Lattice QCD study of inclusive semileptonic decays of heavy mesons

Paolo Gambino,[a] Shoji Hashimoto,[b,c] Sandro Mächler,[a,d] Marco Panero,[a] Francesco Sanfilippo,[e] Silvano Simula,[e] Antonio Smecca[a] and Nazario Tantalo[f]

[a]*Dipartimento di Fisica, Università di Torino & INFN, Sezione di Torino,*
Via Pietro Giuria 1, I-10125 Turin, Italy

[b]*Theory Center, Institute of Particle and Nuclear Studies,*
High Energy Accelerator Research Organization (KEK),
Tsukuba 305-0801, Japan

[c]*School of High Energy Accelerator Science,*
The Graduate University for Advanced Studies (SOKENDAI),
Tsukuba 305-0801, Japan

[d]*Physikinstitut, Universität Zürich,*
Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

[e]*INFN, Sezione di Roma Tre,*
Via della Vasca Navale 84, I-00146 Rome, Italy

[f]*Dipartimento di Fisica, Università di Roma "Tor Vergata"*
& INFN, Sezione di Roma "Tor Vergata",
Via della Ricerca Scientifica 1, I-00133 Rome, Italy

E-mail: paolo.gambino@unito.it, shoji.hashimoto@kek.jp,
sandro.machler@unito.it, marco.panero@unito.it,
francesco.sanfilippo@infn.it, silvano.simula@roma3.infn.it,
antonio.smecca@unito.it, nazario.tantalo@roma2.infn.it

Abstract: We present an *ab initio* study of inclusive semileptonic decays of heavy mesons from lattice QCD. Our approach is based on a recently proposed method, that allows one to address the study of these decays from the analysis of smeared spectral functions extracted from four-point correlators on the lattice, where the smearing is defined in terms of the phase-space integration relevant to the inclusive decays. We present results obtained from gauge-field ensembles from the JLQCD and ETM collaborations, and discuss their relation with theoretical predictions from the operator-product expansion.

Contents

1 Introduction

The theoretical study of semileptonic decays of B mesons continues to be an important and very active area of research in high-energy physics: this interest is mainly driven by the fact that these decays encode direct information on the modulus of two of the elements of the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix [1, 2], namely $|V_{ub}|$ and $|V_{cb}|$, and may be a sensitive probe to new physics beyond the Standard Model (SM). As a matter of fact, many different types of extensions of the SM are expected to affect flavour physics, inducing new flavour-changing interactions, complex phases in the CKM matrix, possible violations of lepton-flavour universality, etc. Even if the mass scales of new particles beyond the SM turned out to be very high, quantum effects of the associated fields could leave detectable imprints onto the physics of bottom and charm quarks.

On the experimental side, recent results from B factories reveal some tension with SM predictions, but also exhibit puzzling discrepancies between exclusive and inclusive channels [3–6]. For theorists, this provides further motivation to improve the understanding of these decays and to refine their predictions. Currently, the most powerful tool to obtain theoretical predictions from the first principles of QCD is the one based on numerical simulations in the lattice regularisation of the theory [7]. It is an intrinsically non-perturbative approach, that allows one to obtain accurate and systematically improvable predictions for a variety of quantities, including those relevant for decays of heavy mesons: for an up-to-date world review of lattice results relevant to flavour physics, see ref. [8]. It should be emphasized, however, that most lattice calculations focus on *exclusive* decays: in a nutshell, this is due to the fact that inclusive processes consist of a potentially very large number of physical states — including states featuring multiple hadrons, which pose their own challenges — and their systematic analysis in numerical calculations is very impractical, if possible at all.

Recently, however, novel approaches have been put forward, that allow one to address inclusive decays in lattice QCD. As an example, in ref. [9] it was pointed out that the differential rate for inclusive decays of the type $B \to X\ell\nu$ (where X denotes all hadronic states that are compatible with the semileptonic decay of the bottom quark) could be evaluated by relating the hadronic tensor

$$W_{\mu\nu}(p,q) = \frac{4\pi^3}{E_B} \sum_X \delta^4(p - q - p_X)\langle B(p)|J_\mu^\dagger|X(p_X)\rangle\langle X(p_X)|J_\nu|B(p)\rangle \qquad (1.1)$$

(where J_μ is the weak current associated with the b quark decay, p and p_X respectively denote the four-momenta of the B meson and of the X state, while q is the transferred four-momentum) to the forward scattering matrix element $T_{\mu\nu}(p,q)$ [10, 11], and by extracting the latter through an analytical continuation of lattice results obtained for this quantity in an unphysical region, where the decay is forbidden by kinematics.

In ref. [12], on the other hand, it was proposed to study decay and transition rates into final states with an arbitrary number of hadrons by reconstructing the spectral function associated with a Euclidean four-point function in a finite volume from lattice correlators, with an appropriate smoothing protocol. A closely related approach was discussed in ref. [13] (see also refs. [14, 15]).

Finally, in ref. [16] it was suggested to study inclusive decays on the lattice by computing a suitable "smeared" spectral density $\rho(w)$ of hadron correlators, where the smearing is defined by the integration over the allowed phase-space region. Also in this case, the strategy involves the lattice determination of a class of four-point correlation functions. This technique allows one to bypass the need for analytical continuation, and, at least in principle, paves the way for the determination of the total semileptonic width as well as of the moments of any kinematic distribution associated with general $B \to X\ell\nu$ decays.

In the present work, we focus on the method proposed in ref. [16], presenting the results of explicit lattice calculations based upon this framework. We discuss results from two different types of ensembles of lattice QCD configurations, and we also compare them with

an analytical calculation based on the operator-product expansion (OPE) [17, 18] within the framework of an expansion in inverse powers of the heavy-quark mass [11, 19, 20].

The structure of this article is the following. In section 2, we recapitulate the formulation of the method, extending the presentation in ref. [16] with additional remarks, and commenting on its application to observables of particular interest (including differential distributions and moments). In section 3, we present an explicit implementation of the method in lattice QCD calculations, using two different ensembles of configurations, generated by the JLQCD collaboration and by the ETM collaboration; the final part of the section is devoted to a technical discussion about the extrapolation to the limit in which the smearing parameter σ tends to zero. The following section 4 presents the analytical calculation based on the OPE, and compares its predictions with the results from lattice QCD. Finally, in section 5 we summarize our results and discuss future prospects.

2 Formulation of the method and application to observables

2.1 Spectral representation of the inclusive decay rate

Here we review the formalism to calculate the inclusive semileptonic decay rate in lattice QCD [16]. To be specific, we consider the semileptonic decay of a B meson to charmed final states X_c with a pair of massless leptons ($\ell\bar{\nu}$) through the flavour-changing current $J_\mu = V_\mu - A_\mu = \bar{c}\gamma_\mu(1 - \gamma_5)b$.

We start from the differential decay rate

$$\frac{d\Gamma}{dq^2 dq^0 dE_\ell} = \frac{G_F^2 |V_{cb}|^2}{8\pi^3} L_{\mu\nu} W^{\mu\nu} \,, \tag{2.1}$$

where G_F is the Fermi constant and $|V_{cb}|$ is the relevant CKM matrix element. Here we work in the rest frame of the initial B meson, so that

$$p = (m_B, \mathbf{0}) \,, \qquad q = p_\ell + p_{\bar{\nu}} = (q_0, \boldsymbol{q}) \,, \qquad r = p - q = (\omega, -\boldsymbol{q}) \,, \tag{2.2}$$

where the differential decay rate is a function of the three kinematical variables q^2, q_0 and the lepton energy E_ℓ, and is proportional to the product of the leptonic tensor,

$$L^{\mu\nu} = p_\ell^\mu p_{\bar{\nu}}^\nu - p_\ell \cdot p_{\bar{\nu}} g^{\mu\nu} + p_\ell^\nu p_{\bar{\nu}}^\mu - i\epsilon^{\mu\alpha\nu\beta} p_{\ell,\alpha} p_{\bar{\nu},\beta} \,, \tag{2.3}$$

and the hadronic tensor,

$$W^{\mu\nu}(p, q) = \sum_{X_c} (2\pi)^3 \delta^{(4)}(p - q - r) \frac{1}{2E_B(\boldsymbol{p})} \langle \bar{B}(\boldsymbol{p})|J^{\mu\dagger}(0)|X_c(\boldsymbol{r})\rangle \langle X_c(\boldsymbol{r})|J^\nu(0)|\bar{B}(\boldsymbol{p})\rangle. \tag{2.4}$$

The sum over the charmed states $X_c(\boldsymbol{r})$ actually includes an integral over $\boldsymbol{r}$.[1]

---

[1] More precisely, the sum over the charmed states $X_c(\boldsymbol{r})$ should be written as

$$\sum_{X_c} |X_c(\boldsymbol{r})\rangle\langle X_c(\boldsymbol{r})| \to \sum_{X_c} \int \frac{d^3\boldsymbol{r}}{(2\pi)^3} \frac{1}{2E_{X_c(\boldsymbol{r})}} |X_c(\boldsymbol{r})\rangle\langle X_c(\boldsymbol{r})|$$

when the standard relativistic normalization for a single-particle state is employed for X_c.

Performing the integral over the lepton energy E_ℓ in its kinematical range, i.e. from $(q_0 - \sqrt{\boldsymbol{q}^2})/2$ to $(q_0 + \sqrt{\boldsymbol{q}^2})/2$, and changing the remaining kinematical variables from (q_0, q^2) to $(\omega, \boldsymbol{q}^2)$, the total rate can be written as

$$\Gamma = \frac{G_F^2 |V_{cb}|^2}{24\pi^3} \int_0^{\boldsymbol{q}^2_{\max}} d\boldsymbol{q}^2 \sqrt{\boldsymbol{q}^2} \bar{X}(\boldsymbol{q}^2) \,, \qquad \bar{X}(\boldsymbol{q}^2) \equiv \int_{\omega_{\min}}^{\omega_{\max}} d\omega \, X(\omega, \boldsymbol{q}^2) \,, \qquad (2.5)$$

where

$$\boldsymbol{q}^2_{\max} = \frac{(m_B^2 - m_D^2)^2}{4m_B^2} \,, \qquad \omega_{\min} = \sqrt{m_D^2 + \boldsymbol{q}^2} \,, \qquad \omega_{\max} = m_B - \sqrt{\boldsymbol{q}^2} \,. \qquad (2.6)$$

The quantity $X(\omega, \boldsymbol{q}^2)$ appearing above is a linear combination, with coefficients depending on the kinematical variables, of the different components of the hadronic tensor. Indeed, Lorentz invariance and time-reversal symmetry allow one to decompose $W^{\mu\nu}$ into invariant structure functions according to

$$W^{\mu\nu}(p,q) = -g^{\mu\nu} W_1(\omega, \boldsymbol{q}^2) + \frac{p^\mu p^\nu}{m_B^2} W_2(\omega, \boldsymbol{q}^2) - i\epsilon^{\mu\nu\alpha\beta} \frac{p_\alpha q_\beta}{m_B^2} W_3(\omega, \boldsymbol{q}^2)$$
$$+ \frac{q^\mu q^\nu}{m_B^2} W_4(\omega, \boldsymbol{q}^2) + \frac{p^\mu q^\nu + p^\nu q^\mu}{m_B^2} W_5(\omega, \boldsymbol{q}^2) \,. \qquad (2.7)$$

By introducing the following basis for three-dimensional space,

$$\hat{\boldsymbol{n}} = \frac{\boldsymbol{q}}{\sqrt{\boldsymbol{q}^2}} \,, \qquad \boldsymbol{\epsilon}^{(a)} \cdot \hat{\boldsymbol{n}} = 0 \,, \qquad \boldsymbol{\epsilon}^{(a)} \cdot \boldsymbol{\epsilon}^{(b)} = \delta^{ab} \,, \qquad \{a, b\} = \{1, 2\} \,, \qquad (2.8)$$

and the hadronic quantities

$$Y^{(1)} = -\sum_{a=1}^2 \sum_{i,j=1}^3 \epsilon_i^{(a)} \epsilon_j^{(a)} W^{ij} \,, \qquad Y^{(2)} = W^{00} \,, \qquad Y^{(3)} = \sum_{i,j=1}^3 \hat{n}^i \hat{n}^j W^{ij} \,,$$
$$Y^{(4)} = \sum_{i=1}^3 \hat{n}^i (W^{0i} + W^{i0}) \,, \qquad Y^{(5)} = \frac{i}{2} \sum_{i,j,k=1}^3 \varepsilon^{ijk} \hat{n}^k W^{ij} \,, \qquad (2.9)$$

it is easy to see that, in the rest frame of the B meson, the information contained in $W^{\mu\nu}$ can be equivalently parametrized in terms of $Y^{(i)} \equiv Y^{(i)}(\omega, \boldsymbol{q}^2)$. A convenient representation of $X(\omega, \boldsymbol{q}^2)$ is then given by

$$X(\omega, \boldsymbol{q}^2) = \sum_{l=0}^2 \left(\sqrt{\boldsymbol{q}^2}\right)^{2-l} (m_B - \omega)^l X^{(l)}(\omega, \boldsymbol{q}^2) \,,$$
$$X^{(0)} = Y^{(1)} + Y^{(2)} \,, \qquad X^{(1)} = -Y^{(4)} \,, \qquad X^{(2)} = Y^{(3)} - Y^{(1)} \,. \qquad (2.10)$$

At this point, some observations are in order. First, we notice that the parity-violating structure function W_3 (or equivalently $Y^{(5)}$) does not contribute to the differential decay rate after the integral over E_ℓ has been performed (this will not be the case for the moments considered below). Then, by rewriting eq. (2.4) as

$$W_{\mu\nu}(\omega, \boldsymbol{q}) = \frac{(2\pi)^3}{2m_B} \langle \bar{B}(\boldsymbol{0}) | J_\mu^\dagger(0) \delta(\hat{H} - \omega) \delta^3(\hat{\boldsymbol{P}} + \boldsymbol{q}) J_\nu(0) | \bar{B}(\boldsymbol{0}) \rangle \,, \qquad (2.11)$$

where $\hat{H}$ and $\hat{\boldsymbol{P}}$ are the QCD Hamiltonian and momentum operators, we explicitly see that in the rest frame of the B meson the different components of the hadronic tensor are functions of ω and $\boldsymbol{q}$ (we already used this information in changing the integration variables from (q_0, q^2) to $(\omega, \boldsymbol{q}^2)$). In the following we refer to eq. (2.11) as the *spectral* representation of the hadronic tensor. Flavour and momentum conservation imply that the hadronic tensor vanishes identically for energies $\omega < \omega_{\min}$. From this observation one obtains $\omega_{\min} = \sqrt{m_D^2 + \boldsymbol{q}^2}$. This means that by introducing the kernels

$$K^{(l)}(\omega, \boldsymbol{q}^2) = (m_B - \omega)^l \theta(\omega_{\max} - \omega)\,, \tag{2.12}$$

the ω integral in eq. (2.5) can be rewritten as

$$\bar{X}(\boldsymbol{q}^2) = \sum_{l=0}^{2} \left(\sqrt{\boldsymbol{q}^2}\right)^{2-l} \int_0^\infty d\omega \, K^{(l)}(\omega, \boldsymbol{q}^2) \, X^{(l)}(\omega, \boldsymbol{q}^2)\,. \tag{2.13}$$

We close this subsection by providing an equivalent representation of $X(\omega, \boldsymbol{q}^2)$ which will be useful later. It is given by

$$X(\omega, \boldsymbol{q}^2) = \sum_{l=0}^{2} \left(\sqrt{\boldsymbol{q}^2}\right)^{2-l} (\omega_{\max} - \omega)^l Z^{(l)}(\omega, \boldsymbol{q}^2)\,, \tag{2.14}$$

where the $Z^{(l)}(\omega, \boldsymbol{q}^2)$ are again linear combinations of the $Y^{(l)}$,

$$Z^{(0)} = Y^{(2)} + Y^{(3)} - Y^{(4)}\,, \quad Z^{(1)} = 2Y^{(3)} - 2Y^{(1)} - Y^{(4)}\,, \quad Z^{(2)} = Y^{(3)} - Y^{(1)}\,. \tag{2.15}$$

By introducing the kernels

$$\Theta^{(l)}(x) = x^l \theta(x)\,, \tag{2.16}$$

that are functions of the single variable $x = \omega_{\max} - \omega$, we thus have

$$\bar{X}(\boldsymbol{q}^2) = \sum_{l=0}^{2} \left(\sqrt{\boldsymbol{q}^2}\right)^{2-l} Z^{(l)}(\boldsymbol{q}^2)\,, \quad Z^{(l)}(\boldsymbol{q}^2) = \int_0^\infty d\omega \, \Theta^{(l)}(\omega_{\max} - \omega) \, Z^{(l)}(\omega, \boldsymbol{q}^2)\,. \tag{2.17}$$

2.2 Decay rate from Euclidean correlators

In order to calculate $\bar{X}(\boldsymbol{q}^2)$, as given in eq. (2.13) or eq. (2.17), we need to evaluate the integral over ω of the different components of the hadronic spectral density (2.11) with the kernels $K^{(l)}(\omega, \boldsymbol{q}^2)$ or $\Theta^{(l)}(\omega_{\max} - \omega)$. To this end, following ref. [9], we first establish the connection between suitably chosen correlation functions that can be calculated on the lattice, and $W_{\mu\nu}$.

