expansion. Because of this, and since both the unconstrained and the constrained fit give compatible results with only small differences, we choose the unconstrained fit as our preferred result.

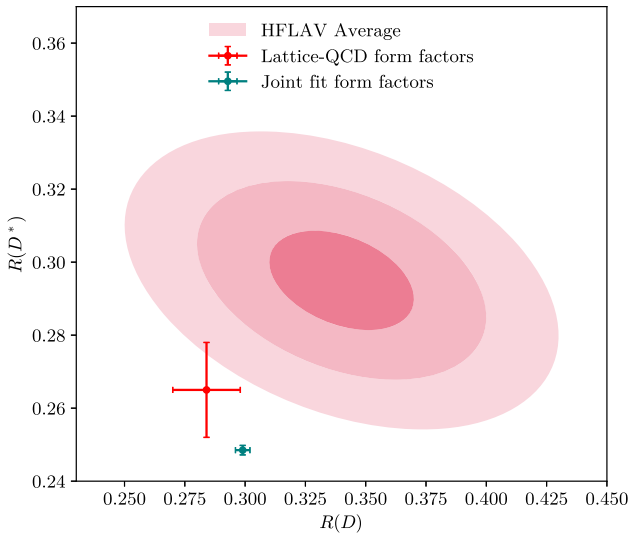

Fig. 11 Current status of the evaluations of $R(D)$ and $R(D^*)$. The contours show the 2019 HFLAV experimental average [1] at 1, 2, and 3σ. The red point with error bars uses our lattice-QCD prediction for $R(D^*)$ and $R(D) = 0.284(14)$ calculated from lattice-QCD data [53] (without correlations, which are not available but could be important). The green point shows $R(D^*)$ from the joint fit yielding $|V_{cb}|$ with $R(D) = 0.299(3)$ from HFLAV [1], which is similarly based on a joint fit to lattice-QCD and experimental data

5.4.2 *The z expansion with an improved CLN parametrization*

For the sake of completeness, we offer an alternative analysis, replacing the BGL parametrization with CLN. In the CLN parametrization, the form factor h_{A_1} is expressed as a polynomial in z, and the other form factors appear as ratios with respect to h_{A_1}:

$$R_0 = \frac{1}{1+r}\left(w + 1 + w\frac{rh_{A_2} - h_{A_3}}{h_{A_1}} - \frac{h_{A_2} - rh_{A_3}}{h_{A_1}}\right), \tag{79}$$

$$R_1 = \frac{h_V}{h_{A_1}}, \tag{80}$$

$$R_2 = \frac{rh_{A_2} + h_{A_3}}{h_{A_1}}. \tag{81}$$

Table 14 Comparison of results of the z expansion fits with and without the kinematic constraint given in Eq. (73). The largest differences appear in the higher-order coefficients. The coefficient c_0 is fixed by the constraint given in Eq. (72), and it is shown for convenience. The Coulomb factors are included in the fits with experimental data input

	Lattice QCD alone		Lattice QCD + Experiment	
	Unconstrained	Constrained	Unconstrained	Constrained
a_0	0.0330(12)	0.0330(12)	0.0321(10)	0.0320(10)
a_1	− 0.156(55)	− 0.156(55)	− 0.147(31)	− 0.142(31)
a_2	− 0.12(98)	− 0.12(98)	− 0.63(20)	− 0.60(19)
b_0	0.01229(23)	0.01228(23)	0.01249(22)	0.01250(22)
b_1	− 0.003(12)	0.001(11)	0.0021(43)	0.0031(42)
b_2	0.07(53)	− 0.23(41)	0.07(11)	0.05(10)
c_0	0.002059(38)	0.002058(38)	0.002092(37)	0.002094(37)
c_1	− 0.0058(25)	− 0.0051(23)	0.00062(86)	0.00090(84)
c_2	− 0.013(91)	− 0.076(58)	0.060(26)	0.061(26)
c_3			− 0.94(48)	− 1.08(47)
d_1	0.0509(15)	0.0509(15)	0.0531(14)	0.0531(14)
d_1	− 0.328(67)	− 0.331(67)	− 0.201(42)	− 0.187(41)
d_2	− 0.02(96)	0.12(95)	0.0007(8980)	0.85(67)
χ^2/dof	0.63/1	1.36/2	126/84	127/85
$\sum_i^N a_i^2$	$0.47^{+1.51}_{-0.36}$	$0.47^{+1.51}_{-0.36}$	$0.43^{+0.30}_{-0.21}$	$0.39^{+0.28}_{-0.20}$
$\sum_i^N (b_i^2 + c_i^2)$	$0.13^{+0.44}_{-0.10}$	$0.12^{+0.33}_{-0.16}$	$0.90^{+1.17}_{-0.57}$	$1.18^{+1.25}_{-0.75}$
$\sum_i^N d_i^2$	$0.54^{+1.40}_{-0.38}$	$0.51^{+1.42}_{-0.33}$	$0.41^{+1.25}_{-0.33}$	$0.77^{+1.60}_{-0.72}$
$\lvert V_{cb}\rvert_c$			38.40(78)	38.36(78)
$R(D^*)$	0.265(13)	0.274(10)	0.2484(13)	0.2492(12)

We include a few improvements to address the weak points of CLN: first we extract the full covariance matrix relating the parameters ρ_{A_1} and c_{A_1} from the original article [8] using the data given for the one-sigma ellipsoids. With the full covariance matrix, we can account for the strong correlations between $\rho^2_{A_1}$, c_{A_1} and d_{A_1}, which allows for small variations of the fixed relations often used in CLN fits. Second, we use updated results for the expansions in $w-1$ of the ratios R_j. The form factors to be fit are

$$h_{A_1}(z) = h_{A_1}(1)\Big[1 - 8\rho^2_{A_1} z + \left(64c_{A_1} - 16\rho^2_{A_1}\right) z^2 + \left(512d_{A_1} + 256c_{A_1} - 24\rho^2_{A_1}\right) z^3\Big], \tag{82}$$

$$R_0(w) = 1.25(35) - 0.183(77)(w-1) + 0.063(23)(w-1)^2, \tag{83}$$

$$R_1(w) = 1.28(36) - 0.101(51)(w-1) + 0.066(24)(w-1)^2, \tag{84}$$

$$R_2(w) = 0.740(44) + 0.128(38)(w-1) - 0.079(19)(w-1)^2, \tag{85}$$

where to fit R_0 to experimental data requires measurements of the τ final state. The full correlation matrix of $\rho^2_{A_1}$ and c_{A_1} is given in Table 15, and we follow Ref. [8] to calculate d_{A_1}. In the following, we refer to the CLN parametrization as "improved" when using Eqs. (82)–(85) and "base" when using the coefficients of the original paper [8].

Fitting our lattice data with the improved CLN parametrization yields a result for $R(D^*)$ that is compatible with Eq. (77), but the χ^2/dof of the fit increases spectacularly to $\chi^2/\text{dof} = 25.7/1$. A base CLN fit is similarly bad, with $\chi^2/\text{dof} = 31.9/7$. The combined fit using lattice and Belle data again yields a compatible $|V_{cb}|$ and is also very bad, with $\chi^2/\text{dof} = 133.5/80$. Here too the base and the improved versions of CLN are equally incompatible with the lattice-QCD form factors and Belle data. The CLN fits involving only lattice-QCD data violate the kinematic constraint in Eq. (73) by 2.7σ. The combined CLN fits satisfy it at around 1σ. In light of these issues, the BGL parametrization provides a much superior z expansion, so we have chosen it in our main analysis.[7]

One can wonder why the improved CLN fit performs so poorly. One possible point of tension between our data and the improved CLN ansatz is the relationship between the slope, the curvature and the cubic coefficient in h_{A_1}, which is much more constrained than in the BGL parametrization. We compare our fit results to only lattice data for both parametrizations by calculating a Taylor expansion around $z = 0$ up to cubic order of our BGL result for $h_{A_1} \propto f$, and we present in Fig. 12 contour plots of the CLN priors, the improved CLN fit result, and the BGL fit result. Our BGL results for $\rho^2_{A_1}$, c_{A_1} and d_{A_1} are compatible with our improved

[7] The fits discussed include the Coulomb factor. Its removal does not change the χ^2/dof significantly.

Table 15 Correlation matrix relating $\rho^2_{A_1}$ and c_{A_1} in the CLN parametrization

	$\rho^2_{A_1}$	c_{A_1}
$\rho^2_{A_1}$	1.0000	0.9834
c_{A_1}	0.9834	1.0000

CLN results within one sigma, suggesting that the tensions with the improved CLN parametrization come from the other form factors.

Figures 13 and 14 show the results for h_{A_1} and $R_{0,1,2}$ in the BGL and the improved CLN cases, along with the lattice data used in the fit. In general, the BGL fit shows a much better agreement with our lattice data, and the ratios $R_{0,2}$ are poorly fitted with the improved CLN ansatz. Also, the improved CLN fit violates the constraint given by Eq. (73). Imposing the constraint does not improve the fit quality, or reduce the tensions in the $R_{0,2}$ ratios.

Following Ref. [108], we advocate either a revision or a deprecation of the CLN parametrization in favor of a truly model-independent parametrization, such as BGL. Even if we had found a good quality of fit with CLN, we would still prefer the BGL results on theoretical grounds. The main reason not to use the CLN parametrization in a first principles analysis like this one is to test the theoretical assumptions that makes CLN different from BGL. CLN imposes constraints through crossing symmetry from the physical region of the cross channel and constraints from heavy-quark effective theory. The BGL parametrization does not. With the latter we are able to verify *ex post facto* the reliability of these constraints.

Other limitations of the base CLN, such as an update of its inputs, and a more careful treatment of the errors and correlations of the CLN coefficients, have been addressed in a few recent papers [119–121]. The HQET parametrization discussed in those works also include corrections of order $1/m_c^2$. We refer the reader to those papers for further discussion.

5.4.3 Comparison with LCSR

We can also test the validity of the light cone sum rules (LCSR), often employed to constrain the form factors at maximum recoil. To this end, we take the latest results from Ref. [122]. They present results for the form factors in a different notation. For the reader's convenience, we provide the conversion formulae:

$$V\,(w_{\rm Max}) = 0.69(13), \quad V\,(w) = h_V(w)\frac{1+r}{2\sqrt{r}}, \tag{86}$$

$$A_1(w_{\rm Max}) = 0.60(9), \quad A_1(w) = h_{A_1}(w)\frac{(1+w)\sqrt{r}}{1+r}, \tag{87}$$

$$A_2(w_{\rm Max}) = 0.51(9),$$

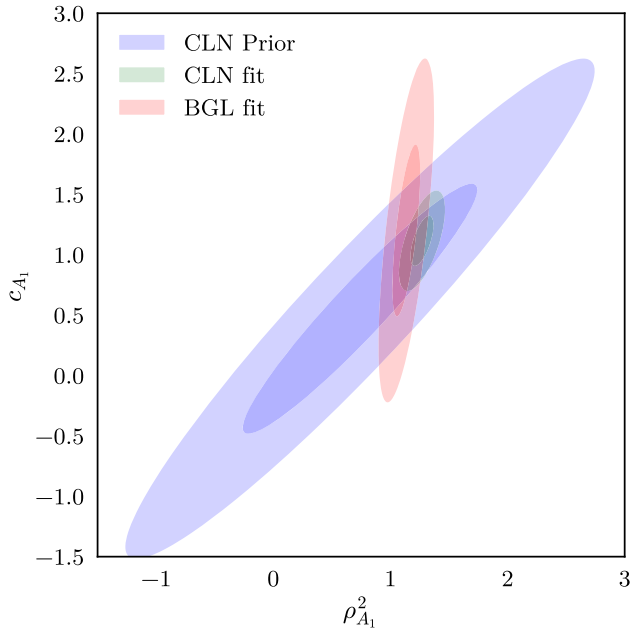

Fig. 12 Contour plot of the slope ($\rho^2_{A_1}$) versus curvature (c_{A_1}) of the h_{A_1} form factor in the CLN and BGL cases. We show the ellipsoids corresponding to the CLN priors extracted from [8] and the results of the improved CLN and BGL fits at one and two sigmas. There is good agreement in all cases, and a numerical test shows that the differences in $\rho^2_{A_1}$, c_{A_1} and d_{A_1} between the fit results of the improved CLN parametrization and the BGL parametrization is close to 1σ

$$A_2(w) = \left(r h_{A_2}(w) + h_{A_3}(w)\right)\frac{1+r}{2\sqrt{r}}. \tag{88}$$

Our lattice-QCD only results for the aforementioned form factors are:

$$V\,(w_{\rm Max}) = 0.65(10), \tag{89}$$

$$A_1(w_{\rm Max}) = 0.608(71), \tag{90}$$

$$A_2(w_{\rm Max}) = 0.71(11). \tag{91}$$

The agreement is excellent for A_1 and V, and the A_2 form factor also agrees within 1.4σ.

6 Discussion and outlook

Using the first unquenched lattice-QCD calculation of the form factors describing the decay $B \to D^*\ell\nu$ at nonzero recoil, together with 2018 Belle [18] and 2019 BaBar [19] measurements, we obtain the following results for the CKM matrix element $|V_{cb}|$:

$$|V_{cb}| = (38.40 \pm 0.78) \times 10^{-3}, \tag{92}$$

which includes the Coulomb correction for neutral B-meson decays. Omitting this correction leads to $|V_{cb}| = (38.74 \pm$

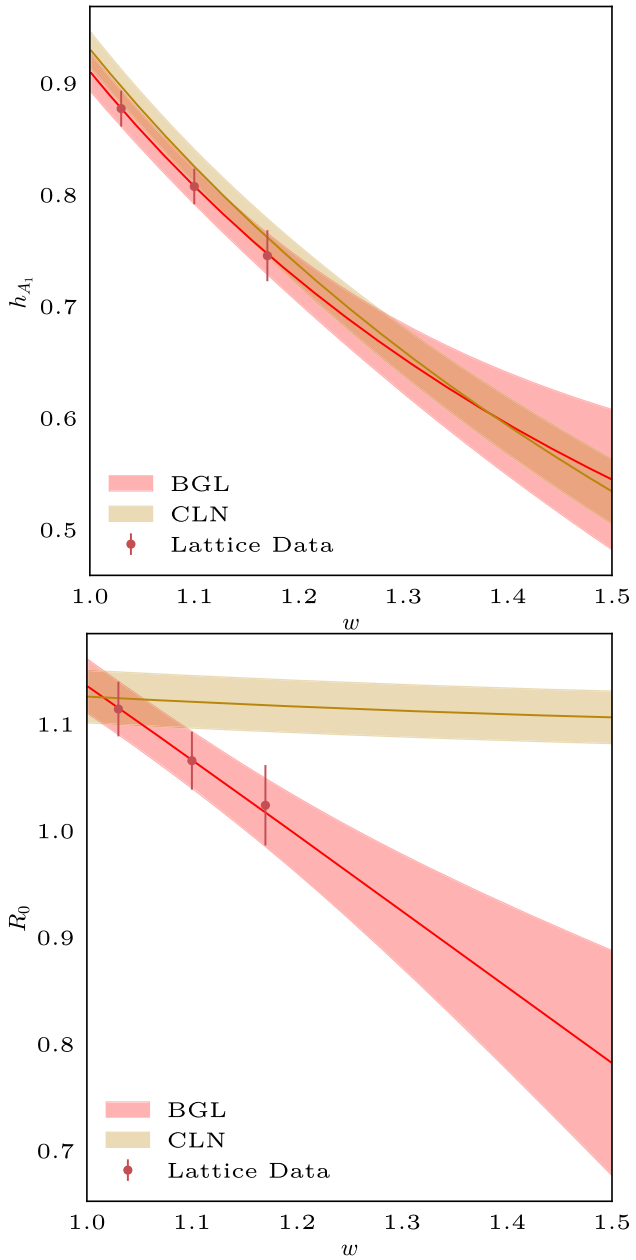

Fig. 13 Comparison between our fits to lattice data using BGL and the improved CLN parametrization. We show h_{A_1} (top) and R_0 (bottom). The differences between BGL and CLN are mild for h_{A_1}, at least in the region where there are data. In contrast, the improved CLN fit to R_0 shows large tensions with the lattice data

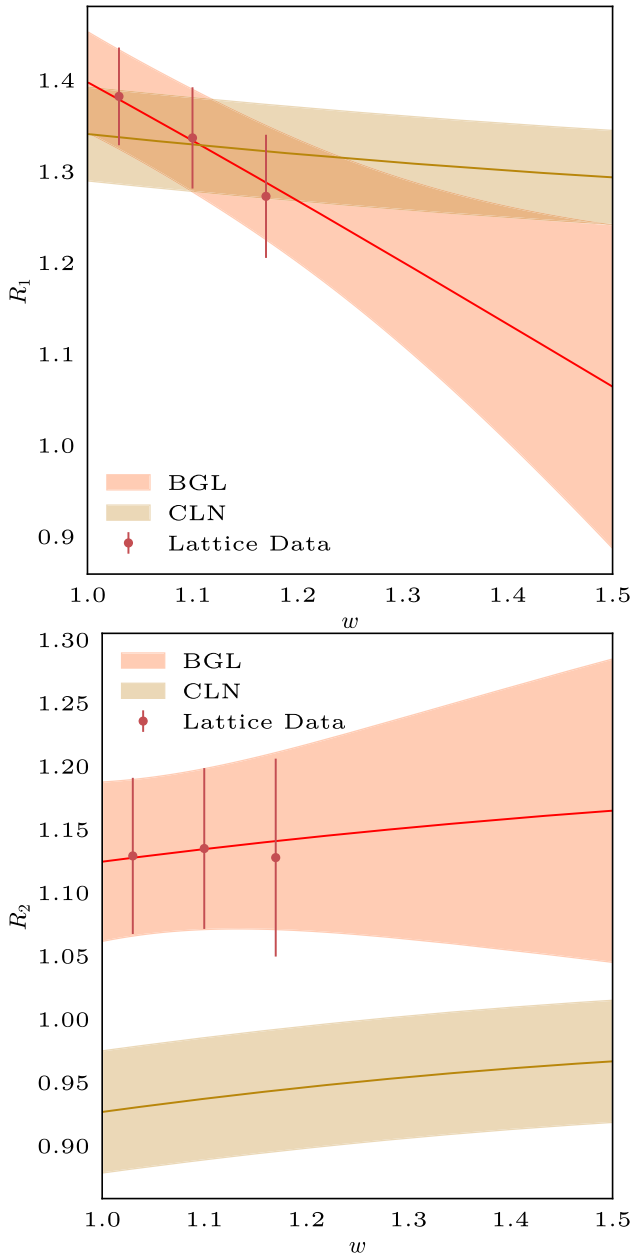

Fig. 14 Comparison between our fits to lattice data using BGL and the improved CLN parametrization. Here we show R_1 (top) and R_2 (bottom). In this case, the fit to the R_1 data looks acceptable in both cases, but the CLN fit to R_2 shows large differences with the data

$0.78) \times 10^{-3}$, which is the result that should be compared with inclusive decays. The quoted uncertainty is from experiment and theory together. As discussed in Sect. 5.2, this result comes from a fit exhibiting tension among the datasets, $\chi^2/\text{dof} = 126/84$. The tension is motivation for better experimental measurements and lattice-QCD calculations, as a $\sim 2\sigma$ effect it is inconclusive. In order to disentangle the contributions from experiment and theory, we run a BGL fit as described in Sect. 5.1, but assuming very small errors on the synthetic lattice-QCD data, and then we run analogous fits assuming very small experimental errors. By looking at how the final error changes, we estimate each contribution, finding 0.34×10^{-3} from experiment, 0.67×10^{-3} from lattice QCD, 0.10×10^{-3} from the truncation of the BGL expansion, and 0.18×10^{-13} due to EW+EM effects. These partial values are an approximation to guide how much improvement we can

expect from future calculations and experiments. The final error in Eq. (92) is an accurate estimate of the full uncertainty.

Belle II is expected to deliver experimental data for this decay, although until there is a better understanding of its detector performance and the systematics of the experiment, it is not clear how much improvement on the error for $|V_{cb}|$ could come from these high statistics data [123]. An improved result can be obtained from this or any other forthcoming data, in combination with the synthetic data points and covariance matrix that fully describe the output of our chiral-continuum extrapolation, given in the ancillary files as described in Appendix D.

The main result of this article is the behavior of the form factors parametrizing semileptonic $B \to D^*\ell\nu$ decays at small, but nonzero recoil. It allows us to perform a much more robust analysis of $|V_{cb}|$ than when using just the value of the decay amplitude at zero recoil. We find excellent agreement between our current and previous results, both from $B \to D^*\ell\nu$ at zero recoil [22] and $B \to D\ell\nu$ at nonzero recoil [53]. Our result also agrees with other recent $B \to D^{(*)}\ell\nu$ exclusive determinations [32,124]. It is also compatible with the recent extraction based on $B_s \to D_s^{(*)}\ell\nu$ decays measured by LHCb [125], and with form factors recently calculated by HPQCD [23,126], but with much smaller errors.

Our result for $|V_{cb}|$ does not change the current status of the inclusive-exclusive puzzle. The inclusive determination was recently updated in [127], slightly reducing the uncertainty, and preliminary results from Belle based on a data-driven approach [128], and compatible with the current inclusive world average [1], have been presented [129]. On the other hand, it has been argued that it is very challenging for BSM physics to accommodate such a tension [3,4]. A further reduction in errors is necessary to extract conclusions. Inclusive calculations of $|V_{cb}|$ might also benefit in the longer term from recent ideas in lattice QCD [130,131], which might lead to calculations with very different systematics.

During the implementation of the joint experimental-lattice fits we found issues in both experimental datasets used: the BaBar Collaboration does not provide unfolded data, and their z expansion uses only five coefficients to describe three form factors, as opposed to the nine coefficients we use for the same form factors in our lattice data only fit. As a result, we are concerned that the truncation errors in the z expansion could be underestimated in the synthetic data generated from BaBar fits. However, the $|V_{cb}|$ and $R(D^*)$ fits are dominated by the Belle and lattice-QCD data, and the effect of including BaBar data is just a small reduction in the total error in Eq. (92). The Belle Collaboration provides unfolded data, but as pointed out in Ref. [116], the statistical correlation matrices given in Ref. [18] seem inconsistent. We also checked that this potential problem has no significant impact on our results for either $|V_{cb}|$ or $R(D^*)$ (with form factors from the $|V_{cb}|$ fit). Thus, an improvement in the presentation of the data from both collaborations would be very welcome.

Another benefit of the knowledge of all form factors at nonzero recoil is the possibility of calculating $R(D^*)$ from first principles. Our result

$$R(D^*)_{\text{Lat}} = 0.265 \pm 0.013 \tag{93}$$

reaches a similar precision to that of the $B \to D\ell\nu$ analysis for $R(D)$. Even though our calculation of $R(D^*)$ involves integrals of extrapolated quantities with large errors, the combined error is relatively small due to the large correlations between $\mathcal{B}(B \to D^*\tau\nu)$ and $\mathcal{B}(B \to D^*\ell\nu)$. In this case, the form factor that enter only in the determination of the decay to a τ was computed with lattice input only. We have also calculated a more precise value using form factors from the $|V_{cb}|$ fit, $R(D^*)_{\text{Lat+Exp}} = 0.2484(13)$. Our preferred SM value is the one given in Eq. (93), which comes exclusively from lattice QCD and avoids any experimental decay-rate input.

The result in Eq. (93) confirms previous theoretical estimates of $R(D^*)$, as well as the current tension between the SM and experiment in the $R(D)$-$R(D^*)$ plane. Recent experimental determinations of $R(D^*)$ tend to reduce the tension, however. In fact, before Belle published results from their untagged dataset [29], the tension was as large as 4σ, but the newest analysis has reduced it to 3σ, and remaining tensions come mainly because of the influence of the BaBar $R(D)$ result [132]. An updated measurement of $R(D)$ could cast some light on the current tensions. Also, future high-precision experimental measurements from Belle II and LHCb are bound to become critical to determine whether these quantities will agree with the SM in the end.