We start by considering the Euclidean correlator

$$C_{\mu\nu}(t_{\text{snk}}, t_2, t_1, t_{\text{src}}; \boldsymbol{q}) = \int d^3x \, e^{i\boldsymbol{q}\cdot\boldsymbol{x}} \, T\langle 0| \, \tilde{\phi}_B(\boldsymbol{0}; t_{\text{snk}}) J_\mu^\dagger(\boldsymbol{x}; t_2) J_\nu(\boldsymbol{0}; t_1) \tilde{\phi}_B^\dagger(\boldsymbol{0}; t_{\text{src}}) |0\rangle\,, \tag{2.18}$$

where $\tilde{\phi}_B(\boldsymbol{0}; t)$ is a B-meson creation/annihilation operator projected onto zero spatial momentum by integrating over space at a time t. A zero-momentum B meson is thus

created at time t_{src} and annihilated at t_{snk}. The two currents are inserted in between, at times t_2 and t_1. The charmed hadrons are created at time t_1 with a momentum insertion $-\boldsymbol{q}$ and propagate until they are transformed back to the B-meson state at time t_2.

The four-point function $C_{\mu\nu}$ is saturated by the B-meson non-local matrix element

$$M_{\mu\nu}(t; \boldsymbol{q}) = e^{-m_B t} \int d^3 x \, \frac{e^{i\boldsymbol{q}\cdot\boldsymbol{x}}}{2m_B} \langle \bar{B}(\boldsymbol{0}) | J_\mu^\dagger(\boldsymbol{x},t) J_\nu(\boldsymbol{0},0) | \bar{B}(\boldsymbol{0}) \rangle \,, \tag{2.19}$$

when the double limit $t_{\text{src}} \to -\infty$, $t_{\text{snk}} \to \infty$ is taken. To include a proper normalization, one can analyse

$$M_{\mu\nu}(t_2 - t_1; \boldsymbol{q}) = Z_B \lim_{\substack{t_{\text{snk}} \to +\infty \\ t_{\text{src}} \to -\infty}} \frac{C_{\mu\nu}(t_{\text{snk}}, t_2, t_1, t_{\text{src}}; \boldsymbol{q})}{C(t_{\text{snk}} - t_2) C(t_1 - t_{\text{src}})} \,, \tag{2.20}$$

where $C(t)$ is the B-meson two-point function

$$C(t) = T \langle 0 | \, \tilde{\phi}_B(\boldsymbol{0}; t) \tilde{\phi}_B^\dagger(\boldsymbol{0}; 0) | 0 \rangle \tag{2.21}$$

and Z_B is its residue when a large time separation is taken, $C(t) \to Z_B e^{-m_B t}$.

Starting from eq. (2.19) we can establish the connection between $M_{\mu\nu}(t; \boldsymbol{q})$ and the hadronic tensor given in eq. (2.11). We have

$$\begin{aligned}
M_{\mu\nu}(t; \boldsymbol{q}) &= \int d^3 x \, \frac{e^{i\boldsymbol{q}\cdot\boldsymbol{x}}}{2m_B} \langle \bar{B}(\boldsymbol{0}) | J_\mu^\dagger(\boldsymbol{0},0) e^{-t\hat{H} + i\hat{\boldsymbol{P}}\cdot\boldsymbol{x}} J_\nu(\boldsymbol{0},0) | \bar{B}(\boldsymbol{0}) \rangle \\
&= \frac{(2\pi)^3}{2m_B} \langle \bar{B}(\boldsymbol{0}) | J_\mu^\dagger(\boldsymbol{0},0) e^{-t\hat{H}} \delta^3(\hat{\boldsymbol{P}} + \boldsymbol{q}) J_\nu(\boldsymbol{0},0) | \bar{B}(\boldsymbol{0}) \rangle \\
&= \int_0^\infty d\omega \, W_{\mu\nu}(\omega, \boldsymbol{q}) \, e^{-\omega t} \,.
\end{aligned} \tag{2.22}$$

The problem of the calculation of $\bar{X}(\boldsymbol{q}^2)$ is now reduced to that of trading the integral of $W_{\mu\nu}(\omega, \boldsymbol{q})$ with the kernels $e^{-t\omega}$ for the integral with the kernels $\Theta^{(l)}(\omega_{\max} - \omega)$ (or $K^{(l)}(\omega, \boldsymbol{q}^2)$).

The general inverse problem represented by the extraction of hadronic spectral densities from Euclidean correlators is notoriously ill-posed. Recently, methods to cope with these problems have been proposed, and they treat the above mentioned integrals with some kernels. In this paper we use two approaches proposed in refs. [13, 21]. The differences between the two methods will be discussed in detail in the following sections. Here we concentrate on the common starting point of the two approaches, which are actually closely related to each other.

We start by introducing an arbitrary length scale a. On the lattice this will be identified with the lattice spacing. The correlators $M_{\mu\nu}(t; \boldsymbol{q})$ will be computed at times $t = a\tau$ where $\tau \geq 0$ is an integer. By introducing the variable $x = e^{-a\omega}$ (and its inverse mapping $\omega = -\log(x)/a$), standard theorems of numerical analysis guarantee that any C_∞ function $f(\omega) \equiv g(x)$ in the interval $\omega \in [0, \infty]$ (corresponding to $x \in [0, 1]$), vanishing at $\omega = \infty$ ($x = 0$), can be approximated with arbitrary precision in terms of polynomials in x according to

$$f(\omega) = \sum_{\tau=1}^\infty g_\tau \, x^\tau \equiv \sum_{\tau=1}^\infty g_\tau \, e^{-a\omega\tau} \,. \tag{2.23}$$

This implies that the integral of the product of $W_{\mu\nu}(\omega, \boldsymbol{q})$ with $f(\omega)$ can be computed, once the coefficients g_τ are known, by using the linear relation

$$\int_0^\infty d\omega \, W_{\mu\nu}(\omega, \boldsymbol{q}) \, f(\omega) = \sum_{\tau=1}^\infty g_\tau \, M_{\mu\nu}(a\tau; \boldsymbol{q}) \,. \tag{2.24}$$

This procedure cannot be applied straightforwardly to the calculation of $X(\boldsymbol{q}^2)$ because the kernels $\Theta^{(l)}(\omega_{\max} - \omega)$ (or $K^{(l)}(\omega, \boldsymbol{q}^2)$) are not smooth, i.e. they contain a discontinuity due to the θ-function. In this case, a sequence of polynomials can still converge to the kernels *in mean*, which would be sufficient for our purposes, but a reasonable approximation would imply a very large number of terms. However, the problem can be solved by introducing smeared C_∞ versions of the θ-function, θ_σ, such that the sharp step-function is recovered in the limit in which the smearing parameter σ is sent to zero, $\lim_{\sigma\to 0} \theta_\sigma(x) = \theta(x)$.

By considering, as suggested in ref. [16], the corresponding smeared versions of the kernels entering the definition of $\bar{X}(\boldsymbol{q}^2)$, that we call $\Theta_\sigma^{(l)}(\omega_{\max} - \omega)$ and $K_\sigma^{(l)}(\omega, \boldsymbol{q})$, we then have

$$\Theta_\sigma^{(l)}(\omega_{\max} - \omega) = m_B^l \sum_{\tau=1}^\infty g_\tau^{(l)}(\omega_{\max}, \sigma) \, e^{-a\omega\tau} \,, \tag{2.25}$$

and

$$\int_0^\infty d\omega \, W_{\mu\nu}(\omega, \boldsymbol{q}) \, \Theta^{(l)}(\omega_{\max} - \omega) = \lim_{\sigma\to 0} m_B^l \sum_{\tau=1}^\infty g_\tau^{(l)}(\omega_{\max}, \sigma) \, M_{\mu\nu}(a\tau; \boldsymbol{q}) \,, \tag{2.26}$$

as well as similar relations in the case of the kernels $K^{(l)}(\omega, \boldsymbol{q}^2)$.

A few observations are now in order. The first concerns a subtle theoretical issue. The smearing procedure, which is algorithmically required to implement the procedure just outlined, is also necessary for theoretical reasons. Hadronic spectral densities, and therefore also $W_{\mu\nu}(\omega, \boldsymbol{q})$, are elements in the space of distributions and their product with another distribution, such as the θ-function, can only be defined through a regularization procedure (when it exists). The issue is particularly important in the case of lattice simulations because they are necessarily performed in a finite volume. Finite-volume spectral functions, due to the quantization of the energy spectrum, are sums of isolated δ-function singularities and their connection with the corresponding physical quantities requires an ordered double-limit procedure: first the infinite volume limit has to be taken and only *after* that, if the quantity is non-singular, can one take the $\sigma \to 0$ limit.

The second observation is related to the fact that the problem we are addressing is particularly hard from the computational point of view. In the limit of very small σ the coefficients $g_\tau^{(l)}(\omega_{\max}, \sigma)$ of eq. (2.25) tend to become arbitrarily large in modulus and oscillate in sign. Since lattice correlators are unavoidably affected by statistical and systematic errors, in these cases the resulting uncertainties on the sums on the left-hand side of eq. (2.26) tend to explode. The two approaches of refs. [13, 21] differ for the procedures used to determine the coefficients $g_\tau^{(l)}(\omega_{\max}, \sigma)$, once the series is truncated at $\tau = \tau_{\max}$, in such a way to keep both statistical and systematic errors under control.

– 7 –

2.3 Kernel approximation

In this subsection we review the methods of refs. [13, 21] by highlighting the differences in the procedures used to approximate the smearing kernels. To simplify the formulae, we shall consider a generic kernel $f(\omega)$, that will then be identified with the kernels $\Theta^{(l)}(\omega_{\max} - \omega)/m_B^l$ or $K^{(l)}(\omega, \boldsymbol{q})/m_B^l$, and a generic correlator

$$C(t) = \int_0^\infty d\omega\, \rho(\omega)\, e^{-\omega t}\,, \tag{2.27}$$

to be identified with $M_{\mu\nu}(t; \boldsymbol{q})$, so that $\rho(\omega)$ will correspond to $W_{\mu\nu}(\omega, \boldsymbol{q})$. In this work we shall not address the systematics associated with the finiteness of the extent of the lattice in the temporal direction, see refs. [13, 15] for an extended discussion of this issue and, in general, for more details concerning the algorithm and its applications.

In the method of ref. [13] the coefficients g_τ corresponding to the approximation of $f(\omega)$ are determined by minimizing the functional

$$W_\lambda[g] = (1-\lambda)\frac{A[g]}{A[0]} + \lambda B[g]\,, \tag{2.28}$$

where $\lambda \in [0, 1]$ is the so-called "trade-off parameter" (see below) and the functionals $A[g]$ and $B[g]$ are given by

$$A[g] = a \int_{E_0}^\infty d\omega \left\{ f(\omega) - \sum_{\tau=1}^{\tau_{\max}} g_\tau\, e^{-a\omega\tau} \right\}^2\,, \quad B[g] = \sum_{\tau,\tau'=1}^{\tau_{\max}} g_\tau g_{\tau'} \frac{\text{Cov}\left[C(a\tau), C(a\tau')\right]}{\left[C(0)\right]^2}\,. \tag{2.29}$$

Here $\text{Cov}\left[C(t), C(t')\right]$ is the statistical covariance of the correlator $C(t)$ and, consequently, the functional $B[g]$ is positive definite. The functional $A[g]$ is also a positive definite quadratic form in the coefficients g_τ. Therefore, the minimum conditions

$$\left.\frac{\partial W_\lambda[g]}{\partial g_\tau}\right|_{g_\tau = g_\tau^\lambda} = 0 \tag{2.30}$$

are a linear system of equations to be solved for the coefficients g_τ^λ. These coefficients define the approximation of $f(\omega)$ and the associated estimator for the integral of $\rho(\omega)$ with $f(\omega)$ according to

$$f^\lambda(\omega) = \sum_{\tau=1}^{\tau_{\max}} g_\tau^\lambda\, e^{-a\omega\tau}\,, \qquad \rho^\lambda[f] = \sum_{\tau=1}^{\tau_{\max}} g_\tau^\lambda C(a\tau) = \int_0^\infty d\omega\, f^\lambda(\omega)\, \rho(\omega)\,. \tag{2.31}$$

The functional $B[g^\lambda]$ is the statistical variance of $\rho^\lambda[f]$ normalized with the square of the correlator in zero and, therefore, vanishes in the ideal case of infinitely precise input data. On the other hand, $A[g^\lambda]$ measures the distance between the target kernel $f(\omega)$ and its approximation $f^\lambda(\omega)$ in the range[2] $\omega \in [E_0, \infty]$. In fact $A[g^\lambda]$ is the squared L_2-norm in function space of the difference $f^\lambda(\omega) - f(\omega)$ and can only vanish in the limit $t_{\max} \to \infty$.

---

[2]The parameter E_0 can be adjusted by exploiting the fact that $\rho(\omega)$ has support only for $\omega > \omega_{\min}$, so that $\rho[f]$ is insensitive to $f(\omega)$ for $\omega < \omega_{\min}$. The same holds for $\rho^\lambda[f]$ so that the functional form of $f^\lambda(\omega)$ can be left unconstrained for $\omega < \omega_{\min}$. Any $E_0 < \omega_{\min}$ is therefore a viable choice in determining the coefficients g_t^λ so E_0 can be chosen to improve the numerical stability of the minimization procedure.