Together with more precise experimental measurements, lattice-QCD form factors with a smaller uncertainty will also be crucial for shedding light to this theory-experiment tension, as well as to the exclusive–inclusive tension in the determination of $|V_{cb}|$. We expect to reduce the uncertainties in the form factors at nonzero recoil in future work, the first of which is already in progress at the time of this publication. The main sources of uncertainty in this work come from statistics and the quark discretization errors. An improvement in this area will require a modification in both the light and the heavy-quark actions to allow for smaller systematics. Chiral-continuum-extrapolation errors can be reduced by using a better discretization for light quarks and physical pion masses. Another area where we can reduce errors is the renormalization, by using a nonperturbative calculation of the renormalization factors. Further validation of our lattice-QCD result will come with independent analyses currently in progress by other lattice-QCD collaborations [133]. Similar improvements, with an expected reduction of errors, can be applied to our calculation of $B \to D\ell\nu$ form factors at nonzero recoil [53]. A correlated analysis of both decays

will allow a correlated determination of $R(D)$ and $R(D^*)$ that could provide tighter theoretical constraints.

In this work, we also assess the impact of using an improved CLN parametrization to describe the shape of the form factors. Instead of using the fixed coefficients published in Ref. [8], we employ the full covariance matrix that relates the slope and curvature coefficients of the reference form factor using data from the original CLN paper [8], and pass this information as priors to a CLN fit. We also compute the errors in the cubic coefficient, and update the values of the ratios with respect to the reference form factor with values coming from one of the latest HQET calculations [15], assuming a 20% error on each coefficient. Our updated CLN parametrization gives a very similar central value and error bar, compared with that of the BGL parametrization, but the quality of fit decreases greatly when the lattice-QCD data are included. CLN is very restrictive with the shape of certain form factors, and because the lattice-QCD data have relatively small errors, they introduce serious constraints in the parametrization. Our findings reinforce the current consensus of the community [108] to abandon CLN in favor of the more flexible and rigorous BGL parametrization. The impact of using improved HQE parametrizations, such as the one in Refs. [119–121], should be nevertheless investigated.

Acknowledgements We thank Claude Bernard for early contributions to this project, in particular for laying the foundations for the chiral-continuum extrapolations employed in this analysis. We thank Biplab Dey, Danny van Dyk, Paolo Gambino, Martin Jung, Laurent Lellouch, William Marciano, Christoph Schwanda, Phillip Urquijo, and Eiasha Waheed for useful discussions. We thank Biplab Dey especially for providing additional information on the BaBar $B \to D^* \ell \nu$ analysis [19], and Martin Jung for his careful reading of the manuscript. We thank Jon Bailey and Chris Bouchard for generating some of the correlator data. Computations for this work were carried out with resources provided by the USQCD Collaboration, the National Energy Research Scientific Computing Center, and the Argonne Leadership Computing Facility, which are funded by the Office of Science of the U.S. Department of Energy; and with resources provided by the National Institute for Computational Science and the Texas Advanced Computing Center, which are funded through the National Science Foundation's Teragrid/XSEDE Program. This work was supported in part by the U.S. Department of Energy under Awards No. DE-FG02-13ER41976 (D.T.), No. DE-SC0009998 (J.L.), No. DE-SC0010120 (S.G.), and No. DE-SC0015655 (A.X.K.,Z.G.); by the U.S. National Science Foundation under Grants No. PHY17-19626 (C.D., A.V.), and No. PHY14-17805 (J.L.); by SRA (Spain) under Grant No. PID2019-106087GB--C21/10.13039/501100011033 (E.G.); by the Consejería de Economía, Innovación, Ciencia y Empleo, Junta de Andalucía (Spain) under Grants No. FQM- 101, A-FQM-467-UGR18, and P18-FR-4314 (FEDER) (E.G.); by the Fermilab Distinguished Scholars Program (A.X.K.). This document was prepared by the Fermilab Lattice and MILC Collaborations using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359.

Data Availability Statement This manuscript has data included as electronic supplementary material. The online version of this article contains supplementary material, which is available to authorized users.

Funded by SCOAP[3]. SCOAP[3] supports the goals of the International Year of Basic Sciences for Sustainable Development.

Appendix A: The chiral logs in our chiral-continuum extrapolation

In this Appendix, we give further details on the chiral-continuum extrapolation. In particular, we discuss the staggered version of the chiral logs corresponding to the $B \to D^* \ell \nu$ case at nonzero recoil.

The chiral logs employed in Eq. (50) are derived following Refs. [87,134]. Since in this analysis we do not use partially quenched ensembles, the expressions employed are those of full QCD with 2 + 1 flavors. Here we reproduce the expression of the logs for the different form factors for the sake of completeness,

$$\begin{aligned}\log s^{Y}_{SU(3)} = \frac{1}{16}\sum_{\Theta}\left(2\bar{F}^{Y}_{\pi_\Theta} + \bar{F}^{Y}_{K_\Theta}\right) - \frac{1}{2}\bar{F}^{Y}_{\pi_I} + \frac{1}{6}\bar{F}^{Y}_{\eta_I} \\ + \sum_{\Xi=V,A} a^2\delta'_\Xi\Bigg(\frac{m^2_{S_\Xi} - m^2_{\pi_\Xi}}{(m^2_{\eta_\Xi} - m^2_{\pi_\Xi})(m^2_{\pi_\Xi} - m^2_{\eta'_\Xi})}\bar{F}^{Y}_{\pi_\Xi} \\ + \frac{m^2_{\eta_\Xi} - m^2_{S_\Xi}}{(m^2_{\eta_\Xi} - m^2_{\eta'_\Xi})(m^2_{\eta_\Xi} - m^2_{\pi_\Xi})}\bar{F}^{Y}_{\eta_\Xi} \\ + \frac{m^2_{S_\Xi} - m^2_{\eta'_\Xi}}{(m^2_{\eta_\Xi} - m^2_{\eta'_\Xi})(m^2_{\eta'_\Xi} - m^2_{\pi_\Xi})}\bar{F}^{Y}_{\eta'_\xi}\Bigg). \end{aligned} \tag{A.1}$$

In the expression above $Y = A_{1,2,3}$, and for the vector form factor $\log s^{V}_{SU(3)} = \log s^{A_1}_{SU(3)}$. The index Θ goes over all pseudoscalar meson fields of the effective theory, whereas Ξ labels vector and axial counterparts. Explicit expressions for the mass of a pseudoscalar meson P with taste Ξ, m_{P_Ξ}, in terms of the parameters of the rooted staggered theory can be found in Ref. [135]. The hairpin parameters, which come from χPT disconnected diagrams, are marked as $\delta'_{V,A}$. Both vector and axial hairpin parameters are defined for the lightest coarse ensemble $a \approx 0.12$ fm to be $\delta'_A = -0.28(6)$ and $\delta'_V = 0.00(7)$, and their value is obtained for other ensembles by rescaling this number, assuming the hairpin parameter scales as the root mean square of the taste splittings. The taste

splittings along with the tree-level LEC B_0 depend solely on the lattice spacing and are given in Table 16. The values for the hairpin parameters, as well as those quoted in Table 16, were determined from fits to light-quark quantities by the MILC Collaboration in Ref. [54]. Our results are insensitive to the errors in the parameters quoted in the table, as well as the errors in the hairpin parameters, since the chiral logs are a very small contribution to our fit function in the region where we have data.

The functions appearing in Eq. (A.1) are given by

$$\bar{F}_j^Y \equiv F^Y(w, m_j, -\Delta^{(c)}/m_j), \tag{A.2}$$

where $\Delta^{(c)}$ is the D-D^* mass splitting that we take from the PDG [2], assuming a charged meson. The error from the PDG uncertainty, even if enlarged to cover a neutral-meson mass difference, also has a negligible impact in the final fit. The F^Y functions are defined as

$$\begin{aligned} F^{A_1}(w,m,x) = -2\Big[& I_1(w,m,x) - \frac{1}{2} I_3(w,m,x) \\ & + (w+1) I_1(w,m,0) \\ & + (w^2-1) I_2(w,m,0) \\ & - \frac{5}{2} I_3(w,m,0) \Big], \end{aligned} \tag{A.3}$$

$$\begin{aligned} F^{A_2}(w,m,x) = -2\Big[& I_1(w,m,x) + (w+1) I_2(w,m,x) \\ & - I_1(w,m,0) - (w+1) I_2(w,m,0) \Big], \end{aligned} \tag{A.4}$$

$$\begin{aligned} F^{A_3}(w,m,x) = -2\Big[& -(w+1) I_2(w,m,x) \\ & - \frac{1}{2} I_3(w,m,x) + (w+2) I_1(w,m,0) \\ & + w(w+1) I_2(w,m,0) - \frac{5}{2} I_3(w,m,0) \Big], \end{aligned} \tag{A.5}$$

with

$$\begin{aligned} I_j(w,m,x) = -\Big[& m^2 x E_j(w) + m^2 x^2 \ln\left(\frac{m^2}{\Lambda_\chi^2}\right) G_j(w) \\ & + m^2 x^2 F_j(w,x) \Big]. \end{aligned} \tag{A.6}$$

The functions $E_j(w)$ and $G_j(w)$ are

$$E_1(w) = \frac{\pi}{w+1}, \quad G_1(w) = \frac{r(w)-w}{2(w^2-1)}, \tag{A.7}$$

$$E_2(w) = \frac{-\pi}{(w+1)^2}, \quad G_2(w) = \frac{w^2+2-3wr(w)}{2(w^2-1)^2}, \tag{A.8}$$

$$E_3(w) = \pi, \quad G_3(w) = -1. \tag{A.9}$$

Here $r(w)$ is

$$r(w) = \frac{1}{\sqrt{w^2-1}} \ln(w + \sqrt{w^2-1}). \tag{A.10}$$

The F_j functions in general cannot be expressed in closed form, and are defined as integrals,

$$\begin{aligned} F_1(w,x) = \frac{1}{x^2} \int_0^{\pi/2} d\theta \frac{a}{1+w\sin 2\theta} \Big\{ & \pi\left(\sqrt{1-a^2}-1\right) \\ & -2\left[\frac{1}{2}\sqrt{a^2-1}\ln\left(1-2a(a+\sqrt{a^2-1})\right) - a\right] \Big\}, \end{aligned} \tag{A.11}$$

$$\begin{aligned} F_2(w,x) = \frac{1}{x^2} \int_0^{\pi/2} d\theta \frac{a\sin 2\theta}{(1+w\sin 2\theta)^2} & \\ \times \Bigg\{ & -\frac{3\pi}{2}\left(\sqrt{1-a^2}-1\right) + \frac{\pi a^2}{2\sqrt{1-a^2}} \\ & + \left[\left(\frac{3-4a^2}{\sqrt{a^2-1}}\right)\left(-\frac{1}{2}\ln\left(1-2a(a+\sqrt{a^2-1})\right)\right) \right. \\ & \left. -3a\right] \Bigg\}, \end{aligned} \tag{A.12}$$

$$\begin{aligned} F_3(w,x) = \frac{1}{x}\Big\{ & \pi\left(\sqrt{1-x^2}-1\right) - 2 \\ & \times \left[\frac{1}{2}\sqrt{x^2-1}\ln\left(1-2x(x+\sqrt{x^2-1})\right) - x\right] \Big\}, \end{aligned} \tag{A.13}$$

where

$$a = \frac{x\cos\theta}{\sqrt{1+w\sin 2\theta}}. \tag{A.14}$$

Appendix B: Estimation of the matching factors and their errors

In this Appendix, we derive the renormalization factors for the vector and axial-vector currents using heavy-quark effective theory (HQET) as an intermediary between lattice gauge theory and continuum QCD [50,103]. We have calculated the renormalization factors at one loop in perturbation theory, sometimes using an expedient approximation described below. These one-loop results are shown in Table 6. Here, we use the notation $p_B = p$, $v_B = v$, $p_{D^*} = p'$, and $v_{D^*} = v'$.