In the absence of errors, the coefficients g_τ^λ that minimize $A[g]$ provide the best polynomial approximation of $f(\omega)$ with respect to the L_2-norm. This has to be compared with the method of ref. [21] that provides the best polynomial approximation of $f(\omega)$ with respect to the L_∞-norm (see below). In the presence of errors, the coefficients g_τ^λ that minimize $W_\lambda[g]$ represent a particular balance between statistical and systematic errors, as dictated by the λ parameter. For small λ the estimator $\rho^\lambda[f]$ is close to $\rho[f]$ but with a large statistical uncertainty. Conversely, for large λ the estimator $\rho^\lambda[f]$ has a small statistical error but differs significantly from $\rho[f]$. When evaluated at the minimum, the functional $W_\lambda[g]$ is a function of λ only, thus defining $W(\lambda) \equiv W_\lambda[g^\lambda]$. The prescription suggested in ref. [13] to choose the optimal value of the trade-off parameter defines $\lambda_\star$ such that

$$\left.\frac{\partial W(\lambda)}{\partial \lambda}\right|_{\lambda=\lambda_\star} = 0\,. \tag{2.32}$$

From eq. (2.30) it follows that at $\lambda_\star$ (the maximum of $W(\lambda)$ where $g_\tau^\lambda = g_\tau^\star$) one has $A[g^\star] = A[0]B[g^\star]$. This can be understood as the condition of "optimal balance" between statistical and systematic errors. The numerical results discussed in subsection 3.2 have been obtained using this method, also monitoring the stability of the results with respect to $\lambda \leq \lambda_\star$.

In ref. [21], on the other hand, the function $f(\omega)$ is approximated using the Chebyshev approximation as

$$f(\omega) \simeq \frac{c_0^*}{2} + \sum_{j=1}^{N} c_j^* T_j^*(e^{-a\omega})\,, \tag{2.33}$$

where $T_j^*(x)$ is a (shifted) Chebyshev polynomial of the j-th order. The coefficients c_j^* are determined only by the function $f(\omega)$:

$$c_j^* = \frac{2}{\pi} \int_0^\pi d\theta\, f\left(-\ln\frac{1+\cos\theta}{2}\right)\cos(j\theta)\,. \tag{2.34}$$

This yields the best approximation in the sense of the L_∞-norm.[3] The approximation of the ω-integral is then constructed as

$$\int d\omega f(\omega)\rho(\omega) = \frac{\langle\psi_\mu|f(\hat{H})|\psi_\nu\rangle}{\langle\psi_\mu|\psi_\nu\rangle} \simeq \frac{c_0^*}{2} + \sum_{j=1}^{N} c_j^* \frac{\langle\psi_\mu|T_j^*(e^{-a\hat{H}})|\psi_\nu\rangle}{\langle\psi_\mu|\psi_\nu\rangle}\,, \tag{2.35}$$

where $|\psi_\mu\rangle \equiv e^{-\hat{H}t_0}J_\mu|B\rangle$ is defined such that the state is evolved for some small time t_0 after applying the current insertion: this allows one to avoid any ultraviolet divergence due to contact terms of two currents. To reflect this change, the kernel $f(\omega)$ is multiplied by $e^{2\omega t_0}$ to cancel the time evolution. The right-hand side of eq. (2.35) can be reconstructed from the matrix elements (2.19) using $M_{\mu\nu}(t+2t_0)/M_{\mu\nu}(2t_0)$.

---

[3]More precisely, the Chebyshev approximation of a generic function $f(y)$ for $y \in [-1,1]$ is in practice equivalent, although not identical, to the optimal polynomial approximation of $f(\vec{g};y) = \sum_{\tau=0}^{N} g_\tau y^\tau$ obtained by minimizing the L_∞-norm $\|f(y) - f(\vec{g};y)\|_\infty = \max_{y \in [-1,1]}|f(y) - f(\vec{g};y)|$ with respect to the coefficients $\vec{g}$. In fact, the Chebyshev approximation is obtained by minimizing the *weighted* squared L_2-norm given by $\int_{-1}^{1} dy\, w(y)|f(y) - f(\vec{g};y)|^2$ with $w(y) = 1/\sqrt{1-y^2}$. By setting instead $w(y) = 1$, as done in the case of the method of ref. [13], one gets the Legendre polynomial approximation.

An advantage of this construction is that the matrix element appearing on the right-hand side of eq. (2.35), $\langle\psi_\mu|T_j^*(e^{-a\hat{H}})|\psi_\nu\rangle/\langle\psi_\mu|\psi_\nu\rangle$, is strictly bound between -1 and $+1$, by construction of the Chebyshev polynomial. This corresponds to the condition that the eigenvalues of $e^{-\hat{H}}$ lie between 0 and 1, or equivalently that the eigenvalues of $\hat{H}$ are positive semi-definite. Then, the convergence of the series appearing in eq. (2.35) is dictated by that of the coefficients c_j^*. Since c_j^* can be easily calculated for arbitrarily large j's, the error due to the truncation in (2.35) can be rigorously estimated.

The constraint $|\langle\psi_\mu|T_j^*(e^{-a\hat{H}})|\psi_\nu\rangle/\langle\psi_\mu|\psi_\nu\rangle| \leq 1$ is not automatically satisfied in the presence of statistical errors. Since the Chebyshev polynomial $T_j^*(x)$ is a sign-alternating series of growing powers of x with (exponentially) large coefficients, this constraint is satisfied after huge cancellations for large j. Therefore, even a small statistical error of the lattice correlator can easily violate the constraint. In the numerical analysis, one should add the constraint when the Chebyshev matrix elements are determined by a fit, see ref. [21] for details. The higher-order terms are then masked by the statistical uncertainties and become basically undetermined within ± 1, so that they only contribute to the truncation error.

In both methods, a good approximation is obtained only when the kernel function is sufficiently smooth. If this is not the case, the truncation error becomes significant, e.g. due to unsuppressed higher-order coefficients c_j^* in the case of the Chebyshev approximation. Unfortunately, the kernel functions $K^{(l)}(\omega, \boldsymbol{q})$ or $\Theta^{(l)}(\omega_{\max} - \omega)$ are not smooth, because they contain the Heaviside function $\theta(\omega_{\max} - \omega)$. We therefore introduce smeared versions of the θ-function and then we take the limit of $\sigma \to 0$ to recover the unsmeared kernel. This has been done by considering three different smeared θ-functions,

$$\theta_\sigma^{\mathsf{s}}(x) = \frac{1}{1 + e^{-\frac{x}{\sigma}}}, \qquad \theta_\sigma^{\mathsf{s1}}(x) = \frac{1}{1 + e^{-\sinh\left(\frac{x}{r^{\mathsf{s1}}\sigma}\right)}}, \qquad \theta_\sigma^{\mathsf{e}}(x) = \frac{1 + \operatorname{erf}\left(\frac{x}{r^{\mathsf{e}}\sigma}\right)}{2}, \qquad (2.36)$$

and by extrapolating the numerical data to the $\sigma \to 0$ limit. In the following we shall refer to $\theta_\sigma^{\mathsf{s}}(x)$ as the "sigmoid function", to $\theta_\sigma^{\mathsf{s1}}(x)$ as the "modified sigmoid function" and to $\theta_\sigma^{\mathsf{e}}(x)$ as the "error function". Any choice of the parameters r^{s1} and r^{e} appearing in the previous formulae corresponds to a legitimate definition of the smearing kernels that approach the same $\sigma \to 0$ limit, i.e. the θ-function. By adjusting the values of these parameters one can change the rate of convergence to the θ-function and balance between statistical and systematic errors. In the following we set $r^{\mathsf{s1}} = 2.2$ and $r^{\mathsf{e}} = 2.0$. This (empirical) choice gives statistical errors of the same order of magnitude for the three kernels at fixed σ and similar (although not identical) shapes for $\theta_\sigma^{\mathsf{s}}(x)$ and $\theta_\sigma^{\mathsf{e}}(x)$ while $\theta_\sigma^{\mathsf{s1}}(x)$ results into a smoother approximation of the θ-function. A combined analysis of smearing kernels that have rather different shapes at fixed σ is in fact helpful in order to quantify the systematics associated with the $\sigma \to 0$ extrapolations (see also ref. [15]).

2.4 Decomposition of the total rate

The expression of the total rate in eq. (2.5) can also be used to compute the differential decay rate in $\boldsymbol{q}^2$, i.e. $d\Gamma/d\boldsymbol{q}^2 = G_F^2|V_{cb}|^2/(24\pi^3)|\boldsymbol{q}|\bar{X}(\boldsymbol{q}^2)$. This can be further decomposed into its contributions from parallel ($\parallel$) and perpendicular ($\perp$) components, where the $\perp$ components are defined as those involving the polarization vector $\epsilon^{*(\alpha)}$, while the $\parallel$ ones are

the rest. In addition, we also separate the contributions from vector (V) and axial-vector (A) current insertions. Since two currents are inserted, we have VV, AA, as well as VA and AV contributions. Among them, VA and AV do not contribute to the differential rate after integrating over E_ℓ, and thus to the total decay rate. We therefore analyze four components: $VV_\parallel$, $VV_\perp$, $AA_\parallel$, $AA_\perp$. For the lepton energy moments, the VA and AV insertions can also appear (see below).

2.5 Moments

It is also interesting to consider the moments of various kinematical quantities. In particular, two types of moments have been studied experimentally: the hadronic mass moments $\langle (M_X^2)^n \rangle$ and the lepton energy moments $\langle E_\ell^{n_\ell} \rangle$. They are defined as

$$\langle (M_X^2)^n \rangle = \frac{\int d\boldsymbol{q}^2 dq_0 dE_\ell \, (\omega^2 - \boldsymbol{q}^2)^n \left[\dfrac{d\Gamma}{d\boldsymbol{q}^2 dq_0 dE_\ell} \right]}{\int d\boldsymbol{q}^2 dq_0 dE_\ell \left[\dfrac{d\Gamma}{d\boldsymbol{q}^2 dq_0 dE_\ell} \right]}, \tag{2.37}$$

$$\langle E_\ell^{n_\ell} \rangle = \frac{\int d\boldsymbol{q}^2 dq_0 dE_\ell \, E_\ell^{n_\ell} \left[\dfrac{d\Gamma}{d\boldsymbol{q}^2 dq_0 dE_\ell} \right]}{\int d\boldsymbol{q}^2 dq_0 dE_\ell \left[\dfrac{d\Gamma}{d\boldsymbol{q}^2 dq_0 dE_\ell} \right]}. \tag{2.38}$$

The strategy to compute these moments on the lattice is the same as in the method described above. For the hadronic mass moments defined in eq. (2.37), the numerator contains extra powers of $\omega^2 - \boldsymbol{q}^2$, with which the ω-dependence of $X^{(0)}$, $X^{(1)}$, $X^{(2)}$ is modified. Otherwise, the basic procedure remains the same. Beside these quantities which require an integration over the whole $\boldsymbol{q}^2$ range, we will also consider moments at fixed values of $\boldsymbol{q}^2$, i.e. *differential moments*:

$$H_n(\boldsymbol{q}^2) \equiv \langle (M_X^2)^n \rangle_{\boldsymbol{q}^2} = \frac{\int dq_0 dE_\ell \, (\omega^2 - \boldsymbol{q}^2)^n \left[\dfrac{d\Gamma}{d\boldsymbol{q}^2 dq_0 dE_\ell} \right]}{\int dq_0 dE_\ell \left[\dfrac{d\Gamma}{d\boldsymbol{q}^2 dq_0 dE_\ell} \right]}, \tag{2.39}$$

$$L_{n_\ell}(\boldsymbol{q}^2) \equiv \langle E_\ell^{n_\ell} \rangle_{\boldsymbol{q}^2} = \frac{\int dq_0 dE_\ell \, E_\ell^{n_\ell} \left[\dfrac{d\Gamma}{d\boldsymbol{q}^2 dq_0 dE_\ell} \right]}{\int dq_0 dE_\ell \left[\dfrac{d\Gamma}{d\boldsymbol{q}^2 dq_0 dE_\ell} \right]}, \tag{2.40}$$

and the second central moment or variance of the lepton energy distribution

$$L_{2c}(\boldsymbol{q}^2) = L_2(\boldsymbol{q}^2) - \left(L_1(\boldsymbol{q}^2) \right)^2.$$

In the case of leptonic moments, the E_ℓ integral is modified with respect to (2.5). The integrand in the denominator is the same as in (2.10); if we set the $\boldsymbol{q}$ momentum direction $\boldsymbol{n}$ along the k-th axis, the two vectors $\boldsymbol{\epsilon}^a$ can be chosen in the perpendicular directions of the i-th and j-th axes, and we can re-express $X(\omega, \boldsymbol{q}^2)$ as

$$X_{n_\ell=0} = \boldsymbol{q}^2 (W^{00} - W^{ii} - W^{jj}) - q_0 q_k (W^{0k} + W^{k0}) + q_0^2 (W^{kk} + W^{ii} + W^{jj}), \tag{2.41}$$

where repeated indices are not summed. The integrand in the numerators of eq. (2.38) and eq, (2.40) depends on the exponent n_ℓ. For $n_\ell = 1$, it reads

$$
\begin{aligned}
X_{n_\ell=1} &= \frac{q_0}{2} \Big[\boldsymbol{q}^2 (W^{00} - W^{ii} - W^{jj}) - q_0 q_k (W^{0k} + W^{k0}) + q_0^2 (W^{kk} + W^{ii} + W^{jj}) \Big] \\
&\quad + \frac{i}{2} q_k (q_0^2 - \boldsymbol{q}^2) W^{ij} \,,
\end{aligned}
\tag{2.42}
$$

where the last term corresponds to the insertion of VA or AV. The other terms are the same as $X_{n_\ell=0}$, up to a factor $q_0/2$. The next orders are more involved:

$$
\begin{aligned}
X_{n_\ell=2} &= \frac{1}{4} \left\{ \left(q_0^2 \boldsymbol{q}^2 + \frac{1}{5} |\boldsymbol{q}|^4 \right) W^{00} + \left(2q_0^4 - \frac{6}{5} q_0^2 \boldsymbol{q}^2 - \frac{4}{5} |\boldsymbol{q}|^4 \right) W^{ii} + \left(q_0^4 + \frac{1}{5} q_0^2 \boldsymbol{q}^2 \right) W^{kk} \right. \\
&\quad \left. - \left(q_0^3 |\boldsymbol{q}| + \frac{1}{5} q_0 |\boldsymbol{q}|^3 \right) (W^{0k} + W^{k0}) + \frac{i}{2} q_0 |\boldsymbol{q}| (q_0^2 - \boldsymbol{q}^2) W^{ij} \right\} \,,
\end{aligned}
\tag{2.43}
$$

$$
\begin{aligned}
X_{n_\ell=3} &= \frac{1}{8} \left\{ \left(q_0^3 \boldsymbol{q}^2 + \frac{3}{5} q_0 |\boldsymbol{q}|^4 \right) W^{00} + \left(q_0^5 + \frac{3}{5} q_0^3 \boldsymbol{q}^2 \right) W^{kk} + \left(2q_0^5 + \frac{2}{5} q_0^3 \boldsymbol{q}^2 - \frac{12}{5} q_0 |\boldsymbol{q}|^4 \right) W^{ii} \right. \\
&\quad \left. - \left(q_0^4 |\boldsymbol{q}| + \frac{3}{5} q_0^2 |\boldsymbol{q}|^3 \right) (W^{0k} + W^{k0}) + i \left(3q_0^2 |\boldsymbol{q}| + \frac{3}{5} |\boldsymbol{q}|^3 \right) (q_0^2 - \boldsymbol{q}^2) W^{ij} \right\} \,.
\end{aligned}
\tag{2.44}
$$

Again, the term with W^{ij} survives for VA and AV insertions, while the others are from VV or AA.