Appendix B.1: HQET matching

The HQET description of lattice-QCD currents is [50]

$$V^\mu \doteq \bar{C}_{V_\parallel}^{\mathrm{LGT}}(w) v^\mu \bar{c}_{v'} b_v + \bar{C}_{V_\perp}^{\mathrm{LGT}}(w) \bar{c}_{v'} i\gamma_\perp^\mu b_v$$

Table 16 Taste splittings and B_0 LEC used to compute the meson masses at tree level in χPT

a (fm)	$r_1^2 a^2 \Delta_I$	$r_1^2 a^2 \Delta_V$	$r_1^2 a^2 \Delta_T$	$r_1^2 a^2 \Delta_A$	$r_1 B_0$
0.15	0.9851	0.7962	0.6178	0.3915	6.761
0.12	0.6008	0.4803	0.3662	0.2270	6.832
0.09	0.2207	0.1593	0.1238	0.0747	6.639
0.06	0.0704	0.0574	0.0430	0.0263	6.487
0.045	0.0278	0.0227	0.0170	0.0104	6.417
Continuum	–	–	–	–	6.015

$$+ \bar{C}^{\mathrm{LGT}}_{V_{v'}}(w) v'^{\mu}_{\perp} \bar{c}_{v'} b_v, \tag{B.15}$$

$$A^{\mu} \doteq \bar{C}^{\mathrm{LGT}}_{V_\perp}(w) \bar{c}_{v'} i\gamma^{\mu}_{\perp}\gamma^5 b_v - \bar{C}_{V_\parallel}(w)^{\mathrm{LGT}} v^{\mu} \bar{c}_{v'}\gamma^5 b_v - \bar{C}^{\mathrm{LGT}}_{V_{v'}}(w) v'^{\mu}_{\perp} \bar{c}_{v'}\gamma^5 b_v. \tag{B.16}$$

where $\doteq$ means "has the same matrix elements as". The currents on the left-hand side are lattice operators, while the bilinears on the right-hand side are HQET operators (e.g., built from fields satisfying $\not{v} b_v = i b_v$). Similarly, HQET can be used to describe continuum-QCD currents.

$$\mathcal{V}^{\mu} \doteq \bar{C}_{V_\parallel}(w) v^{\mu} \bar{c}_{v'} b_v + \bar{C}_{V_\perp}(w) \bar{c}_{v'} i\gamma^{\mu}_{\perp} b_v + \bar{C}_{V_{v'}}(w) v'^{\mu}_{\perp} \bar{c}_{v'} b_v, \tag{B.17}$$

$$\mathcal{A}^{\mu} \doteq \bar{C}_{V_\perp}(w) \bar{c}_{v'} i\gamma^{\mu}_{\perp}\gamma^5 b_v - \bar{C}_{V_\parallel}(w) v^{\mu} \bar{c}_{v'}\gamma^5 b_v - \bar{C}_{V_{v'}}(w) v'^{\mu}_{\perp} \bar{c}_{v'}\gamma^5 b_v, \tag{B.18}$$

The only difference between the lattice and the continuum currents lies in the short-distance coefficients; for Wilson-like fermions, as used here,

$$\lim_{a\to 0} \bar{C}^{\mathrm{LGT}}_{J}(w) = \bar{C}_J(w). \tag{B.19}$$

Equations (B.15)–(B.18) are valid at zeroth order in $1/m_Q$, $Q = c, b$, which suffices here.[8]

To continue with the analysis, it is advantageous to write explicitly the velocity and polarization vectors of our mesons. Assuming the B meson to be at rest, and the D^* to have momentum p' along the z axis:

$$v = (0, 0, 0, i) = p/M_B, \tag{B.20}$$

$$v' = (0, 0, \sqrt{w^2-1}, iw) = p'/M_{D^*}, \tag{B.21}$$

$$v'_\perp = (0, 0, \sqrt{w^2-1}, 0), \tag{B.22}$$

$$\hat{v}'_\perp \equiv v'_\perp/(v'_\perp \cdot v'_\perp)^{1/2} = (0, 0, 1, 0), \tag{B.23}$$

$$\epsilon^{\pm} = (\boldsymbol{\epsilon}^{\pm}, 0, 0), \tag{B.24}$$

$$\epsilon^3 = (0, 0, w, i\sqrt{w^2-1}), \tag{B.25}$$

$$\epsilon^s = v', \tag{B.26}$$

where $\boldsymbol{\epsilon}^{\pm}$ are two-component unit vectors in the xy plane. For momenta in other directions, such as $\boldsymbol{v}' \propto (1, 1, 0)$ or $(1, -1, 1)$, one can rotate the spatial components accordingly. Note that we pick $\hat{v}'_\perp$, and then we deduce $\epsilon^{\pm}$.

In this notation, $v \cdot v = -1$, $v' \cdot v' = -1$, and $v \cdot v' = -w$.[9] The polarization vectors satisfy $\bar{\epsilon}^m \cdot \epsilon^n = g^{mn}$ for $(m, n) \in \{+, -, 3, s\}$ with $g^{mn} = \mathrm{diag}(1, 1, 1, -1)$. The bar on a polarization vector means to complex-conjugate the spatial components, which arises if (as usual) $\epsilon^{\pm}$ corresponds to D^* helicity ± 1. (We use linear polarizations with real $\boldsymbol{\epsilon}$ but keep the $\pm$ notation.) The polarization vectors satisfy $v' \cdot \epsilon^{\pm,3} = 0$, and also $v \cdot \epsilon^{\pm} = 0$.

These vectors can be used to isolate parts of the currents with different matching factors:

$$-v \cdot V \doteq \bar{C}^{\mathrm{LGT}}_{V_\parallel}(w) \bar{c}_{v'} b_v, \tag{B.27}$$

$$\hat{v}'_\perp \cdot V \doteq \bar{C}^{\mathrm{LGT}}_{V_\perp}(w) \bar{c}_{v'} i \not{\hat{v}}'_\perp b_v + \sqrt{w^2-1}\, \bar{C}^{\mathrm{LGT}}_{V_{v'}}(w) \bar{c}_{v'} b_v, \tag{B.28}$$

$$\epsilon^{\pm} \cdot V \doteq \bar{C}^{\mathrm{LGT}}_{V_\perp}(w) \bar{c}_{v'} i \not{\epsilon}^{\pm}_{\perp} b_v, \tag{B.29}$$

and similarly for $\mathcal{V}$ (without the superscript "LGT") and for A and $\mathcal{A}$ with $\bar{c}_{v'} \to \bar{c}_{v'}\gamma^5$.

In HQET, the matrix elements can be worked out with the "trace formalism" [43]:

$$\langle D^*|\bar{c}_{v'} \Gamma b_v|B\rangle = \sqrt{M_{D^*} M_B}\, \mathrm{tr}\left[\bar{\mathcal{M}}_{D^*} \Gamma \mathcal{M}_B\right] \xi(w), \tag{B.30}$$

where $\xi(w)$ is a form factor known as the "Isgur-Wise function", and

$$\mathcal{M}_B = -\frac{1}{2}(1 - i\not{v})\gamma^5, \qquad \bar{\mathcal{M}}_{D^*} = \bar{\not{\epsilon}}\frac{1}{2}(1 - i\not{v}'), \tag{B.31}$$

[8] Matching at the next order in $1/m_Q$ is possible in principle but cumbersome beyond the tree level.

[9] These conventions also hold in Minkowski space when using the metric $g^{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ [103]. NB: in Minkowski space, $\mu \in \{0, 1, 2, 3\}$; in Euclidean space, $g^{\mu\nu} = \delta^{\mu\nu}$ and $\mu \in \{1, 2, 3, 4\}$; $x^4 = ix^0$.

and similarly for $\bar{\mathcal{M}}_D$ and $\mathcal{M}_{D^*}$. Then

$$\mathrm{tr}\left[\bar{\mathcal{M}}_{D^*}\mathcal{M}_B\right] = 0, \tag{B.32a}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_{D^*} i\gamma_\perp^\mu \mathcal{M}_B\right] = \varepsilon^{\mu\nu\alpha\beta}\bar{\epsilon}_\nu v'_\alpha v_\beta, \tag{B.32b}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_{D^*}(-\gamma^5)\mathcal{M}_B\right] = i\bar{\epsilon}\cdot v, \tag{B.32c}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_{D^*} i\gamma_\perp^\mu\gamma^5\mathcal{M}_B\right] = -i\left[(w+1)\bar{\epsilon}_\perp^\mu + \bar{\epsilon}\cdot v v'_\perp\right], \tag{B.32d}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_D\mathcal{M}_B\right] = w+1, \tag{B.32e}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_D i\gamma_\perp^\mu\mathcal{M}_B\right] = v'^{\,\mu}_\perp, \tag{B.32f}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_D(-\gamma^5)\mathcal{M}_B\right] = 0, \tag{B.32g}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_D i\gamma_\perp^\mu\gamma^5\mathcal{M}_B\right] = 0, \tag{B.32h}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_{D^*}\mathcal{M}_{B^*}\right] = (w+1)\bar{\epsilon}'\cdot\epsilon + \bar{\epsilon}'\cdot v\,\epsilon\cdot v', \tag{B.32i}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_{D^*} i\gamma_\perp^\mu\mathcal{M}_{B^*}\right] = v'^{\,\mu}_\perp\bar{\epsilon}'\cdot\epsilon - \epsilon_\perp^\mu\bar{\epsilon}'\cdot v - \bar{\epsilon}'^{\,\mu}_\perp\epsilon\cdot v', \tag{B.32j}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_{D^*}(-\gamma^5)\mathcal{M}_{B^*}\right] = \varepsilon^{\nu\rho\alpha\beta}\bar{\epsilon}'_\nu\epsilon_\rho v'_\alpha v_\beta, \tag{B.32k}$$
$$\mathrm{tr}\left[\bar{\mathcal{M}}_{D^*} i\gamma_\perp^\mu\gamma^5\mathcal{M}_{B^*}\right] = (\delta^\mu_\tau + v^\mu v_\tau)\varepsilon^{\tau\nu\rho\alpha}\bar{\epsilon}'_\nu(v+v')_\alpha\epsilon_\rho. \tag{B.32l}$$

With these results, physical combinations of the decompositions in Eqs. (B.15)–(B.18) can be more easily tracked. Note that the sign convention for ε cancels in ratios introduced below.

These expressions can be used to relate matrix elements of continuum and LGT currents to each other, via HQET and Eqs. (B.15)–(B.18). Thus, $Z_V V$ and $Z_A A$ have the same matrix elements as $\mathcal{V}$ and $\mathcal{A}$ if one chooses matching factors such that

$$\bar{Z}_{V_\parallel}(w) v\cdot V \doteq v\cdot\mathcal{V}, \tag{B.33a}$$
$$\bar{Z}_{V_{v'}}(w)\hat{v}'_\perp\cdot V \doteq \hat{v}'_\perp\cdot\mathcal{V}, \tag{B.33b}$$
$$\bar{Z}_{V_\perp}(w)\epsilon^\pm\cdot V \doteq \epsilon^\pm\cdot\mathcal{V}, \tag{B.33c}$$
$$\bar{Z}_{A_\parallel}(w) v\cdot A \doteq v\cdot\mathcal{A}, \tag{B.33d}$$
$$\bar{Z}_{A_{v'}}(w)\hat{v}'_\perp\cdot A \doteq \hat{v}'_\perp\cdot\mathcal{A}, \tag{B.33e}$$
$$\bar{Z}_{A_\perp}(w)\epsilon^\pm\cdot A \doteq \epsilon^\pm\cdot\mathcal{A}. \tag{B.33f}$$

From these requirements one finds

$$\bar{Z}_{V_\parallel}(w) = \frac{\bar{C}_{V_\parallel}(w)}{\bar{C}^{\mathrm{LGT}}_{V_\parallel}(w)}, \tag{B.34a}$$
$$\bar{Z}_{V_{v'}}(w) = \frac{\bar{C}_{V_\perp}(w) + (w+1)\bar{C}_{V_{v'}}}{\bar{C}^{\mathrm{LGT}}_{V_\perp}(w) + (w+1)\bar{C}^{\mathrm{LGT}}_{V_{v'}}}, \tag{B.34b}$$
$$\bar{Z}_{V_\perp}(w) = \frac{\bar{C}_{V_\perp}(w)}{\bar{C}^{\mathrm{LGT}}_{V_\perp}(w)}, \tag{B.34c}$$
$$\bar{Z}_{A_\parallel}(w) = \frac{\bar{C}_{A_\parallel}(w)}{\bar{C}^{\mathrm{LGT}}_{A_\parallel}(w)}, \tag{B.34d}$$
$$\bar{Z}_{A_{v'}}(w) = \frac{\bar{C}_{A_\perp}(w) + (w-1)\bar{C}_{A_{v'}}}{\bar{C}^{\mathrm{LGT}}_{A_\perp}(w) + (w-1)\bar{C}^{\mathrm{LGT}}_{A_{v'}}}, \tag{B.34e}$$
$$\bar{Z}_{A_\perp}(w) = \frac{\bar{C}_{A_\perp}(w)}{\bar{C}^{\mathrm{LGT}}_{A_\perp}(w)}, \tag{B.34f}$$

In anticipation of one-loop perturbative calculations, it is convenient to define

$$\bar{\rho}_J(w) = \frac{\bar{Z}_J(w)}{\bar{Z}^{1/2}_{V_\parallel,\bar{b}b}(1)\bar{Z}^{1/2}_{V_\parallel,\bar{c}c}(1)} \tag{B.35}$$

to cancel conventional field-normalization factors as well as potentially large tadpole diagrams. The denominators can be computed nonperturbatively. With a quantitative method to compute matrix elements, one can obtain the matching factors. For example, one can use quark states and expand them in perturbative QCD.