The contributions in eq. (2.42) can be rearranged in such a way that the ω-integral contributing to the numerator of eq. (2.40) takes the form

$$
\begin{aligned}
\bar{X}_{n_\ell=1}(\boldsymbol{q}^2) &= \sum_{l=0}^{3} \left(\sqrt{\boldsymbol{q}^2} \right)^{3-l} Z_{n_\ell=1}^{(l)}(\boldsymbol{q}^2) \,, \\
Z_{n_\ell=1}^{(l)}(\boldsymbol{q}^2) &= \int_0^\infty d\omega \, \Theta^{(l)}(\omega_{\max} - \omega) \, Z_{n_\ell=1}^{(l)}(\omega, \boldsymbol{q}^2) \,,
\end{aligned}
\tag{2.45}
$$

where the $Z_{n_\ell=1}^{(l)}(\omega, \boldsymbol{q}^2)$ are given by

$$
\begin{aligned}
Z_{n_\ell=1}^{(0)} &= \frac{Y^{(2)} + Y^{(3)} - Y^{(4)}}{2} \,, & Z_{n_\ell=1}^{(1)} &= \frac{-2Y^{(1)} + Y^{(2)} + 3Y^{(3)} - 2Y^{(4)} + 2Y^{(5)}}{2} \,, \\
Z_{n_\ell=1}^{(2)} &= \frac{-3Y^{(1)} + 3Y^{(3)} - Y^{(4)} + Y^{(5)}}{2} \,, & Z_{n_\ell=1}^{(3)} &= \frac{-Y^{(1)} + Y^{(3)}}{2} \,.
\end{aligned}
\tag{2.46}
$$

The previous expressions are analogous to the corresponding expressions for the differential decay rate, eq. (2.17) and eq. (2.15), but include the sum of four terms with the one corresponding to $l = 3$ that involves the kernel $\Theta^{(3)}(\omega_{\max} - \omega)$. In this basis the second leptonic moment is given by

$$
\begin{aligned}
\bar{X}_{n_\ell=2}(\boldsymbol{q}^2) &= \sum_{l=0}^{4} \left(\sqrt{\boldsymbol{q}^2} \right)^{4-l} Z_{n_\ell=2}^{(l)}(\boldsymbol{q}^2) \,, \\
Z_{n_\ell=2}^{(l)}(\boldsymbol{q}^2) &= \int_0^\infty d\omega \, \Theta^{(l)}(\omega_{\max} - \omega) \, Z_{n_\ell=2}^{(l)}(\omega, \boldsymbol{q}^2) \,,
\end{aligned}
\tag{2.47}
$$

where

$$Z^{(0)}_{n_\ell=2} = 3\frac{Y^{(2)} + Y^{(3)} - Y^{(4)}}{10},$$

$$Z^{(1)}_{n_\ell=2} = \frac{7Y^{(1)} - 5Y^{(2)} - 11Y^{(3)} + 8Y^{(4)} - 10Y^{(10)}}{10},$$

$$Z^{(2)}_{n_\ell=2} = \frac{-27Y^{(1)} + 5Y^{(2)} + 31Y^{(3)} - 15Y^{(4)} + 30Y^{(5)}}{20},$$

$$Z^{(3)}_{n_\ell=2} = \frac{4Y^{(1)} - 4Y^{(3)} + Y^{(4)} - 2Y^{(5)}}{4},$$

$$Z^{(4)}_{n_\ell=2} = \frac{-Y^{(1)} + Y^{(3)}}{4}, \tag{2.48}$$

and the first hadronic moment is

$$\bar{X}_{n=1}(\boldsymbol{q}^2) = \sum_{l=0}^{4} Z^{(l)}_{n=1}(\boldsymbol{q}^2), \qquad Z^{(l)}_{n=1}(\boldsymbol{q}^2) = \int_0^\infty d\omega\, \Theta^{(l)}(\omega_{\max} - \omega)\, Z^{(l)}_{n=1}(\omega, \boldsymbol{q}^2), \tag{2.49}$$

where the $Z^{(l)}_{n=1}(\omega, \boldsymbol{q}^2)$ are given by

$$Z^{(0)}_{n=1} = m_B|\boldsymbol{q}|^3(m_B - 2|\boldsymbol{q}|)\left(Y^{(2)} + Y^{(3)} - Y^{(4)}\right),$$

$$Z^{(1)}_{n=1} = 2|\boldsymbol{q}|^4\left(Y^{(2)} + Y^{(3)} - Y^{(4)}\right) + m_B^2|\boldsymbol{q}|^2\left(-2Y^{(1)} + 2Y^{(3)} - Y^{(4)}\right)$$
$$+ m_B|\boldsymbol{q}|^3\left[-2\left(Y^{(2)} + Y^{(3)} - Y^{(4)}\right) + 2\left(2Y^{(1)} - 2Y^{(3)} + Y^{(4)}\right)\right],$$

$$Z^{(2)}_{n=1} = m_B^2|\boldsymbol{q}|\left(-Y^{(1)} + Y^{(3)}\right) + |\boldsymbol{q}|^3\left[Y^{(2)} + Y^{(3)} - Y^{(4)} - 2\left(2Y^{(1)} - 2Y^{(3)} + Y^{(4)}\right)\right]$$
$$+ m_b|\boldsymbol{q}|^2\left[-2\left(-Y^{(1)} + Y^{(3)}\right) + 2\left(2Y^{(1)} - 2Y^{(3)} + Y^{(4)}\right)\right],$$

$$Z^{(3)}_{n=1} = -2m_B|\boldsymbol{q}|\left(-Y^{(1)} + Y^{(3)}\right) + |\boldsymbol{q}|^2\left[-2Y^{(1)} + 2Y^{(3)} + 2\left(-Y^{(1)} + Y^{(3)}\right) - Y^{(4)}\right],$$

$$Z^{(4)}_{n=1} = |\boldsymbol{q}|\left(-Y^{(1)} + Y^{(3)}\right). \tag{2.50}$$

3 Numerical implementation in lattice QCD

In this section, we discuss in detail two different implementations of the method in lattice QCD calculations. First, in subsection 3.1 we present an implementation based on configurations generated within the JLQCD collaboration. Then, in subsection 3.2 we discuss an analogous calculation based on an ensemble generated by the ETM collaboration (ETMC). In both cases, we specify the technical details of the lattice calculations, and discuss the different types of uncertainties affecting the results. Finally, in subsection 3.3, we discuss a few technical aspects related to the extrapolation to the $\sigma \to 0$ limit.

3.1 Lattice implementation with JLQCD configurations

One dataset used to demonstrate the lattice computation of the inclusive semileptonic decay rate is based on the ensemble generated by the JLQCD collaboration. See the supplementary materials of ref. [22] for details of the gauge configurations. It employs Möbius domain-wall fermions for both valence and sea quarks. In the sea, 2+1 flavors of light and strange

quarks are included. The light quark mass corresponds to a pion of mass around 300 MeV; the strange quark mass is slightly heavier than its physical value. The gauge action is the tree-level $O(a^2)$-improved Symanzik at $\beta = 4.35$. The corresponding lattice spacing is $a \simeq 0.055$ fm, corresponding to inverse lattice spacing $1/a = 3.610(9)$ GeV. The lattice volume is $48^3 \times 96$, so that the spatial volume is about $L^3 = (2.6 \, \text{fm})^3$. This ensemble corresponds to "M-$ud3$-sa" of ref. [22]. (See also ref. [23].)

The valence quarks are also described by Möbius domain-wall fermions. The charm quark mass is tuned to its physical value (see ref. [23] for details), while the bottom quark mass is set at $1.25^4 \simeq 2.44$ times the charm quark mass. The spectator quark is a strange quark, so the process corresponds to the inclusive semileptonic decay of a B_s meson, albeit with a light B_s meson mass of $\simeq 3.45$ GeV.

The measurement is carried out on 100 gauge configurations and is replicated four times on each configuration with shifted position of the initial source. The B_s meson is created by a interpolating pseudoscalar operator, which is spatially smeared by a gauge-invariant operator $(1 - (\alpha/N)\Delta)^N$ with a discretized Laplacian Δ and parameters $\alpha = 20$ and $N = 200$. The source points are spread over the source time slice $t_{\text{src}} = 0$ with $\mathbb{Z}_2$ noises to improve statistics. The initial B_s meson is thus projected to zero spatial momentum. The B_s meson on the other end is created by another pseudoscalar operator of the same type placed at the time slice $t_{\text{snk}} = 42$ using a sequential source from the spectator strange quark propagator. The bottom quark propagates from there to a time slice t_2, where the first $b \to c$ current is inserted with momentum $\boldsymbol{q}$ and is fixed at $t_2 = 26$. The charm quark propagator then connects the time slice t_2 to t_1 where the other $b \to c$ current is contracted with momentum insertion $-\boldsymbol{q}$. The charm quark propagator is computed repeatedly for each choice of the current operator and momentum insertion at t_1. We fix the time separation between t_{src} and t_{snk} under an assumption that the ground-state B_s meson state dominates the signal between t_{src} and t_1 or between t_2 and t_{snk}. This separation is at least 16 in units of the lattice spacing, which corresponds to 0.9 fm. The saturation is confirmed in ref. [9].

The matrix element (2.19) is then constructed as in eq. (2.20). For the analysis of this ensemble we applied the Chebyshev polynomial approximation, following ref. [16]. The polynomial order is set to $N = 15$, but the results are unchanged within the statistical error with other choices beyond $N = 12$. The $\sigma \to 0$ limit is taken for each point assuming a polynomial in σ with data points at $\sigma a = 0.02$, 0.05, 0.10 and 0.20. For all the cases, the extrapolation is small compared to the statistical error on the finite values of σ.

The results are shown in figure 1. The left panel is $\bar{X}$ as a function of $\boldsymbol{q}^2$, while the integrand to produce the numerator of $\langle E_\ell \rangle$ is shown in the right panel. The lattice data are obtained at momentum transfer $\boldsymbol{q}$ at $(0,0,0)$, $(0,0,1)$, $(0,1,1)$, $(1,1,1)$, $(0,0,2)$ in units of $2\pi/L$. Data points represent different channels as discussed in section 2.4.

Also shown in figure 1 are dashed curves that represent the contributions from the ground-states, i.e. D_s and D_s^* mesons. They are computed using the form factors obtained by JLQCD for the same quark mass parameters. The necessary formulae and the lattice data are presented in the appendix A.

The lattice data with different momentum insertion $\boldsymbol{q}$ are analyzed together to account for the statistical correlations among them. We then fit $\bar{X}$ in a polynomial of $\boldsymbol{q}^2$ including terms up to $(\boldsymbol{q}^2)^2$.
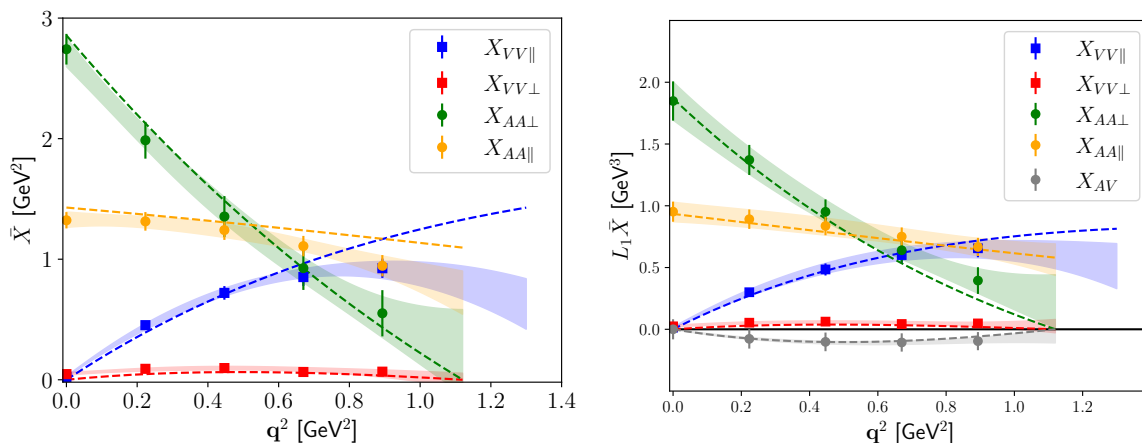
Figure 1. $\bar{X}$ (left panel) and $L_1\bar{X}$ (right panel, corresponding to the numerator of eq. (2.38)) as functions of $\boldsymbol{q}^2$. The results are shown for each channel. X^{AV} is non-vanishing only for $L_1\bar{X}$. The dashed curves are the estimated contributions from the ground states of D_s and D_s^*.
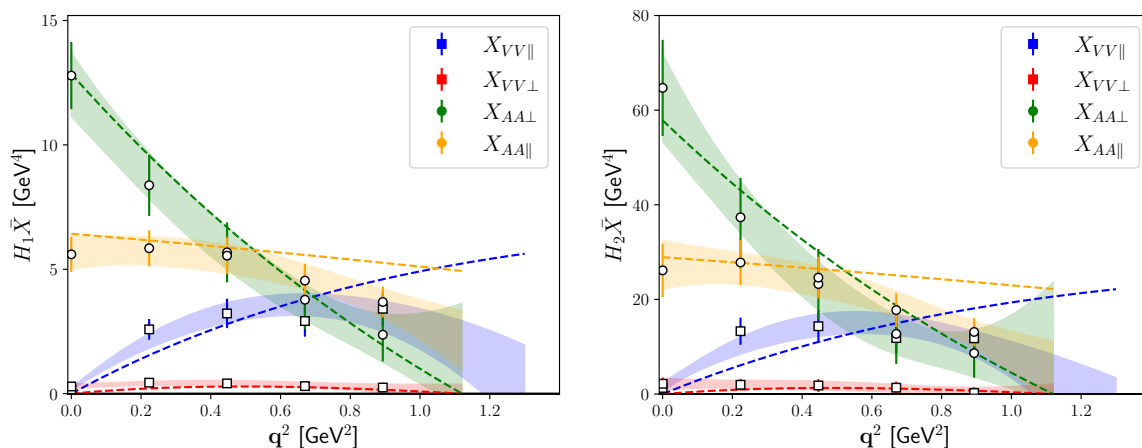

Figure 2. $H_1\bar{X}$ (left panel) and $H_2\bar{X}$ (right panel), which are numerators of eq. (2.37), as a function of $\boldsymbol{q}^2$. The results are shown for each channel. The dashed curves are estimated contributions from the ground state of D and D^*.