Appendix B.2: Useful ratios

We start with the original double ratio

$$|R_{A_1}(1)|^2 = \frac{\langle D^*(\mathbf{0},\epsilon)|\epsilon\cdot A|B(\mathbf{0})\rangle\,\langle B(\mathbf{0})|\bar{\epsilon}\cdot A|D^*(\mathbf{0},\epsilon)\rangle}{\langle D^*(\mathbf{0},\epsilon)|v\cdot V|D^*(\mathbf{0},\epsilon)\rangle\langle B(\mathbf{0})|v\cdot V|B(\mathbf{0})\rangle}, \tag{B.36}$$

which requires the matching factor

$$\bar{\rho}^2_{A_\perp}(1) = \frac{\bar{Z}^2_{A_\perp}(1)}{\bar{Z}_{V_\parallel,\bar{b}b}(1)\bar{Z}_{V_\parallel,\bar{c}c}(1)}. \tag{B.37}$$

To obtain the w dependence of all four form factors, we define several further ratios:

$$Q_{A_1}(w) = \frac{\langle D^*(\boldsymbol{p},\epsilon)|\epsilon^\pm\cdot A|B(\mathbf{0})\rangle}{\langle D^*(\mathbf{0},\epsilon)|\epsilon^\pm\cdot A|B(\mathbf{0})\rangle}, \tag{B.38a}$$
$$X_0(w) = \frac{\langle D^*(\boldsymbol{p},\epsilon)|(-v\cdot A)|B(\mathbf{0})\rangle}{\langle D^*(\boldsymbol{p},\epsilon)|\epsilon^\pm\cdot A|B(\mathbf{0})\rangle}, \tag{B.38b}$$
$$X_1(w) = \frac{\langle D^*(\boldsymbol{p},\epsilon)|\hat{v}'_\perp\cdot A|B(\mathbf{0})\rangle}{\langle D^*(\boldsymbol{p},\epsilon)|\epsilon^\pm\cdot A|B(\mathbf{0})\rangle}, \tag{B.38c}$$
$$X_V(w) = \frac{\langle D^*(\boldsymbol{p},\epsilon)|\epsilon^\pm\cdot V|B(\mathbf{0})\rangle}{\langle D^*(\boldsymbol{p},\epsilon)|\epsilon^\pm\cdot A|B(\mathbf{0})\rangle}, \tag{B.38d}$$

which require, respectively, the matching factors

$$\frac{\bar{Z}_{A_\perp}(w)}{\bar{Z}_{A_\perp}(1)} = \frac{\bar{\rho}_{A_\perp}(w)}{\bar{\rho}_{A_\perp}(1)}, \tag{B.39a}$$
$$\frac{\bar{Z}_{A_\parallel}(w)}{\bar{Z}_{A_\perp}(w)} = \frac{\bar{\rho}_{A_\parallel}(w)}{\bar{\rho}_{A_\perp}(w)}, \tag{B.39b}$$
$$\frac{\bar{Z}_{A_{v'}}(w)}{\bar{Z}_{A_\perp}(w)} = \frac{\bar{\rho}_{A_{v'}}(w)}{\bar{\rho}_{A_\perp}(w)}, \tag{B.39c}$$

$$\frac{\bar{Z}_{V_\perp}(w)}{\bar{Z}_{A_\perp}(w)} = \frac{\bar{\rho}_{V_\perp}(w)}{\bar{\rho}_{A_\perp}(w)}, \tag{B.39d}$$

Note that the form factor $A_1(q^2) \propto h_{A_1}(w)$, defined in Eq. (87), comes directly from $R_{A_1} Q_{A_1}(w)$, and $V(q^2) \propto h_V(w)$, defined in Eq. (86), comes directly from $R_{A_1} Q_{A_1}(w) X_V(w)$, while the helicity amplitudes H_0 and H_s are linear combinations of $R_{A_1} Q_{A_1}(w) X_0(w)$ and $R_{A_1} Q_{A_1}(w) X_1(w)$. The helicity amplitudes $H_\pm$ come from A_1 and V, but it is presumably more convenient (or just as convenient) to keep the axial and vector parts separate.

For the dynamic velocity consider the vector-current matrix element

$$\begin{aligned}\langle D^*(\boldsymbol{p},\epsilon')|\mathcal{V}^\mu|D^*(\boldsymbol{0},\epsilon)\rangle \\ = \ \bar{\epsilon}'\cdot\epsilon(p'+p)^\mu f_1(w) \\ + \text{terms proportional to } v'\cdot\epsilon,\, v\cdot\epsilon'.\end{aligned} \tag{B.40}$$

If we take the same transverse polarization, $\epsilon' = \epsilon = \epsilon^\pm$, for the initial and final D^*s, then

$$\langle D^*(\boldsymbol{p})|(-v\cdot\mathcal{V})|D^*(\boldsymbol{0})\rangle = M_{D^*}(w+1)f_1(w), \tag{B.41}$$

$$\langle D^*(\boldsymbol{p})|\hat{v}'_\perp\cdot\mathcal{V}|D^*(\boldsymbol{0})\rangle = M_{D^*}\sqrt{w^2-1}\,f_1(w), \tag{B.42}$$

where $\hat{v}'_\perp$, defined in Eq. (B.23), is a unit vector in the direction of $\boldsymbol{p}$, such as $(0,0,1)$ or $(1,1,0)/\sqrt{2}$. On the right-hand side of the second equation, $\sqrt{w^2-1}$ is nothing but $|v'_\perp|$, the magnitude of $v'_\perp$; that is, $v'_\perp = \sqrt{w^2-1}\hat{v}'_\perp$. Because $\mathcal{V}$ is properly normalized, it measures the flavor charge, so the form factor satisfies
$f_1(1) = 1$.

Using the trace formalism, it is easy to show

$$\langle D^*(\boldsymbol{p})|(-v\cdot V)|D^*(\boldsymbol{0})\rangle = M_{D^*}(w+1)\bar{C}^{\mathrm{LGT}}_{V_\parallel}(w)\xi(w), \tag{B.43}$$

$$\begin{aligned}\langle D^*(\boldsymbol{p})|\hat{v}'_\perp\cdot V|D^*(\boldsymbol{0})\rangle = M_{D^*}\sqrt{w^2-1} \\ \left[\bar{C}^{\mathrm{LGT}}_{V_\perp}(w) + (w+1)\bar{C}^{\mathrm{LGT}}_{V_{v'}}(w)\right]\xi(w).\end{aligned} \tag{B.44}$$

Taking the ratio, the form factor drops out:

$$\begin{aligned}\frac{\langle D^*(\boldsymbol{p})|\hat{v}'_\perp\cdot V|D^*(\boldsymbol{0})\rangle}{\langle D^*(\boldsymbol{p})|(-v\cdot V)|D^*(\boldsymbol{0})\rangle} \\ = \frac{|v'_\perp|}{w+1}\frac{\bar{C}^{\mathrm{LGT}}_{V_\perp}(w) + (w+1)\bar{C}^{\mathrm{LGT}}_{V_{v'}}(w)}{\bar{C}^{\mathrm{LGT}}_{V_\parallel}(w)},\end{aligned} \tag{B.45}$$

but a matching factor remains.[10] The analogous equations for $\mathcal{V}$ shows that $f_1(w) = \bar{C}_{V_\parallel}(w)\xi(w)$ and $\bar{C}_{V_\parallel}(w) = \bar{C}_{V_\perp}(w) + (w+1)\bar{C}_{V_{v'}}(w)$.

Appendix B.3: Expedient approximation

Calculating the full w dependence of the $\bar{Z}_J$ at the one-loop level is, in general, very cumbersome. From Ref. [50], however, we have

$$\lim_{m_c a\to 0}\bar{Z}_{J_\parallel}(w) = Z_{J_\parallel}, \tag{B.46}$$

$$\lim_{m_c a\to 0}\bar{Z}_{J_\perp}(w) = Z_{J_\perp}, \tag{B.47}$$

$$\begin{aligned}\lim_{m_c a\to 0}\bar{Z}_{J_\perp}\bar{C}^{\mathrm{LGT}}_{J_{v'}}(w) = Z_{J_\perp}\bar{C}_{J_{v'}}(w) \\ \Rightarrow \lim_{m_c a\to 0}\bar{Z}_{J_{v'}}(w) = Z_{J_\perp}.\end{aligned} \tag{B.48}$$

Then we can approximate,

$$\frac{\bar{\rho}_{A_\perp}(w)}{\bar{\rho}_{A_\perp}(1)} \approx 1 \pm \alpha_V(q^*)\rho^{[1]}_{\max}(w-1)m_{2c}a, \tag{B.49a}$$

$$\frac{\bar{\rho}_{A_\parallel}(w)}{\bar{\rho}_{A_\perp}(w)} \approx \frac{\rho_{A_\parallel}}{\rho_{A_\perp}} \pm \alpha_V(q^*)\rho^{[1]}_{\max}m_{2c}a, \tag{B.49b}$$

$$\frac{\bar{\rho}_{A_{v'}}(w)}{\bar{\rho}_{A_\perp}(w)} \approx 1 \pm \alpha_V(q^*)\rho^{[1]}_{\max}m_{2c}a, \tag{B.49c}$$

$$\frac{\bar{\rho}_{V_\perp}(w)}{\bar{\rho}_{A_\perp}(w)} \approx \frac{\rho_{V_\perp}}{\rho_{A_\perp}} \pm \alpha_V(q^*)\rho^{[1]}_{\max}m_{2c}a, \tag{B.49d}$$

where $\rho^{[1]}_{\max} = 0.352$ is the largest one-loop coefficient that we find among the computable one-loop coefficients. For lack of a better choice, we set $q^* = 2/a$, as in other papers.

In the limit $m_c a \to 0$, the matching factor in the velocity tends to 1. We could use an uncertainty like that in Eq. (B.49c), but that seems to be an unecessary complication. A little algebra shows that the mismatch in w is very small:

$$w \approx w^{\mathrm{LGT}} \pm \alpha_V(q^*)\rho^{[1]}_{\max}(w^2-1)m_{2c}a, \tag{B.50}$$

while the mismatch in z is

$$z \approx z^{\mathrm{LGT}}\left[1 \pm \alpha_V(q^*)2\rho^{[1]}_{\max}\frac{1+z}{1-z}m_{2c}a\right]. \tag{B.51}$$

Appendix C: Heavy quark mistuning corrections procedure

We tune the heavy-quark masses (κ_b and κ_c) for clover quarks in the Fermilab interpretation using the procedure described in Ref. [22]. In brief, in a mass-independent scheme for the

[10] We overlooked this factor in previous papers [52,53,136].

Table 17 Simulation values κ_c' and κ_b' employed in the calculation of the form factors compared with their more precisely tuned values κ_c and κ_b [22]. The first error in κ_b and κ_c includes statistical and fitting contributions; the second one comes from scale setting

a (fm)	am_l/am_s	κ_b'	κ_b	κ_c'	κ_c
≈ 0.15	0.0097/0.0484	0.0781	0.0775(16)(3)	0.1218	0.12237(26)(20)
≈ 0.12	0.02/0.05	0.0918	0.0868(9)(3)	0.1259	0.12423(15)(16)
	0.01/0.05	0.0901	0.0868(9)(3)	0.1254	0.12423(15)(16)
	0.007/0.05	0.0901	0.0868(9)(3)	0.1254	0.12423(15)(16)
	0.005/0.05	0.0901	0.0879(9)(3)	0.1254	0.12452(15)(16)
≈ 0.09	0.0124/0.031	0.0982	0.0964(7)(3)	0.1277	0.12710(9)(14)
	0.0062/0.031	0.0979	0.0965(7)(3)	0.1276	0.12714(9)(14)
	0.00465/0.031	0.0977	0.0966(7)(3)	0.1275	0.12718(9)(14)
	0.0031/0.031	0.0976	0.0967(7)(3)	0.1275	0.12722(9)(14)
	0.00155/0.031	0.0976	0.0972(7)(3)	0.1275	0.12737(9)(14)
≈ 0.06	0.0072/0.018	0.1048	0.1050(5)(2)	0.1295	0.12955(4)(11)
	0.0036/0.018	0.1052	0.1051(5)(2)	0.1296	0.12957(4)(11)
	0.0025/0.018	0.1052	0.1052(5)(2)	0.1296	0.12960(4)(11)
	0.0018/0.018	0.1052	0.1054(5)(2)	0.1296	0.12964(4)(11)
≈ 0.045	0.0028/0.014	0.1143	0.1116(3)(2)	0.1310	0.130921(16)(7)

Table 18 Ensemble and kappa parameters used to calculate the mistuning correction [22]. The values in bold are the uncorrected simulation values for this ensemble. Not all combinations are available, since we change only one quark mass at a time while keeping the other quark mass fixed to its uncorrected value

a (fm)	am_l/am_s	Available κ_b'	Available κ_c'
≈ 0.12	0.01/0.05	**0.0901**, 0.860, 0.820	**0.1254**, 0.1280, 0.1230

lattice scale, set by $r_1 = 0.3117(22)$ fm [64], we tune the heavy-quark masses so that the kinetic masses M_2 of the B_s and D_s mesons agree with their experimental values [64,70]. Historically, the tuning was done in two stages. A preliminary, lower-statistics study set the heavy-quark masses used in the simulation to generate the two- and three-point correlators in the present study. A subsequent, higher-statistics study was then possible, and it gave slightly different kappa values, as shown in Table 17. We denote the simulation values with κ_c' and κ_b' and the refined values with κ_c and κ_b. As in Ref. [22], we proceed to correct our results for the slight mistuning and, in the process, estimate the systematic error associated with the correction.

The correction procedure is based on a single $a \approx 0.12$ fm ensemble. A full set of two- and three-point correlation functions were calculated at slightly shifted values of κ_b' and κ_c', as shown in Table 18. The shifted values were chosen close to both the corrected and simulation kappa values. With these data, we then estimate the derivatives of the form factors and the recoil parameter with respect to the heavy quark masses. By expressing these results in dimensionless terms, we can extrapolate them to other ensembles with different light-quark masses and lattice spacings. As a departure from Ref. [53], we perform the correction after the renormalization factors have been applied.

The first step is to construct the combinations of ratios we need to correct. Clever combinations can remove a dependence on the recoil parameter or vanish at zero recoil. We can use this information to improve the precision of the corrections. In this case, we define the following quantities:

$$A = (1 - x_f^2) R_{A_1}, \tag{C.52}$$

$$V = \frac{X_v}{x_f}, \tag{C.53}$$

$$B_0 = \frac{X_0}{x_f}, \tag{C.54}$$

$$B_1 = \frac{X_1 - 1}{2x_f^2}, \tag{C.55}$$

$$C_1 = \frac{X_1 + 1}{2}, \tag{C.56}$$

At zero recoil we expect $C_1 \to 1$, since $X_1 \to 1$ in that limit. Although x_f vanishes when $w \to 1$, the combinations are designed to give a finite value at zero recoil. In fact, it is easy to reconstruct the form factors using these building blocks,

$$h_{A_1} = A, \tag{C.57}$$

$$h_V = AV, \tag{C.58}$$

$$h_{A_2} = A(C_1 + B_1 - B_0), \tag{C.59}$$

$$h_{A_3} = A(C_1 - B_1). \tag{C.60}$$

Table 19 Results of the κ correction fits

Heavy quark	a_w	a_A	b_A	b_V	b_{B_0}	b_{B_1}	a_{C_1}
Bottom	–	1.6(1.5)	− 0.23(13)	0.37(14)	− 0.65(15)	− 0.35(37)	− 0.17(17)
Charm	1.14(11)	− 0.79(55)	− 0.010(31)	0.186(84)	0.297(78)	− 0.47(30)	− 0.38(14)

Table 20 Effect of the mistuning adjustments in the form factors and the recoil parameter of the $a \approx 0.12$ fm ensemble used to compute the correction. Only the final error is shown in the table. The error for w seems to decrease with the correction, but the relevant quantity is $w-1$, for which the error either stays the same or increases. The correction (and its error) is highly correlated with $w-1$, as it is directly proportional to it

Momentum	(0, 0, 0)		(1, 0, 0)		(2, 0, 0)	
	Raw	Tuned	Raw	Tuned	Raw	Tuned
w	1.0(0)	1.0(0)	1.0402(40)	1.0374(37)	1.134(21)	1.125(20)
h_{A_1}	0.901(13)	0.906(13)	0.858(21)	0.863(21)	0.816(30)	0.824(30)
h_{A_2}	–	–	− 0.43(15)	− 0.40(16)	− 0.49(17)	− 0.46(17)
h_{A_3}	–	–	0.98(12)	0.96(12)	1.02(12)	1.01(12)
h_V	–	–	1.106(85)	1.098(85)	1.05(11)	1.04(11)

The last quantity corrected is the recoil parameter, defined as in Eq. 21, with a trivial dependence on κ_b.

We estimate the derivative by taking finite differences between the observables in Eqs. (C.52)–(C.56) calculated in the standard run and one of the correction runs listed in Table 18. The derivative is taken with respect to $\xi_\alpha = 1/am_{2\alpha}$ with $\alpha = b, c$, where the kinetic mass $am_{2\alpha}$ is defined as

$$\frac{1}{am_{2\alpha}} = \frac{2}{am_{0\alpha}(2+am_{0\alpha})} + \frac{1}{1+am_{0\alpha}}, \tag{C.61}$$

and the bare quark mass is calculated using the tadpole-improved, tree-level formula

$$am_{0\alpha} = \frac{1}{u_0}\left(\frac{1}{2\kappa_\alpha} - \frac{1}{2\kappa_{cr}}\right). \tag{C.62}$$

Here u_0 is the tadpole parameter, and κ_{cr} is the value of κ at which the clover-quark pion becomes massless. We calculate the derivatives of each observable for each available value of the recoil parameter. When computing the correction for the b quark, since the $\kappa_b' = 0.860$ simulation value is so close to the tuned value of κ_b, we do not use the data coming from $\kappa_b' = 0.820$.

In general, the calculated derivative is small for all recoil values, and the errors are large enough that we can consider a linear dependence on $w-1$ for all derivatives. Nonetheless, we can exploit some good properties of our observables in order to reduce the number of free parameters. For instance, $dC_1/d\xi_\alpha(w=1) = 0$, since $C_1 = 1$ in that limit, so we can set the constant term to zero. Also, the derivatives of V, $B_{0,1}$ change very little with $w-1$ compared with their errors. Thus, we can safely assume these derivatives behave as constants in the small recoil range we are considering. In fact, these quantities are derived from ratios of form factors h_X/h_{A_1}, which should depend only weakly on w. These simplifications are not only welcome, they are necessary since for V, $B_{0,1}$, C_1 we have only two values of the recoil parameter available, so our fits can have one degree of freedom. The observable A is not a ratio *per se*, but it is calculated at zero recoil as well, and a clear dependence with w arises. Finally, it is obvious that $dw/d\xi_c(w=1) = 0$, and we can apply the same treatment as for C_1. Summarizing, we fit our data to the expressions,

$$\frac{dw}{d\xi_c} = a_{w,c}(w-1), \tag{C.63}$$
$$\frac{dA}{d\xi_c} = a_{A,c}(w-1) + b_{A,c}, \tag{C.64}$$
$$\frac{dV}{d\xi_c} = b_{V,c}, \tag{C.65}$$
$$\frac{dB_0}{d\xi_c} = b_{B_0,c}, \tag{C.66}$$
$$\frac{dB_1}{d\xi_c} = b_{B_1,c}, \tag{C.67}$$
$$\frac{dC_1}{d\xi_c} = a_{C_1,c}(w-1), \tag{C.68}$$
$$\frac{dw}{d\xi_b} = 0, \tag{C.69}$$
$$\frac{dA}{d\xi_b} = a_{A,b}(w-1) + b_{A,b}, \tag{C.70}$$
$$\frac{dV}{d\xi_b} = b_{V,b}, \tag{C.71}$$
$$\frac{dB_0}{d\xi_b} = b_{B_0,b}, \tag{C.72}$$

$$\frac{dB_1}{d\xi_b} = b_{B_1,b}, \tag{C.73}$$

$$\frac{dC_1}{d\xi_b} = a_{C_1,b}(w-1). \tag{C.74}$$

All fits are unconstrained and result in a p value exceeding 0.5. Our final corrections are gathered in Table 19. These values are used to correct the form factors for all ensembles. As an example, we show in Table 20 how this tuning modifies the calculated form factors for the ensemble used in the correction.

Appendix D: Full covariance matrix of the chiral-continuum extrapolation results

Future work with the lattice-QCD form factors computed here can start with the synthetic data generated for the z expansion. We provide these data for the form factors in the BGL and HQET notation, along with their correlation matrix, in the ancillary files included in this paper. We also include the values and correlation matrix of the BGL z expansion coefficients resulting from our preferred fits from Sects. 5.1.1 and 5.2, using only lattice-QCD data (the quadratic fit in Table 11) and lattice QCD plus experimental data (the last column in Table 12). Our BGL lattice-QCD-only fit can be reproduced by fitting the included synthetic data in the BGL notation to the BGL expansion described in Sect. 5.1, and using the constraint at zero recoil given in Eq. (72) to remove c_0.

Appendix D.1: Reading the data

The data is provided in Python format using the `gvar` package [137]. The synthetic data points can be read from the file `SynthData.PyDat` using the following code:

```
import gvar

data = gvar.load("SynthData.PyDat")
corr = gvar.evalcorr(data)
```

After execution, `data` holds a dictionary whose keys are in the format `F(w)`, where `F = g, f, F1 or F2` for the g, f, $\mathcal{F}_1$ and $\mathcal{F}_2$ form factors in the BGL notation, or `F = hA1, hA2, hA3 or V` for the h_{A_i} and h_V form factors in the HQET notation, and `w` is the recoil parameter, $w = 1.03, 1.10, 1.17$. The h_X form factors are also available at zero recoil, but the extra data point at $w = 1.00$ is not independent and should be used only if one of the other available points is removed. The last line stores the correlations between the different synthetic data points in a dictionary called `corr`. One can access any correlation by invoking the pair `corr[F(w1),G(w2)]`, where `F` and `G` are the two relevant form factors, and `w1` and `w2` are the recoil parameters at which the correlation should be evaluated.

The z-fit results are stored in the file `FitResults.PyDat` and can be read following the same procedure. The dictionary keys in this case are in the format `FIT_xj`, where `FIT` can be either `LQCD` for the lattice-QCD-only fit, or `LQCDEXP` for the fit including lattice-QCD and experimental data, and `xj = a0, a1...` are the different coefficients of the z expansion $a_0, a_1 \ldots$ Our final result for $|\eta_{\rm EW}|^2|V_{cb}|^2$ is stored under the key `LQCDEXP_eVcb`, and the results for $R(D^*)$ are stored as `LQCD_RDst` and `LQCDEXP_RDst`.