We observe that the inclusive results for each channel are consistent with the expected ground-state contributions. This means that the excited-state contributions are small, which is consistent with our expectation from the $B \to D^{**}\ell\nu$ form factors based on heavy-quark effective theory (HQET) [24]. Also phenomenologically, it is plausible because the mass of the initial bottom quark is smaller than its physical value. The heavy quark symmetry predicts that the wave-function overlap is 1 at zero recoil when the initial and final masses are degenerate.

We also calculate the differential moments. The numerators for the hadronic mass moments $\langle M_X^2 \rangle$ and $\langle (M_X^2)^2 \rangle$ are shown in figure 2, while that for $\langle E_\ell \rangle$ is in figure 1 (right panel). The corresponding differential moments, evaluated for each channel at individual momentum $\boldsymbol{q}^2$, are shown in figure 3 and in figure 4.
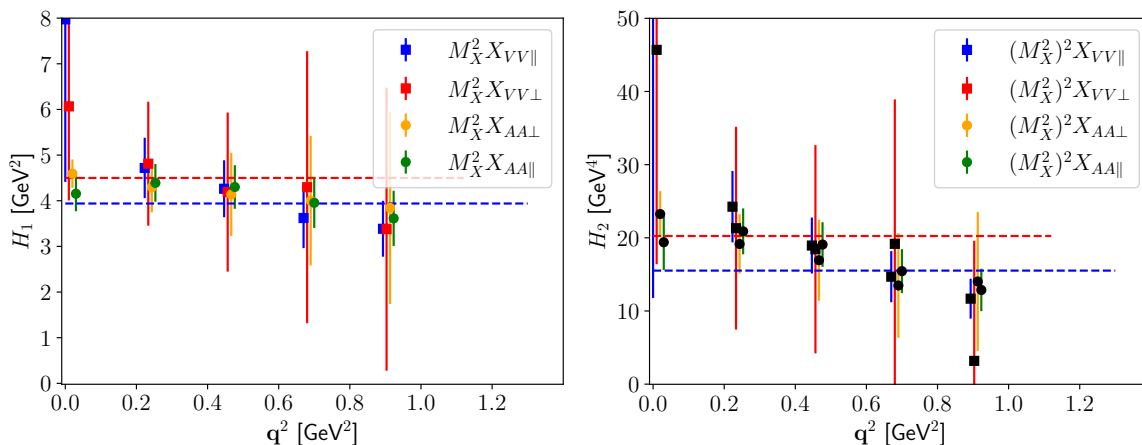
Figure 3. $\langle H_1 \rangle_{\boldsymbol{q}^2}$ (left panel) and $\langle H_2 \rangle_{\boldsymbol{q}^2}$ (right panel) for each channel. The dashed lines are those of the expected contribution from the ground state.
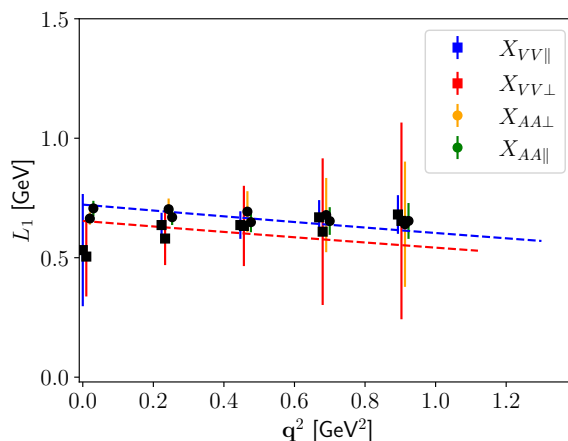
Figure 4. $\langle L_1 \rangle_{\boldsymbol{q}^2}$ for each channel. The dashed lines are those of the expected contribution from the ground state.

3.2 Lattice implementation with ETMC configurations

The ETMC gauge ensemble used in this work is the one named B55.32, generated by ETMC together with other 14 ensembles with $N_f = 2 + 1 + 1$ dynamical quarks in refs. [25, 26] for determining the average up/down, strange and charm quark masses. The Iwasaki action [27] and the Wilson twisted-mass action [28–30] were used for gluons and sea quarks, respectively. Using the mass renormalization constants determined in ref. [31] the physical light, strange, and charm quark masses were found to be $m_{ud}^{\text{phys}}(\overline{\text{MS}}, 2\,\text{GeV}) = 3.70(17)\,\text{MeV}$, $m_s^{\text{phys}}(\overline{\text{MS}}, 2\,\text{GeV}) = 99.6(4.3)\,\text{MeV}$, and $m_c^{\text{phys}}(\overline{\text{MS}}, 2\,\text{GeV}) = 1176(39)\,\text{MeV}$, respectively.

In order to avoid the mixing of K- and D-meson states in the correlation functions a non-unitary setup [32] is used in the valence sectors: the strange and the charm valence quarks are regularised as Osterwalder-Seiler fermions [33], while the up and down valence quarks have the same action as the sea. Working at maximal twist, such a setup guarantees an automatic $\mathcal{O}(a)$-improvement [30, 32].
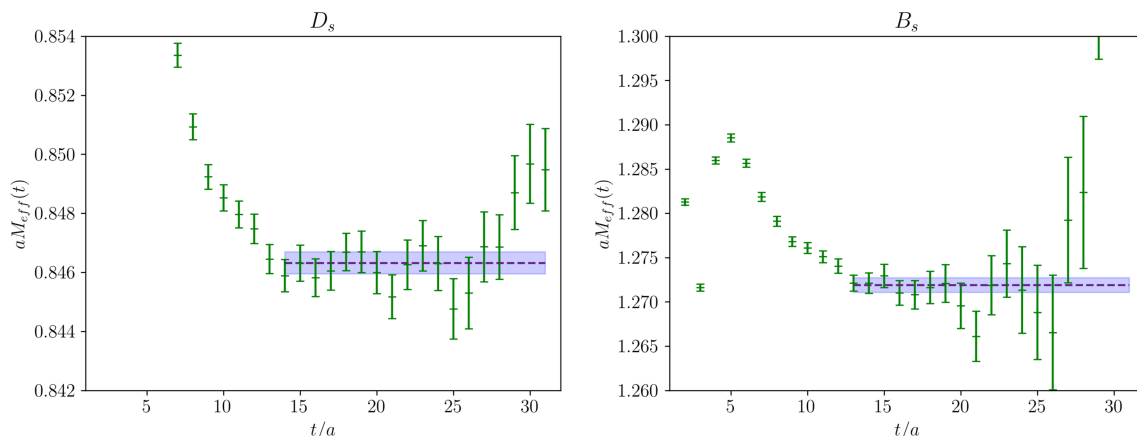
Figure 5. Effective mass $aM_{\text{eff}}(t) \equiv \log\left(C(t)/C(t+a)\right)$ in lattice units for the D_s-meson (left panel) and the B_s-meson (right panel) correlation function (2.21), evaluated using the ETMC gauge ensemble B55.32 for bare quark masses equal to $a\mu_b = 0.50$, $a\mu_c = 0.25$ and $a\mu_s = 0.021$, corresponding to renormalised quark masses $m_b(\overline{\text{MS}}, 2\,\text{GeV}) \simeq 2.4\,\text{GeV}$, $m_c(\overline{\text{MS}}, 2\,\text{GeV}) \simeq 1.2\,\text{GeV}$ and $m_s(\overline{\text{MS}}, 2\,\text{GeV}) \simeq 100\,\text{MeV}$. The values of the Wilson r-parameter of the two valence quarks are opposite, i.e. $r_c = -r_s$ in the D_s meson and $r_b = -r_s$ in the B_s meson.

The ensemble B55.32 has a lattice volume $L^3 \times T = (32^3 \times 64)\,a^4$ with a lattice spacing equal to $a = 0.0815(30)\,\text{fm}$ and a bare light-quark mass equal to $a\mu_\ell = 0.0055$, corresponding to a simulated pion mass $m_\pi = 375(13)$ MeV [31] with $m_\pi L \simeq 5.0$. The number of analyzed gauge configurations, separated by 20 trajectories, is 150. We have carried out our simulations using the values $a\mu_s = 0.021$ and $a\mu_c = 0.25$ for the bare valence strange and charm quark masses, which correspond to renormalised strange and charm quark masses very close to their physical values.

We have calculated the two-point function $C(t)$, defined in eq. (2.21), using the interpolating operator $\bar{b}(x)\gamma_5 s(x)$ with a simulated b-quark mass equal to twice the physical charm mass, i.e. $m_b(\overline{\text{MS}}, 2\,\text{GeV}) \simeq 2.4\,\text{GeV}$, and a physical strange quark. We set opposite Wilson parameters for the two valence quarks in order to guarantee that cutoff effects on the pseudoscalar mass are $O(a^2\mu_f)$ [30, 34, 35]. To improve the statistical precision we have made use of the "one-end trick" stochastic method [36, 37] and employed 10 spatial stochastic sources at a randomly chosen time-slice per gauge configuration. Moreover, in order to suppress contributions of the excited states in the B_s-meson correlation function, we have used Gaussian smeared interpolating quark fields [38] both at the source and at the sink. For the values of the smearing parameters we set $k_G = 4$ and $N_G = 30$. In addition, we apply APE smearing to the gauge links [39] in the interpolating fields with parameters $\alpha_{\text{APE}} = 0.5$ and $N_{\text{APE}} = 20$.

Smearing leads to improved projection onto the lowest-energy eigenstate at smaller Euclidean time separations. As shown by the effective mass $aM_{\text{eff}}(t) \equiv \log\left(C(t)/C(t+a)\right)$ in figure 5, the dominance of the ground-state signal starts around $t/a \simeq 13$ for both the D_s and B_s mesons. By averaging over the plateau regions shown in figure 5 the ground-state masses are respectively found to be $m_{D_s} = 2.05(8)\,\text{GeV}$ and $m_{B_s} = 3.08(11)\,\text{GeV}$.

We have calculated the four-point function $C_{\mu\nu}(t_{\rm snk}, t_2, t_1, t_{\rm src}; \boldsymbol{q})$, given by eq. (2.18), as a function of t_1, the time at which the first weak current is inserted with momentum $\boldsymbol{q}$, for fixed values of t_2, where the second weak current is contracted with momentum insertion $-\boldsymbol{q}$, fixing $t_{\rm src} = 0$ and $t_{\rm snk} = T/2 = 32a$. The momentum $\boldsymbol{q}$ is inserted along one spatial direction, namely $\boldsymbol{q} = (0, 0, q)$ and we have considered eleven values for q ranging from $q = 0$ up to $q = q_{\max} \simeq 0.9 \, {\rm GeV}$. On the lattice these values are injected through the use of twisted boundary conditions (BC's) [40–42] in the spatial directions and anti-periodic BC's in time. The sea dynamical quarks, on the contrary, are simulated with periodic BC's in the spatial directions and anti-periodic ones in time. The twisted BC's for the valence quark fields lift the severe limitations, arising from the use of periodic BC's, on the accessible kinematical regions of momentum-dependent quantities. Furthermore we remark that, as shown in refs. [43, 44], for physical quantities that do not involve final-state interactions (like, e.g., meson masses, decay constants and form factors), the use of different BC's for valence and sea quarks produces only finite-size effects that are exponentially small.

For the $b \to c$ weak current we use the local vector and axial-vector quark currents, $\bar{b}(x)\gamma_\mu c(x)$ and $\bar{b}(x)\gamma_\mu\gamma_5 c(x)$. The value of the Wilson r-parameter for the charm quark is chosen to be opposite to that of the b quark, i.e. $r_c = -r_b$, and therefore in our maximally twisted setup the vector and axial-vector currents renormalise respectively with the axial and vector renormalization constants, Z_A and Z_V, determined in ref. [31].

We extract the matrix elements $M_{\mu\nu}(t_2 - t_1; \boldsymbol{q})$ using eq. (2.20). In order to calculate $\bar{X}(\boldsymbol{q}^2)$, as defined in eq. (2.17), we apply the smearing kernel $\Theta^{(l)}(\omega_{\max} - \omega)$ to the quantities $Z^{(l)}(\omega, \boldsymbol{q}^2)$. These in turn are defined in terms of the quantities $Y^{(a)}(\omega, \boldsymbol{q})$ in eq. (2.15). To this end we start from the linear combinations of the correlators $M_{\mu\nu}(t; \boldsymbol{q})$ with the kinematical coefficients of eqs. (2.9). We call these objects

$$
\begin{aligned}
Y^{(a)}(t; \boldsymbol{q}^2) &= \int_0^\infty d\omega \, Y^{(a)}(\omega, \boldsymbol{q}^2) \, e^{-\omega t}, && a = 1, \cdots, 5, \\
Z^{(l)}(t; \boldsymbol{q}^2) &= \int_0^\infty d\omega \, Z^{(l)}(\omega, \boldsymbol{q}^2) \, e^{-\omega t}, && l = 0, 1, 2.
\end{aligned} \tag{3.1}
$$

To show the quality of the numerical data, in figure 6 we plot the correlators $Y^{(a)}(t; \boldsymbol{q}^2)$ corresponding to $|\boldsymbol{q}| \simeq 0.5 \, {\rm GeV}$. Notice that the correlators $Z^{(l)}(t; \boldsymbol{q}^2)$ are linear combinations of the $Y^{(a)}(t; \boldsymbol{q}^2)$'s, see eq. (2.15). Similar results are obtained for the other momenta considered in this work.

The central values for all the physical quantities extracted from the $Y^{(a)}(t, \boldsymbol{q})$ correlators have been extracted by setting $t_2 = 22a$ in eq. (2.20) and by using the data up to $t = 18a$, which corresponds to $t_1 - t_{\rm src} = 4a$. To check the approach to the $t_{\rm src} \to -\infty$ and $t_{\rm snk} \to \infty$ limits we have repeated the analysis by setting $t_2 = \{18a, 20a, 22a, 26a, 28a\}$ and by varying the maximum value of t used to reconstruct the smearing kernels. Figure 7 shows the comparison of the correlator $Y^{(1)}(t, \boldsymbol{q})$ at $|\boldsymbol{q}| \simeq 0.5 \, {\rm GeV}$ for different values of $t = t_2 - t_1$ and t_2. In the following analysis, we chose the value $(t_{\rm snk} - t_2) = 10a$, corresponding to $t_2 = 22a$. Similar results are obtained for the other correlators ($Y^{(2)}$, $Y^{(3)}$, $Y^{(4)}$ and $Y^{(5)}$), and, in all cases, we observe that the onset of the $t_{\rm snk} \to \infty$ limit is reached within the uncertainties already for $t_{\rm snk} - t_2 = 4a$.