References

1. Y.S. Amhis et al., Eur. Phys. J. C **81**(3), 226 (2021). https://doi.org/10.1140/epjc/s10052-020-8156-7
2. P.A. Zyla et al., PTEP **2020**(8), 083C01 (2020). https://doi.org/10.1093/ptep/ptaa104
3. A. Crivellin, S. Pokorski, Phys. Rev. Lett. **114**(1), 011802 (2015). https://doi.org/10.1103/PhysRevLett.114.011802
4. M. Jung, D.M. Straub, JHEP **01**, 009 (2019). https://doi.org/10.1007/JHEP01(2019)009
5. C.G. Boyd, B. Grinstein, R.F. Lebed, Nucl. Phys. B **461**, 493 (1996). https://doi.org/10.1016/0550-3213(95)00653-2
6. C.G. Boyd, B. Grinstein, R.F. Lebed, Phys. Lett. B **353**, 306 (1995). https://doi.org/10.1016/0370-2693(95)00480-9
7. C.G. Boyd, B. Grinstein, R.F. Lebed, Phys. Rev. D **56**, 6895 (1997). https://doi.org/10.1103/PhysRevD.56.6895
8. I. Caprini, L. Lellouch, M. Neubert, Nucl. Phys. B **530**, 153 (1998). https://doi.org/10.1016/S0550-3213(98)00350-2
9. B. Aubert et al., Phys. Rev. Lett. **100**, 231803 (2008). https://doi.org/10.1103/PhysRevLett.100.231803
10. N.E. Adam et al., Phys. Rev. D **67**, 032001 (2003). https://doi.org/10.1103/PhysRevD.67.032001
11. B. Aubert et al., Phys. Rev. D **79**, 012002 (2009). https://doi.org/10.1103/PhysRevD.79.012002
12. Y. Amhis, et al. (2012). arXiv:1207.1158
13. A. Abdesselam, et al. (2017). arXiv:1702.01521
14. D. Bigi, P. Gambino, S. Schacht, Phys. Lett. B **769**, 441 (2017). https://doi.org/10.1016/j.physletb.2017.04.022
15. D. Bigi, P. Gambino, S. Schacht, JHEP **11**, 061 (2017). https://doi.org/10.1007/JHEP11(2017)061
16. B. Grinstein, A. Kobach, Phys. Lett. B **771**, 359 (2017). https://doi.org/10.1016/j.physletb.2017.05.078
17. S. Jaiswal, S. Nandi, S.K. Patra, JHEP **12**, 060 (2017). https://doi.org/10.1007/JHEP12(2017)060
18. E. Waheed et al., Phys. Rev. D **100**(5), 052007 (2019). https://doi.org/10.1103/PhysRevD.100.052007
19. J. Lees et al., Phys. Rev. Lett. **123**(9), 091801 (2019). https://doi.org/10.1103/PhysRevLett.123.091801
20. P. Gambino, M. Jung, S. Schacht, Phys. Lett. B **795**, 386 (2019). https://doi.org/10.1016/j.physletb.2019.06.039
21. S. Jaiswal, S. Nandi, S.K. Patra, JHEP **06**, 165 (2020). https://doi.org/10.1007/JHEP06(2020)165
22. J.A. Bailey et al., Phys. Rev. D **89**(11), 114504 (2014). https://doi.org/10.1103/PhysRevD.89.114504
23. E. McLean, C.T.H. Davies, A.T. Lytle, J. Koponen, Phys. Rev. D **99**(11), 114512 (2019). https://doi.org/10.1103/PhysRevD.99.114512

24. S. Bifani, S. Descotes-Genon, A. Romero Vidal, M.H. Schune, J. Phys. G **46**(2), 023001 (2019). https://doi.org/10.1088/1361-6471/aaf5de
25. S. Hirose et al., Phys. Rev. Lett. **118**(21), 211801 (2017). https://doi.org/10.1103/PhysRevLett.118.211801
26. S. Hirose et al., Phys. Rev. D **97**(1), 012004 (2018). https://doi.org/10.1103/PhysRevD.97.012004
27. R. Aaij et al., Phys. Rev. Lett. **120**(17), 171802 (2018). https://doi.org/10.1103/PhysRevLett.120.171802
28. R. Aaij et al., Phys. Rev. D **97**(7), 072013 (2018). https://doi.org/10.1103/PhysRevD.97.072013
29. G. Caria et al., Phys. Rev. Lett. **124**(16), 161803 (2020). https://doi.org/10.1103/PhysRevLett.124.161803
30. S. Hashimoto, A.S. Kronfeld, P.B. Mackenzie, S.M. Ryan, J.N. Simone, Phys. Rev. D **66**, 014503 (2002). https://doi.org/10.1103/PhysRevD.66.014503
31. C. Bernard et al., Phys. Rev. D **79**, 014506 (2009). https://doi.org/10.1103/PhysRevD.79.014506
32. J. Harrison, C. Davies, M. Wingate, Phys. Rev. D **97**(5), 054502 (2018). https://doi.org/10.1103/PhysRevD.97.054502
33. G.M. de Divitiis, R. Petronzio, N. Tantalo, Nucl. Phys. B **807**, 373 (2009). https://doi.org/10.1016/j.nuclphysb.2008.09.013
34. S.W. Qiu, C. DeTar, A.X. El-Khadra, A.S. Kronfeld, J. Laiho, R.S. Van de Water, PoS **LATTICE2013**, 385 (2014). https://doi.org/10.22323/1.187.0385
35. A. VaqueroAvilés-Casco, C. DeTar, D. Du, A. El-Khadra, A.S. Kronfeld, J. Laiho, R.S. Van deWater, EPJ Web Conf **175**, 13003 (2018). https://doi.org/10.1051/epjconf/201817513003
36. A.V. Avilés-Casco, C. DeTar, A.X. El-Khadra, A.S. Kronfeld, J. Laiho, R.S. Van de Water, PoS **LATTICE2018**, 282 (2019). https://doi.org/10.22323/1.334.0282
37. A. Vaquero, C. DeTar, A.X. El-Khadra, A.S. Kronfeld, J. Laiho, R.S. Van de Water, in *17th Conference on Flavor Physics and CP Violation* (2019). arXiv:1906.01019
38. A.V. Avilés-Casco, C. DeTar, A.X. El-Khadra, A.S. Kronfeld, J. Laiho, R.S. Van de Water, PoS **LATTICE2019**, 049 (2019). https://doi.org/10.22323/1.363.0049
39. J.A. Bailey et al., Phys. Rev. D **92**(1), 014024 (2015). https://doi.org/10.1103/PhysRevD.92.014024
40. J.A. Bailey et al., Phys. Rev. Lett. **115**(15), 152002 (2015). https://doi.org/10.1103/PhysRevLett.115.152002
41. J.A. Bailey et al., Phys. Rev. D **93**(2), 025026 (2016). https://doi.org/10.1103/PhysRevD.93.025026
42. A. Bazavov et al., Phys. Rev. D **100**(3), 034501 (2019). https://doi.org/10.1103/PhysRevD.100.034501
43. A.F. Falk, M. Neubert, Phys. Rev. D **47**, 2965 (1993). https://doi.org/10.1103/PhysRevD.47.2965
44. J.G. Körner, G.A. Schuler, Z. Phys. C **46**, 93 (1990). https://doi.org/10.1007/BF02440838
45. A. Sirlin, Nucl. Phys. B **196**, 83 (1982). https://doi.org/10.1016/0550-3213(82)90303-0
46. E.S. Ginsberg, Phys. Rev. **171**, 1675 (1968). https://doi.org/10.1103/PhysRev.171.1675. https://doi.org/10.1103/PhysRev.174.2169.3 [Erratum: Phys. Rev. 174, 2169 (1968)]
47. E.S. Ginsberg, Phys. Rev. **162**, 1570 (1967). https://doi.org/10.1103/physrev.187.2280.2. https://doi.org/10.1103/PhysRev.162.1570 [Erratum: Phys. Rev. 187, 2280 (1969)]
48. D. Atwood, W.J. Marciano, Phys. Rev. D **41**, 1736 (1990). https://doi.org/10.1103/PhysRevD.41.1736
49. A.X. El-Khadra, A.S. Kronfeld, P.B. Mackenzie, Phys. Rev. D **55**, 3933 (1997). https://doi.org/10.1103/PhysRevD.55.3933
50. J. Harada, S. Hashimoto, A.S. Kronfeld, T. Onogi, Phys. Rev. D **65**, 094514 (2002). https://doi.org/10.1103/PhysRevD.65.094514
51. A.X. El-Khadra, A.S. Kronfeld, P.B. Mackenzie, S.M. Ryan, J.N. Simone, Phys. Rev. D **64**, 014502 (2001). https://doi.org/10.1103/PhysRevD.64.014502
52. J.A. Bailey et al., Phys. Rev. D **85**, 114502 (2012). https://doi.org/10.1103/PhysRevD.85.114502 [Erratum: Phys. Rev. D 86, 039904 (2012)]
53. J.A. Bailey et al., Phys. Rev. D **92**(3), 034506 (2015). https://doi.org/10.1103/PhysRevD.92.034506
54. A. Bazavov et al., Rev. Mod. Phys. **82**, 1349 (2010). https://doi.org/10.1103/RevModPhys.82.1349
55. C. Aubin, C. Bernard, C. DeTar, J. Osborn, S. Gottlieb, E.B. Gregory, D. Toussaint, U.M. Heller, J.E. Hetrick, R. Sugar, Phys. Rev. D **70**, 094505 (2004). https://doi.org/10.1103/PhysRevD.70.094505
56. C.W. Bernard, T. Burch, K. Orginos, D. Toussaint, T.A. DeGrand, C.E. Detar, S. Datta, S.A. Gottlieb, U.M. Heller, R. Sugar, Phys. Rev. D **64**, 054506 (2001). https://doi.org/10.1103/PhysRevD.64.054506
57. M. Wingate, J. Shigemitsu, C.T.H. Davies, G.P. Lepage, H.D. Trottier, Phys. Rev. D **67**, 054505 (2003). https://doi.org/10.1103/PhysRevD.67.054505
58. A. Bazavov et al., Phys. Rev. D **93**(11), 113016 (2016). https://doi.org/10.1103/PhysRevD.93.113016
59. A. Bazavov et al., Phys. Rev. D **97**(3), 034513 (2018). https://doi.org/10.1103/PhysRevD.97.034513
60. A. Bazavov et al., Phys. Rev. D **81**, 114501 (2010). https://doi.org/10.1103/PhysRevD.81.114501
61. R. Li et al., PoS **LATTICE2018**, 269 (2019). https://doi.org/10.22323/1.334.0269
62. A. Bazavov et al., Phys. Rev. D **98**(7), 074512 (2018). https://doi.org/10.1103/PhysRevD.98.074512
63. J.L. Richardson, Phys. Lett. B **82**, 272 (1979). https://doi.org/10.1016/0370-2693(79)90753-6
64. A. Bazavov et al., Phys. Rev. D **85**, 114506 (2012). https://doi.org/10.1103/PhysRevD.85.114506
65. G.P. Lepage, B. Clark, C.T.H. Davies, K. Hornbostel, P.B. Mackenzie, C. Morningstar, H. Trottier, Nucl. Phys. B Proc. Suppl. **106**, 12 (2002). https://doi.org/10.1016/S0920-5632(01)01638-3
66. D. Toussaint, W. Freeman, (2008). arXiv:0808.2211
67. M.B. Oktay, A.S. Kronfeld, Phys. Rev. D **78**, 014504 (2008). https://doi.org/10.1103/PhysRevD.78.014504
68. B.P.G. Mertens, A.S. Kronfeld, A.X. El-Khadra, Phys. Rev. D **58**, 034505 (1998). https://doi.org/10.1103/PhysRevD.58.034505
69. A.S. Kronfeld, Nucl. Phys. B Proc. Suppl. **53**, 401 (1997). https://doi.org/10.1016/S0920-5632(96)00671-8
70. C. Bernard et al., Phys. Rev. D **83**, 034503 (2011). https://doi.org/10.1103/PhysRevD.83.034503
71. S.J. Brodsky, G.P. Lepage, P.B. Mackenzie, Phys. Rev. D **28**, 228 (1983). https://doi.org/10.1103/PhysRevD.28.228
72. G.P. Lepage, P.B. Mackenzie, Phys. Rev. D **48**, 2250 (1993). https://doi.org/10.1103/PhysRevD.48.2250
73. J.A. Bailey et al., Phys. Rev. D **79**, 054507 (2009). https://doi.org/10.1103/PhysRevD.79.054507
74. J. Harada, S. Hashimoto, K.I. Ishikawa, A.S. Kronfeld, T. Onogi, N. Yamada, Phys. Rev. D **65**, 094513 (2002). https://doi.org/10.1103/PhysRevD.71.019903. https://doi.org/10.1103/PhysRevD.65.094513 [Erratum: Phys. Rev. D 71, 019903 (2005)]
75. C. Aubin, C. Bernard, Phys. Rev. D **73**, 014515 (2006). https://doi.org/10.1103/PhysRevD.73.014515
76. S. Prelovsek, Phys. Rev. D **73**, 014506 (2006). https://doi.org/10.1103/PhysRevD.73.014506
77. C.W. Bernard, C.E. DeTar, Z. Fu, S. Prelovsek, PoS **LAT2006**, 173 (2006). https://doi.org/10.22323/1.032.0173
78. C. Bernard, Phys. Rev. D **73**, 114503 (2006). https://doi.org/10.1103/PhysRevD.73.114503