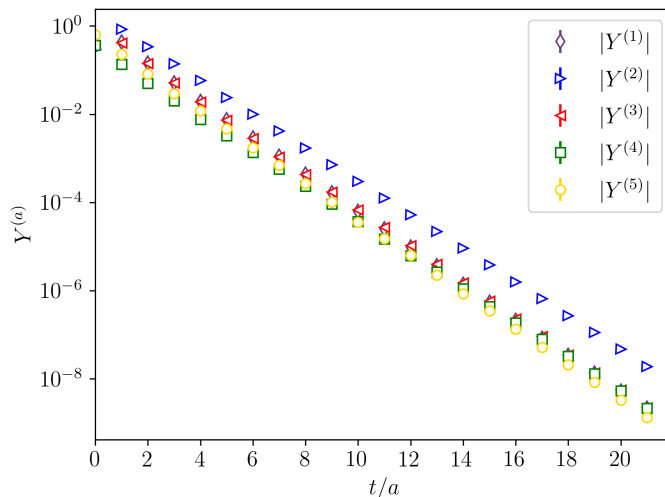**Figure 6**. Time dependence of the correlators $Y^{(a)}(t; \boldsymbol{q})$ for $|\boldsymbol{q}| \simeq 0.5\,\text{GeV}$ calculated on the ETMC ensemble B55.32. The error bars are smaller than the point markers on this scale and a similar quality of the numerical signal is observed for the other momentum values considered in this work.
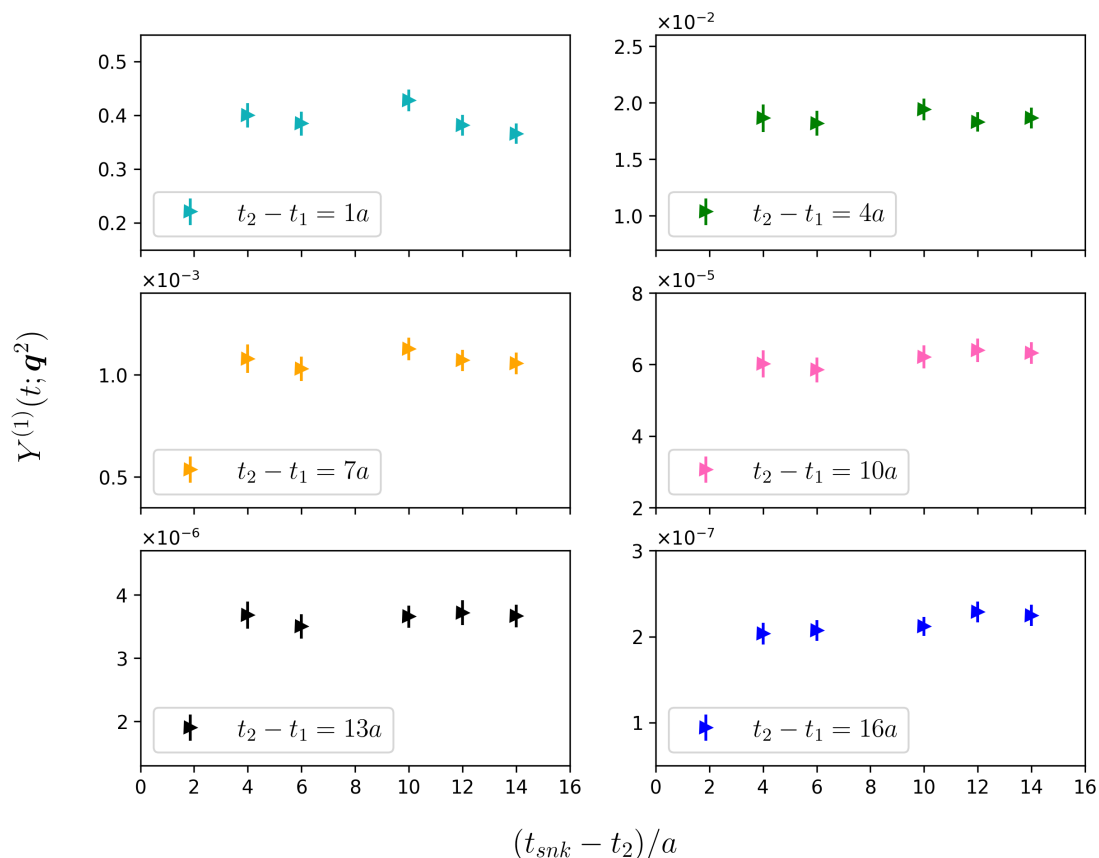

Figure 7. Correlator $Y^{(1)}(t, \boldsymbol{q}^2)$ at various time separations $t_2 - t_1$ for $|\boldsymbol{q}| \simeq 0.5\,\text{GeV}$. The points in each subplot are obtained for different values of t_2, with the x-axis showing the distance between t_{snk} and the time t_2 at which the current is inserted.
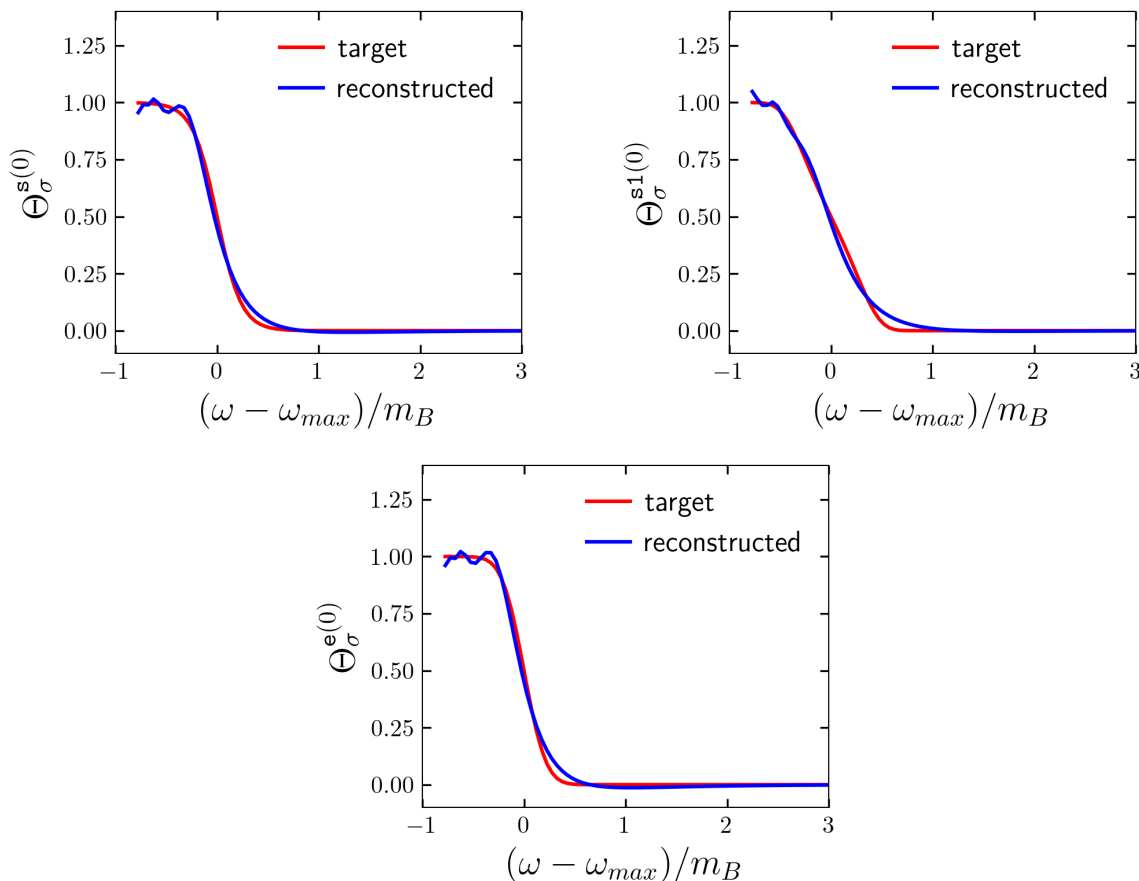
Figure 8. Reconstruction of the kernels $\Theta_\sigma^{(0)}(\omega_{\max} - \omega)$ defined with the three smearing types $\mathtt{s}$, $\mathtt{s1}$ and $\mathtt{e}$, see eq. (2.36), at $\lambda = \lambda_\star$. The data correspond to $|\boldsymbol{q}| \simeq 0.7\,\mathrm{GeV}$ and $\sigma = 0.12 m_{B_s}$, the smallest value of the smearing parameter that we used.

We now turn to the discussion of the systematics associated with the approximation of the kernels of eq. (2.36) by using the method of ref. [13]. This is an important issue because, on the one hand, the reconstruction of a given kernel can never be exact with a finite number of time-slices and in the presence of errors. On the other hand, one can (and must) quantify the systematic error associated with an approximate reconstruction.

In order to illustrate this point we consider the quantity $Z_\sigma^{(0)}(\boldsymbol{q}^2)$ (see eq. (2.17)) for three smooth approximations of the θ-function given in eq. (2.36). The kernels are approximated as described in section 2, see in particular eq. (2.31), with $\tau_{\max} = 18$. The quantity $Z_\sigma^{(0)}(\boldsymbol{q}^2)$ is then obtained by applying the coefficients g_τ^λ that represent the approximated kernel at a fixed value of λ to the correlator $Z^{(0)}(t; \boldsymbol{q}^2)$. Figure 8 shows the comparison of the reconstructed kernels with the target ones for $|\boldsymbol{q}| \simeq 0.7\,\mathrm{GeV}$ and $\sigma = 0.12 m_{B_s}$ at the values $\lambda = \lambda_\star$ determined independently for each kernel. The values of $\lambda_\star$ are marked with red points in figure 9, where we show the dependence of $Z_\sigma^{(0)}(\boldsymbol{q}^2)$ on the normalised L_2-norm $A[g^\lambda]/A[0]$. As explained in section 2, for smaller values of λ one obtains a more accurate reconstruction of the kernels and thus smaller $A[g^\lambda]/A[0]$ values. There is no significant difference on the final results for $Z_\sigma^{(0)}(\boldsymbol{q})$ by decreasing λ with respect to $\lambda_\star$.
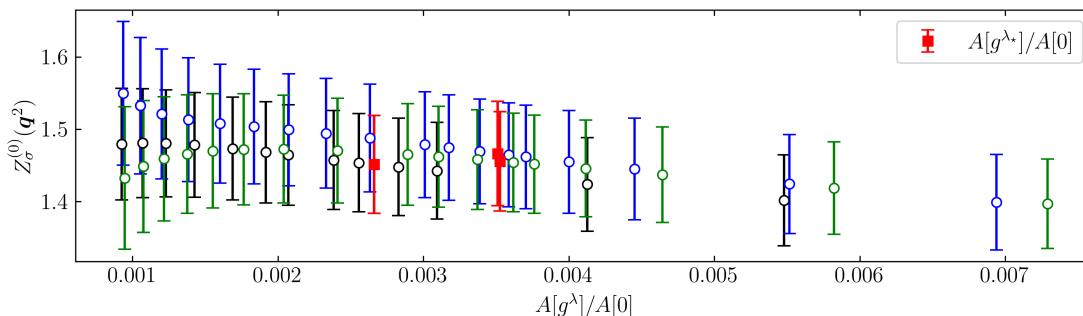
Figure 9. Integral $\bar{Z}_\sigma^{(0)}(\boldsymbol{q})$ of the hadronic correlator with three kernels, plotted as a function of $A[g^\lambda]/A[0]$. No significant difference is observed within the statistical errors for values $A[g^\lambda]/A[0]$ smaller than $A[g^{\lambda_\star}]/A[0]$.

By implementing this strategy, proposed in ref. [15], we have checked that the estimated errors on the different quantities that enter our determinations of the physical observables discussed below properly take into account the systematics associated with the kernel approximation.

In figure 10 we show our results for the total decay rate, with the different points corresponding to different input parameters used in the analysis, as described in the figure's caption. The plot shows clearly that all results are compatible with each other. In order to take into account all the results showed in the figure, we use eq. (28) of ref. [31] to get an estimate of the central value and its standard deviation, corresponding to the filled red dots in the plot, and we quote that value as our final result for the total decay rate. This procedure is repeated for all other observables considered in this work.

The final ETMC results for all the physical observables, divided into four different channels, are shown together with the OPE results in figures 14, 17, 19, and 20.

3.3 Extrapolation to $\sigma = 0$

The ETMC data are produced at several values of the smearing parameter σ and, for each of the target kernels $\Theta^{(l)}(x)$ with three different smeared versions of the θ-function in eq. (2.36). These are used in a combined $\sigma \to 0$ extrapolation for each contribution to the differential decay rate and to the leptonic and hadronic moments.

Before presenting the results of the $\sigma \to 0$ extrapolation an important remark is needed. As discussed in section 2 the limits of zero smearing radius and of infinite volume do *not* commute. Because of the quantized energy spectrum in a finite volume, the $\sigma \to 0$ extrapolation must be performed only after the infinite-volume limit. Under the reasonable assumption that smeared QCD spectral densities are affected by exponentially suppressed finite-volume effects, and given the exploratory nature of the present work, we shall assume below that finite-volume effects are negligible with respect to our statistical uncertainties. This assumption can only be verified with simulations on larger volumes, a task that we leave for future work on the subject. Taking this issue into account, the $\sigma \to 0$ extrapolation discussed below has to be considered as a feasibility study that, as we work at unphysical meson masses and fixed cutoff, we consider interesting and promising.
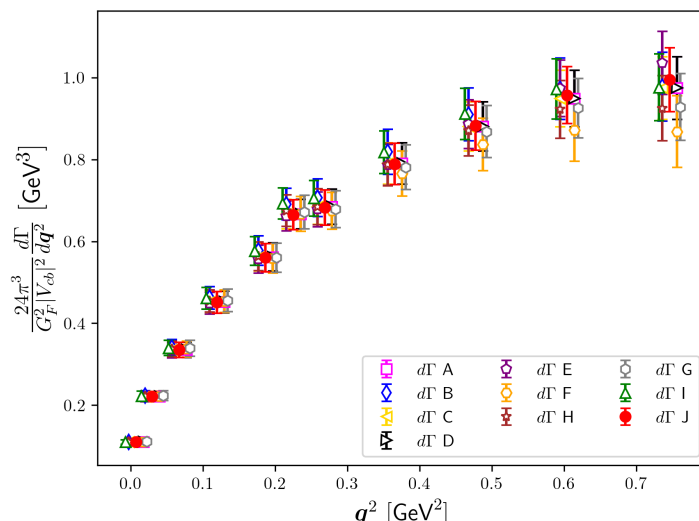
Figure 10. Results for $\frac{24\pi^3}{G_F^2|V_{cb}|^2}\frac{d\Gamma}{dq^2}$, obtained changing the parameters given as input to our analysis. The default values are: $A_{tr} = 1 \times 10^{-3}$, $\tau_{\max} = 18$, extrapolations to $\sigma = 0$ using 5 values of σ. The letters in the legend stand for: A) All parameter equal to default, the final result is given by extrapolating to $\sigma = 0$ the single components $X^{(i)}$ and then summing the extrapolations together. B) The same as case (A) but with extrapolations done employing all 10 values of σ, as quoted in the caption of figure 11. C) A threshold changed to $A_{tr} = 1 \times 10^{-2}$. D) A threshold changed to $A_{tr} = 5 \times 10^{-3}$. E) All parameters equal to default, final result given by summing all the single contributions $X^{(i)}$ together and then extrapolation the sum to $\sigma = 0$. F) $\tau_{\max}$ changed to $\tau_{\max} = 15$. G) $\tau_{\max}$ changed to $\tau_{\max} = 16$. H) $\tau_{\max}$ changed to $\tau_{\max} = 17$. I) Same as default, analysis performed using the bootstrap method. J) Final results obtained considering all previous results listed here. Central value and standard deviation are calculated using the average procedure given by eq. (28) of ref. [31]. It is important to note that the analysis of all the cases listed above is performed taking the result corresponding to $\lambda = \lambda_\star$ as discussed in subsection 3.2, the only exception being when we change the A_{tr} parameter. In these two cases we take the results corresponding to values of $A[g^\lambda]/A[0]$ smaller than A_{tr}.