79. C. Bernard, C.E. DeTar, Z. Fu, S. Prelovsek, Phys. Rev. D **76**, 094504 (2007). https://doi.org/10.1103/PhysRevD.76.094504
80. C. Aubin, J. Laiho, R.S. Van de Water, Phys. Rev. D **77**, 114501 (2008). https://doi.org/10.1103/PhysRevD.77.114501
81. C. Bernard, M. Golterman, Y. Shamir, Phys. Rev. D **73**, 114511 (2006). https://doi.org/10.1103/PhysRevD.73.114511
82. Y. Shamir, Phys. Rev. D **75**, 054503 (2007). https://doi.org/10.1103/PhysRevD.75.054503
83. Y. Shamir, Phys. Rev. D **71**, 034509 (2005). https://doi.org/10.1103/PhysRevD.71.034509
84. S. Dürr, PoS LAT2005, **021** (2006). https://doi.org/10.22323/1.020.0021
85. S.R. Sharpe, PoS **LAT2006**, 022 (2006). https://doi.org/10.22323/1.032.0022
86. A.S. Kronfeld, PoS **016**, LATTICE2007 (2007). https://doi.org/10.22323/1.042.0016
87. J. Laiho, R.S. Van de Water, Phys. Rev. D **73**, 054501 (2006). https://doi.org/10.1103/PhysRevD.73.054501
88. A. Anastassov et al., Phys. Rev. D **65**, 032003 (2002). https://doi.org/10.1103/PhysRevD.65.032003
89. J.P. Lees et al., Phys. Rev. D **88**(5), 052003 (2013). https://doi.org/10.1103/PhysRevD.88.079902. https://doi.org/10.1103/PhysRevD.88.052003 [Erratum: Phys. Rev. D 88, no. 7, 079902 (2013)]
90. J.P. Lees et al., Phys. Rev. Lett. **111**(11), 111801 (2013). https://doi.org/10.1103/PhysRevLett.111.111801, https://doi.org/10.1103/PhysRevLett.111.169902
91. W. Detmold, C.J.D. Lin, S. Meinel, Phys. Rev. Lett. **108**, 172003 (2012). https://doi.org/10.1103/PhysRevLett.108.172003
92. K.U. Can, G. Erkol, M. Oka, A. Ozpineci, T.T. Takahashi, Phys. Lett. B **719**, 103 (2013). https://doi.org/10.1016/j.physletb.2012.12.050
93. D. Becirevic, F. Sanfilippo, Phys. Lett. B **721**, 94 (2013). https://doi.org/10.1016/j.physletb.2013.03.004
94. J.M. Flynn, P. Fritzsch, T. Kawanai, C. Lehner, B. Samways, C.T. Sachrajda, R.S. Van de Water, O. Witzel, Phys. Rev. D **93**(1), 014510 (2016). https://doi.org/10.1103/PhysRevD.93.014510
95. F. Bernardoni, J. Bulava, M. Donnellan, R. Sommer, Phys. Lett. B **740**, 278 (2015). https://doi.org/10.1016/j.physletb.2014.11.051
96. W. Detmold, C.J.D. Lin, S. Meinel, Phys. Rev. D **85**, 114508 (2012). https://doi.org/10.1103/PhysRevD.85.114508
97. S. Aoki et al., Eur. Phys. J. C **80**(2), 113 (2020). https://doi.org/10.1140/epjc/s10052-019-7354-7
98. M.E. Luke, Phys. Lett. B **252**, 447 (1990). https://doi.org/10.1016/0370-2693(90)90568-Q
99. O. Ledoit, M. Wolf, J. Empir. Finance **10**(5), 603 (2003). https://doi.org/10.1016/S0927-5398(03)00007-0
100. J. Schaefer, K. Strimmer, Statistical Applications in Genetics and Molecular Biology **4**, 32 (2005)
101. O. Ledoit, M. Wolf, (264) (2017). https://doi.org/10.5167/uzh-139880. http://hdl.handle.net/10419/173422. Accessed 27 May 2021
102. J.N. Simone. Improved data covariance estimation techniques in lattice QCD (2017). https://cafpe.ugr.es/lattice2017/indico/session/102/contribution/386.html. talk at https://cafpe.ugr.es/lattice2017/. 35th International Symposium on Lattice Field Theory
103. A.S. Kronfeld, Phys. Rev. D **62**, 014505 (2000). https://doi.org/10.1103/PhysRevD.62.014505
104. a Bazavov et al., PoS **LATTICE2010**, 074 (2010). https://doi.org/10.22323/1.105.0074
105. R. Sommer, Nucl. Phys. B **411**, 839 (1994). https://doi.org/10.1016/0550-3213(94)90473-1
106. C.W. Bernard, T. Burch, K. Orginos, D. Toussaint, T.A. DeGrand, C.E. DeTar, S.A. Gottlieb, U.M. Heller, J.E. Hetrick, B. Sugar, Phys. Rev. D **62**, 034503 (2000). https://doi.org/10.1103/PhysRevD.62.034503
107. D. Arndt, C.J.D. Lin, Phys. Rev. D **70**, 014503 (2004). https://doi.org/10.1103/PhysRevD.70.014503
108. P. Gambino et al., Eur. Phys. J. C **80**(10), 966 (2020). https://doi.org/10.1140/epjc/s10052-020-08490-x
109. C. Bourrely, I. Caprini, L. Lellouch, Phys. Rev. D **79**, 013008 (2009). https://doi.org/10.1103/PhysRevD.82.099902 [Erratum: Phys. Rev. D 82, 099902 (2010)]
110. D. Ferlewicz, P. Urquijo, E. Waheed, Phys. Rev. D **103**(7), 073005 (2021). https://doi.org/10.1103/PhysRevD.103.073005
111. D. Bigi, P. Gambino, Phys. Rev. D **94**(9), 094008 (2016). https://doi.org/10.1103/PhysRevD.94.094008
112. W. Dungel et al., Phys. Rev. D **82**, 112007 (2010). https://doi.org/10.1103/PhysRevD.82.112007
113. S. de Boer, T. Kitahara, I. Nisandzic, Phys. Rev. Lett. **120**(26), 261804 (2018). https://doi.org/10.1103/PhysRevLett.120.261804
114. S. Calí, S. Klaver, M. Rotondo, B. Sciascia, Eur. Phys. J. C **79**(9), 744 (2019). https://doi.org/10.1140/epjc/s10052-019-7254-x
115. Y. Amhis et al., Eur. Phys. J. C **77**(12), 895 (2017). https://doi.org/10.1140/epjc/s10052-017-5058-4
116. C. Bobeth, M. Bordone, N. Gubernari, M. Jung, D. van Dyk, Eur. Phys. J. C **81**(11), 984 (2021). https://doi.org/10.1140/epjc/s10052-021-09724-2
117. J.P. Lees et al., Phys. Rev. Lett. **116**(4), 041801 (2016). https://doi.org/10.1103/PhysRevLett.116.041801
118. S. Fajfer, J.F. Kamenik, I. Nisandzic, Phys. Rev. D **85**, 094025 (2012). https://doi.org/10.1103/PhysRevD.85.094025
119. F.U. Bernlochner, Z. Ligeti, M. Papucci, D.J. Robinson, Phys. Rev. D **95**(11), 115008 (2017). https://doi.org/10.1103/PhysRevD.95.115008 [Erratum: Phys. Rev. D 97, 059902 (2018)]
120. M. Bordone, M. Jung, D. van Dyk, Eur. Phys. J. C **80**(2), 74 (2020). https://doi.org/10.1140/epjc/s10052-020-7616-4
121. M. Bordone, N. Gubernari, D. van Dyk, M. Jung, Eur. Phys. J. C **80**(4), 347 (2020). https://doi.org/10.1140/epjc/s10052-020-7850-9
122. N. Gubernari, A. Kokulu, D. van Dyk, JHEP **01**, 150 (2019). https://doi.org/10.1007/JHEP01(2019)150
123. W. Altmannshofer, et al., PTEP **2019**(12), 123C01 (2019). https://doi.org/10.1093/ptep/ptz106 [Erratum: PTEP 2020, 029201 (2020)]
124. H. Na, C.M. Bouchard, G.P. Lepage, C. Monahan, J. Shigemitsu, Phys. Rev. D **92**(5), 054510 (2015). https://doi.org/10.1103/PhysRevD.93.119906. [Erratum: Phys. Rev. D 93, 119906 (2016)]
125. R. Aaij et al., Phys. Rev. D **101**(7), 072004 (2020). https://doi.org/10.1103/PhysRevD.101.072004
126. E. McLean, C.T.H. Davies, J. Koponen, A.T. Lytle, Phys. Rev. D **101**(7), 074513 (2020). https://doi.org/10.1103/PhysRevD.101.074513
127. M. Bordone, B. Capdevila, P. Gambino, Phys. Lett. B **822**, 136679 (2021). https://doi.org/10.1016/j.physletb.2021.136679
128. M. Fael, T. Mannel, K.K. Vos, JHEP **02**, 177 (2019). https://doi.org/10.1007/JHEP02(2019)177
129. F. Bernlochner, M. Fael, K. Olschewsky, E. Persson, R. van Tonder, K.K. Vos, M. Welsch, JHEP **10**, 068 (2022). https://doi.org/10.1007/JHEP10(2022)068
130. S. Hashimoto, PTEP **2017**(5), 053B03 (2017). https://doi.org/10.1093/ptep/ptx052
131. P. Gambino, S. Hashimoto, Phys. Rev. Lett. **125**(3), 032001 (2020). https://doi.org/10.1103/PhysRevLett.125.032001
132. J. Lees et al., Phys. Rev. Lett. **109**, 101802 (2012). https://doi.org/10.1103/PhysRevLett.109.101802
133. T. Kaneko, Y. Aoki, G. Bailas, B. Colquhoun, H. Fukaya, S. Hashimoto, J. Koponen, PoS **LATTICE2019**, 139 (2019). https://doi.org/10.22323/1.363.0139

134. C.K. Chow, M.B. Wise, Phys. Rev. D **48**, 5202 (1993). https://doi.org/10.1103/PhysRevD.48.5202
135. C. Aubin, C. Bernard, Phys. Rev. D **68**, 034014 (2003). https://doi.org/10.1103/PhysRevD.68.034014
136. J.A. Bailey et al., Phys. Rev. Lett. **109**, 071802 (2012). https://doi.org/10.1103/PhysRevLett.109.071802
137. P. Lepage, C. Gohlke, D. Hackett. gplepage/gvar: gvar version 11.9.2 (2021). https://doi.org/10.5281/zenodo.4695132