In figure 11 we show the $\sigma \to 0$ extrapolations of the three contributions $Z_\sigma^{(l)}(\boldsymbol{q}^2)$ to the differential decay rate for $|\boldsymbol{q}| \simeq 0.5\,\text{GeV}$ (plots on the left) and $|\boldsymbol{q}| \simeq 0.7\,\text{GeV}$ (plots on the right). The reconstruction of the kernels $\Theta_\sigma^{(0)}(\omega_{\max} - \omega)$ is more challenging from the numerical point of view with respect to the case of the kernels $\Theta_\sigma^{(l)}(\omega_{\max} - \omega)$ with $l = 1, 2$. In all cases studied in this work we have obtained results at 10 different values of σ that, for the kernel $\Theta_\sigma^{(0)}(\omega_{\max} - \omega)$ span the region $\sigma \in [0.12 m_{B_s}, 0.3 m_{B_s}]$ while for the other kernels we have $\sigma \in [0.03 m_{B_s}, 0.16 m_{B_s}]$. For all the values of $\boldsymbol{q}^2$ we have included the five smallest σ values into a combined linear extrapolation to obtain the central values and statistical errors that we quote for our results at $\sigma = 0$. As evident from the plots in figure 11 there is a reassuring convergence of the results corresponding to the different kernels for small values of σ. The five points included in the fit are always in the linear regime and the $\chi^2/\text{d.o.f.}$ for all the combined $\sigma \to 0$ linear extrapolations performed in this work never exceed 1.
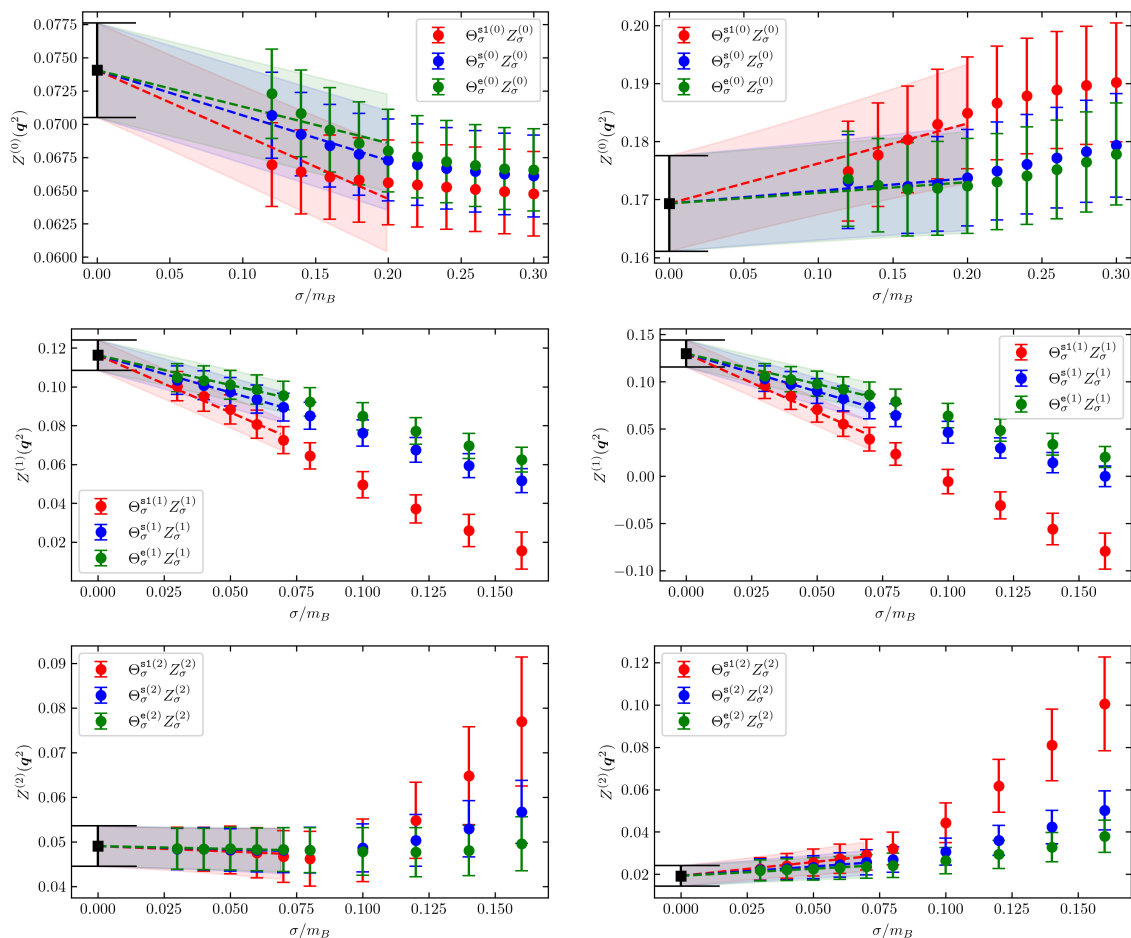
Figure 11. Combined $\sigma \to 0$ extrapolations of three contributions $Z^{(l)}(\boldsymbol{q}^2)$ to the differential decay rate, see eq. (2.17). The plots on the left correspond to $|\boldsymbol{q}| \simeq 0.5\,\mathrm{GeV}$ while those on the right to $|\boldsymbol{q}| \simeq 0.7\,\mathrm{GeV}$. The reconstruction of the kernels $\Theta_\sigma^{(0)}(\omega_{\max} - \omega)$ is more difficult from the numerical point of view w.r.t. the case of the kernels $\Theta_\sigma^{(l)}(\omega_{\max} - \omega)$ with $l = 1, 2$. In all cases we have obtained results at 10 different values of σ that, in the case of $\Theta_\sigma^{(0)}(\omega_{\max} - \omega)$ span the region $\sigma \in [0.12 m_{B_s}, 0.3 m_{B_s}]$ while in the other case we have $\sigma \in [0.03 m_{B_s}, 0.16 m_{B_s}]$. In all cases we include the five smallest values of σ into a combined linear extrapolation to quote our results at $\sigma = 0$.

The systematics associated with the $\sigma \to 0$ extrapolations has been quantified (see also the caption of figure 10) by performing unconstrained linear extrapolations of the five points at the smaller values of σ and combined quadratic extrapolations of all points, i.e. with ten values of σ. This procedure is illustrated in figure 12 where we show, for the same set of data appearing in the top-left panel of figure 11, the unconstrained linear extrapolations and the result of the combined quadratic extrapolation (violet point). As can be seen in this plot, the results of the three different unconstrained extrapolations are compatible within the quoted errors and also compatible with our central value result (black point). Following the procedure explained in the caption of figure 10, i.e. estimating the systematics associated with the extrapolation by adding in quadrature the statistical error of the black point and
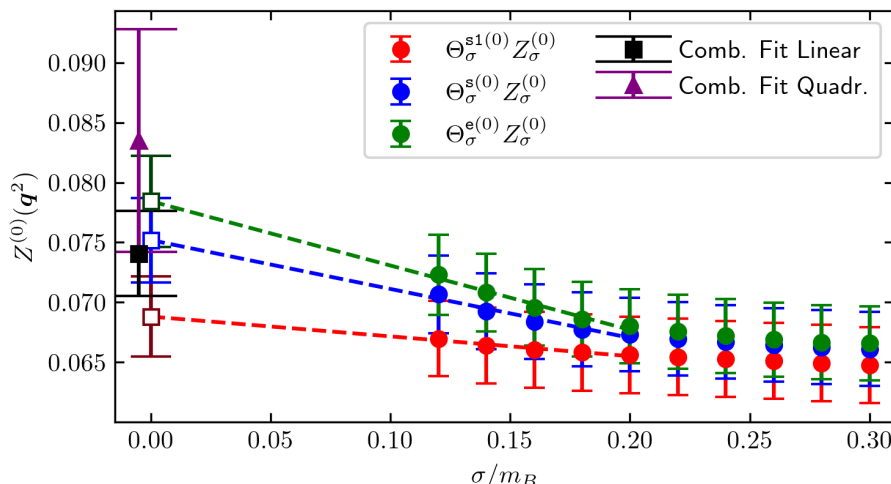
Figure 12. Systematics associated with the $\sigma \to 0$ extrapolation of $Z^{(0)}(\boldsymbol{q}^2)$ at $|\boldsymbol{q}| \simeq 0.5\,\mathrm{GeV}$, the same set of data shown in the top-left panel of figure 11. The unconstrained linear extrapolations of the different sets of data, corresponding to the three different smearing kernels, are shown together with the results of the combined linear extrapolation of the five points at the smaller values of σ (black point) and of the combined quadratic extrapolation including all ten values of σ (violet point). The black and violet points have been slightly displaced on the horizontal axis to help the eye.

the difference between the central values of the black and violet points, largely takes into account the spread of the results coming from the different extrapolations, including the unconstrained ones. The same procedure has been repeated for all the sets of data analyzed in this work and similar plots can be shown in all cases.

In figure 13 we show the $\sigma \to 0$ extrapolations of the four different terms that enter the calculation of the leptonic moment $L_1(\boldsymbol{q}^2)$.

4 Operator-product expansion and comparison with lattice results

As inclusive semileptonic B decays are described by an OPE, observables that are sufficiently inclusive admit a double expansion in α_s and in inverse powers of m_b [10, 11, 19, 20, 45], or more precisely of the energy release, which is of the order of $m_b - m_c$. Schematically, for an observable M we have

$$M = M^{(0)} + M^{(1)}a_s + M^{(2)}a_s^2 + \left(M_\pi^{(0)} + M_\pi^{(1)}a_s\right)\frac{\mu_\pi^2}{m_b^2}$$

$$+ \left(M_G^{(0)} + M_G^{(1)}a_s\right)\frac{\mu_G^2}{m_b^2} + M_D^{(0)}\frac{\rho_D^3}{m_b^3} + M_{LS}^{(0)}\frac{\rho_{LS}^3}{m_b^3} + \dots \tag{4.1}$$

where $a_s = \alpha_s(\mu)/\pi$ is the QCD coupling evaluated at a scale $\mu \sim m_b$ and the ellipsis represents higher-order terms in a_s and in $1/m_b$. The parameters μ_π^2, μ_G^2, ρ_D^3, ρ_{LS}^3 are expectation values of dimension-5 and dimension-6 local operators in the physical B meson. For instance,

$$\mu_\pi^2(\mu_k) = \frac{1}{2M_B}\langle B|\bar{b}_v\,\vec{\pi}^2\,b_v|B\rangle_{\mu_k}, \qquad \mu_G^2(\mu_k) = \frac{1}{2M_B}\langle B|\bar{b}_v\frac{i}{2}\sigma_{\mu\nu}G^{\mu\nu}b_v|B\rangle_{\mu_k} \tag{4.2}$$
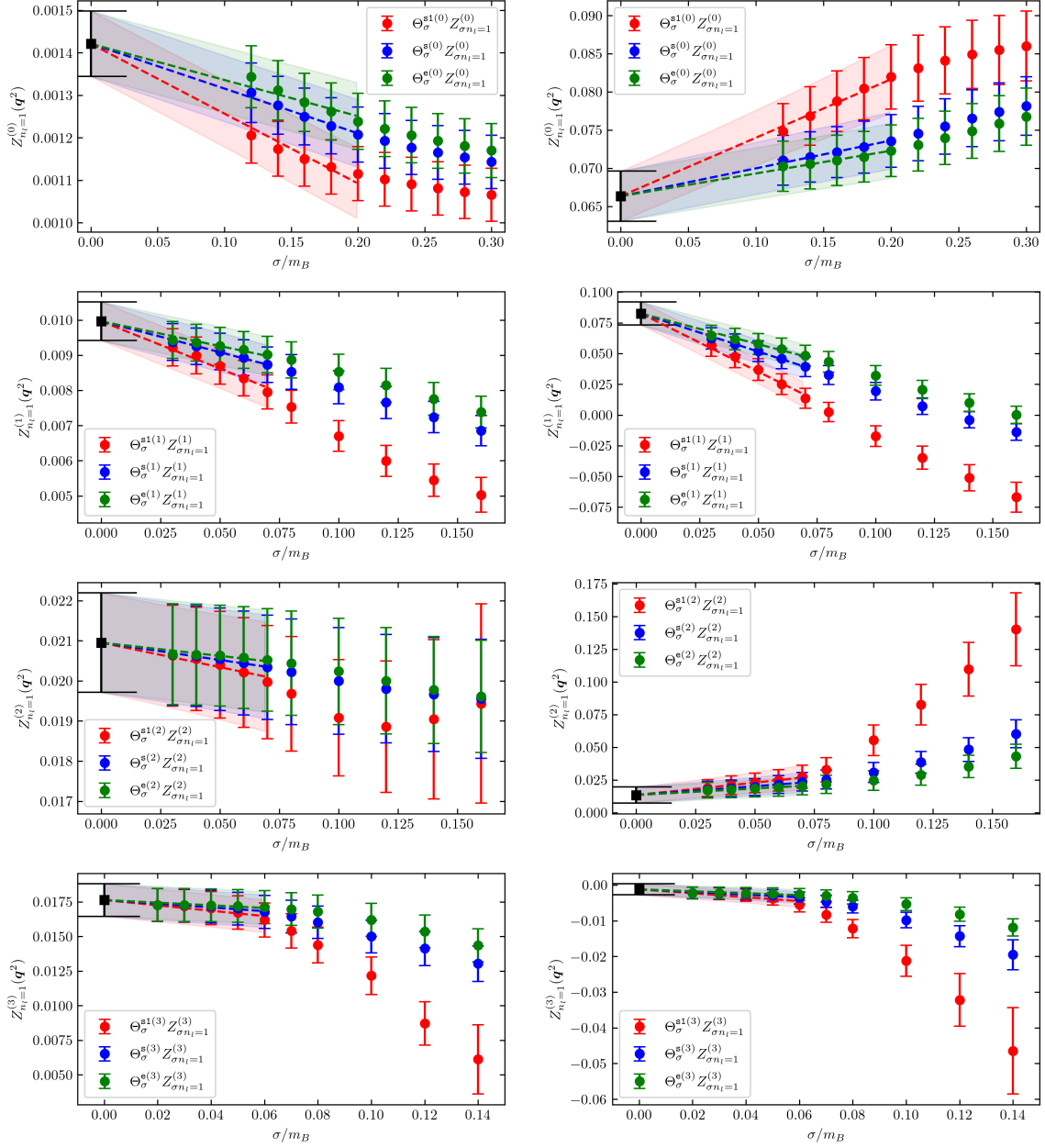
Figure 13. Combined $\sigma \to 0$ extrapolations of the four contributions $Z_{n_\ell=1}^{(l)}(\boldsymbol{q}^2)$ to the first leptonic moment, see eq. (2.45). The plots on the left correspond to $|\boldsymbol{q}| \simeq 0.26\,\mathrm{GeV}$ while those on the right to $|\boldsymbol{q}| \simeq 0.78\,\mathrm{GeV}$.

where $\vec{\pi} = -i\vec{D}$, while D^μ is the covariant derivative, $b_v(x) = e^{-im_b v \cdot x} b(x)$ is the b field deprived of its high-frequency modes, and $G^{\mu\nu}$ is the gluon-field tensor. In the so-called kinetic scheme [46–48], the Wilsonian cutoff $\mu_k \sim 1\,\mathrm{GeV}$ is introduced to factorise long- and short-distance contributions. Indeed, the OPE disentangles the physics associated with *soft* scales of order Λ_{QCD} (described by the above parameters) from that associated with *hard* scales $\sim m_b$, which determine the Wilson coefficients M_i that admit an expansion in α_s. Quite importantly, the power corrections start at $O(\Lambda_{\mathrm{QCD}}^2/m_b^2)$ and are therefore comparatively suppressed. The kinetic scheme provides a short-distance, renormalon-free definition of m_b and of the OPE parameters by introducing the cutoff μ_k to factor out the infrared contributions from the perturbative calculation.

The smearing provided by the phase-space integration, discussed in section 2, is in general sufficient to guarantee the convergence of the OPE for the quantities introduced in eqs. (2.6) and (2.37)–(2.40), which can then be expressed in the form (4.1). The OPE calculation proceeds therefore as in refs. [10, 11, 49]. There are however two specific points related to the kinematics chosen in the lattice calculation that need to be mentioned. First, while the hard scale that governs the OPE is generally $m_b - m_c$, there are regions of the phase space, e.g. at small recoil $|\boldsymbol{q}| \sim 0$, where it is rather m_c, possibly implying a slower convergence of the expansion. Second, near the maximum value of $\boldsymbol{q}^2$ the smearing interval in ω closes up and one cannot expect the OPE to provide reliable results.

4.1 Details of the OPE calculation and related uncertainties

From a technical point of view, the OPE provides a double expansion like the one in eq. (4.1) for the hadronic tensor $W^{\mu\nu}$ defined in eq. (2.4) that can be used to compute the total rate, the moments, and any sufficiently inclusive quantity. The coefficients of the expansion involve the Dirac delta $\delta(r^2 - m_c^2)$ and its derivatives, which upon integration over the *quark* (partonic) phase space lead to results valid for sufficiently inclusive observables. It is customary to use the decomposition of $W^{\mu\nu}$ into Lorentz-invariant form factors as in eq. (2.7) and to identify the four-velocities of the B meson and of the b quark, $v = p/m_B = p_b/m_b$. In this section we will use eq. (2.7) replacing m_B with the b quark mass m_b and employing a hat for quantities that are normalised to m_b. In the case of massless leptons considered in this work, the form factors $W_{4,5}$ do not contribute to the decay amplitude.

The lowest order of the expansion for the relevant W_i and the $1/m_b^2$ corrections can be found in refs. [10, 11], while analytic expressions for the $O(\alpha_s)$ terms are given in refs. [50, 51]. The $O(1/m_b^3)$ corrections have been first computed in ref. [52]. Higher power corrections have been investigated in ref. [53], but involve a large number of new and poorly known parameters. They appear to be sufficiently suppressed at the physical m_b [54]; we will not consider them but they represent an important source of theoretical uncertainty in our low m_b setup. The $O(\alpha_s/m_b^2)$ corrections to the W_i are also known [51, 55], while for the total rate we also have $O(\alpha_s/m_b^3)$ corrections [56, 57]. Numerical results for the $O(\alpha_s^2 \beta_0)$ contributions are also available [50], while the complete $O(\alpha_s^2)$ are available only for the total rate and for a few moments [58–61]. Finally, the $O(\alpha_s^3)$ correction to the total rate has been recently computed in ref. [62].

While these corrections have generally been computed in the $V - A$ case realised in the SM, the decomposition in VV, AA and $AV = VA$ components is potentially useful in our case, and has been made manifest for the $O(1/m_b^{2,3})$ and $O(\alpha_s)$ corrections, see refs. [11, 63, 64]. In the calculation of the $\boldsymbol{q}^2$ spectrum and of the differential moments we will therefore consider only power corrections up to and including $O(1/m_b^3)$ and the $O(\alpha_s)$ perturbative corrections. However in the calculation of the total width and of the total moments we will restrict to the SM case and will employ all the known corrections.

Following section 2, we take the three-momentum $\boldsymbol{q}$ to point along the k direction and the i and j directions to be perpendicular to that. The components of the hadronic tensor along these directions are given by

$$
\begin{aligned}
W^{00} &= -W_1 + W_2 + \hat{q}_0^2 W_4 + 2\hat{q}_0 W_5 \,, \\
W^{ii} &= W^{jj} = W_1 \,, \\
W^{kk} &= W_1 + \hat{q}_k^2 W_4 \,, \\
W^{0i} &= W^{i0} = W^{ik} = W^{ki} = W^{jk} = W^{kj} = 0 \,, \\
W^{0k} &= W^{k0} = \hat{q}_0 \hat{q}_k W_4 + \hat{q}_k W_5 \,, \\
W^{ij} &= -W^{ji} = -i\epsilon^{ij0k} \hat{q}_k W_3 \,.
\end{aligned}
$$

In the OPE the decay occurs at the quark level: $p_b = p' + p_\ell + p_\nu$, where p_b and p' are the momenta of the initial b quark and of a final hadronic state made of a c quark and $n \geq 0$ perturbative gluons. At the leading order in α_s and in $1/m_b$, this is a free-quark decay into an on-shell c quark, which implies that the W_i are proportional to $\delta(p'^2 - m_c^2) = \delta(\hat{u})/m_b^2$, where $\hat{u} = (p'^2 - m_c^2)/m_b^2$. We can rewrite this δ distribution in terms of the energy of the final c quark,

$$
\delta(\hat{u}) = \frac{1}{2\sqrt{\hat{\boldsymbol{q}}^2 + \rho}} \left[\delta\left(\hat{\chi} - \sqrt{\hat{\boldsymbol{q}}^2 + \rho}\right) + \delta\left(\hat{\chi} + \sqrt{\hat{\boldsymbol{q}}^2 + \rho}\right) \right] \,, \tag{4.3}
$$

where $\rho = m_c^2/m_b^2$ and $\hat{\chi}$ is the parton-level energy of the final hadronic state in units of m_b, which is related to the total hadronic energy ω by $\omega = m_b \hat{\chi} + \Lambda$, with $\Lambda = M_B - m_b$. Similarly, the invariant hadronic mass M_X^2 is related to the partonic variables by

$$
M_X^2 = (p_B - q)^2 = (p_b + \Lambda v - q)^2 = m_b^2 \hat{u} + 2m_b \Lambda \hat{\chi} + \Lambda^2 + m_c^2 \,.
$$

Only the first term of eq. (4.3) contributes to the physical process of interest and can be readily integrated over $\hat{\chi}$. At $O(1/m_b^{2,3})$ one has to deal with $\delta'(\hat{u})$, $\delta''(\hat{u})$ and $\delta'''(\hat{u})$ that upon integration subject to kinematic constraints lead to new singularities. A typical case is provided by the interplay between the δ' and the requirement that $q^2 \geq 0$:

$$
\begin{aligned}
\int f(\hat{u})\, \theta(q^2)\, \delta'(\hat{u})\, d\hat{u} &= \int f(\hat{u})\, \theta\left(1 + \rho + \hat{u} - 2\sqrt{\rho + \hat{\boldsymbol{q}}^2 + \hat{u}}\right) \delta'(\hat{u})\, d\hat{u} \\
&= -f'(0)\theta\left(1 + \rho - 2\sqrt{\rho + \hat{\boldsymbol{q}}^2}\right) + f(0)\frac{1+\rho}{2}\delta(\hat{\boldsymbol{q}}^2 - \hat{\boldsymbol{q}}_{\max}^2)\,. \tag{4.4}
\end{aligned}
$$

The singularity at the partonic endpoint of the $\boldsymbol{q}^2$ spectrum, $\hat{\boldsymbol{q}}_{\max}^2 = (1 - \rho)^2/4$, appears because one reaches the maximum energy exactly on the mass-shell of the charm quark.

We apply exactly the same setup to compare with both JLQCD and ETMC data, adjusting only the heavy-quark masses to the two cases. The unphysically light b quark mass and the OPE parameters are expressed in the kinetic scheme with $\mu_k = 1\,\text{GeV}$, while the c quark mass is expressed in the $\overline{\text{MS}}$ scheme at $2\,\text{GeV}$. In the case of the JLQCD data we employ $m_b(1\,\text{GeV}) = 2.70(4)\,\text{GeV}$, obtained from matching the observed m_{B_s} with the results of refs. [65, 66], and $\overline{m_c}^{(4)}(2\,\text{GeV}) = 1.10(2)\,\text{GeV}$. In the case of the ETMC data we employ $\overline{m_c}^{(4)}(2\,\text{GeV}) = 1.186(41)\,\text{GeV}$ and $\overline{m_b}^{(4)}(2\,\text{GeV}) = 2.372(82)\,\text{GeV}$ (with 100% correlated uncertainties), and translate the latter into the kinetic scheme using the three-loop conversion formula [48] implemented in version 3.1 of *RunDec* [67] obtaining $m_b(1\,\text{GeV}) = 2.39(8)\,\text{GeV}$. The strong coupling employed in the conversion and elsewhere is $\alpha_s^{(4)}(2\,\text{GeV}) = 0.301$.

For the OPE parameters that appear in eq. (4.1) we start from the results of the most recent fit to the semileptonic moments [68], which refer to the physical B meson, with a much heavier b quark and without a strange spectator. The difference induced in these parameters by the strange spectator at the physical point has been investigated in refs. [66, 69, 70], where it was found that spectroscopic and lattice data approximately suggest a 20% upward shift in μ_π^2 and μ_G^2, while heavy-quark sum rules hint at a similar or even stronger SU(3) flavour-symmetry breaking in ρ_D^3. The dependence on the mass of the heavy quark, on the other hand, can be analysed by observing that μ_π^2 and μ_G^2 satisfy a heavy-quark expansion

$$\mu_\pi^2 = \mu_\pi^2|_\infty - \frac{\rho_{\pi\pi}^3 + \frac{1}{2}\rho_{\pi G}^3}{m_b} + \dots, \qquad \mu_G^2 = \mu_G^2|_\infty + \frac{\rho_S^3 + \rho_A^3 + \frac{1}{2}\rho_{\pi G}^3}{m_b} + \dots \qquad (4.5)$$

where $\rho_{\pi\pi}^3$, $\rho_{\pi G}^3$, ρ_S^3, ρ_A^3 are expectation values of non-local operators, of which little is known, see ref. [65]. If they were of the same order of magnitude of ρ_D^3 and ρ_{LS}^3, i.e. about 0.1–0.2 GeV^3, they could shift μ_π^2 and μ_G^2 by 0.02–0.1 GeV in going from the physical value of m_b to $m_b \sim 2.5\,\text{GeV}$, which amounts to a 5–25% shift. We show the inputs of our calculation in table 1. While the heavy-quark masses are slightly different between the two setups, we adopt the same expectation values in both cases. Their central values take into account the shift related to the strange spectator, while the uncertainties follow from the uncertainty of the fit of ref. [68], the SU(3) symmetry breaking, and the lower b mass.

Beside the parametric uncertainty of the inputs, our results are subject to an uncertainty due to the truncation of the expansion in eq. (4.1) and to possible violations of quark-hadron duality. We estimate the former by varying the OPE parameters, the heavy-quark masses, and α_s in an uncorrelated way and adding the relative uncertainties in quadrature. In particular, we shift $m_{b,c}$ by 6 MeV, $\mu_{\pi,G}^2$ by 15%, and $\rho_{D,LS}^3$ by 25%. These corrections should mimic the effect of higher-power corrections. Since in the case of the q^2 spectrum and differential moments we restrict ourselves to $O(\alpha_s)$ corrections, we include the relative uncertainty in the same way, shifting α_s by 0.15, which corresponds to a 50% uncertainty. In the case of the total width and total moments, higher-order perturbative corrections are known and the perturbative uncertainty can be reduced, as discussed below.

$m_b^{\rm kin}$ (JLQCD)	2.70 ± 0.04
$\overline{m}_c(2\,{\rm GeV})$ (JLQCD)	1.10 ± 0.02
$m_b^{\rm kin}$ (ETMC)	2.39 ± 0.08
$\overline{m}_c(2\,{\rm GeV})$ (ETMC)	1.19 ± 0.04
μ_π^2	0.57 ± 0.15
ρ_D^3	0.22 ± 0.06
$\mu_G^2(m_b)$	0.37 ± 0.10
ρ_{LS}^3	-0.13 ± 0.10
$\alpha_s^{(4)}(2\,{\rm GeV})$	0.301 ± 0.006

Table 1. Inputs for our OPE calculation. All parameters are in GeV to the appropriate power and all, except m_c, in the kinetic scheme at $\mu = 1\,{\rm GeV}$. The heavy-quark masses for the ETMC setup are 100% correlated. As a remnant of the semileptonic fit, we include a 50% correlation between μ_π^2 and ρ_D^3.

4.2 Comparison with lattice results

4.2.1 q^2 spectrum and differential moments

We start our comparison of lattice and OPE results with the q^2 spectrum and the differential moments introduced in eq. (2.39) and in eq. (2.40). Only the $O(\alpha_s)$ perturbative corrections are included in this case. Figure 14 shows the q^2 spectrum in the SM, namely with a $V - A$ current. Despite the large uncertainty of the OPE prediction, about 30% in the JLQCD case and 50% in the ETMC case, the overall agreement is good. The OPE uncertainty is dominated by the power corrections. We also stress that close to the partonic endpoint, corresponding to $1.27\,{\rm GeV}^2$ and $0.82\,{\rm GeV}^2$ in the two cases, we do not expect the OPE calculation to be reliable, as discussed above. The corresponding hadronic endpoints are $1.35\,{\rm GeV}^2$ and $0.75\,{\rm GeV}^2$, respectively.

The uncertainties affecting both calculations can be greatly reduced by considering the differential moments. In particular, the OPE uncertainty becomes smaller because of the cancellations between power corrections to the numerator and to the denominator. To expose the cancellations we expand the ratios in powers of α_s and $1/m_b$. In figure 15 we show the first differential lepton energy moment, $L_1(q^2)$, in the SM, comparing the OPE with ETMC data. As expected, the relative uncertainty of both the OPE calculation and of the lattice data is much smaller than in the bottom panel of figure 14 and we observe good agreement at low and moderate q^2.

Figures 16 and 17 show the q^2 spectrum in the individual channels. Comparing them with figure 14 we see that in the individual channels the agreement between OPE and lattice results is poorer than in their sum, especially at large q^2. This is to be expected and (unless discretisation and/or finite-volume effects turn out to have a sizeable impact on the lattice results) is likely to be a manifestation of duality violations. Notice that the OPE central predictions for the $AA_\perp$ and $VV_\perp$ channels turn negative at large and moderate q^2, respectively, and that for $q^2 > 0.6\,{\rm GeV}^2$ the spectrum is always negative within errors. This unphysical feature suggests that our error estimates are not adequate at large q^2. The contribution to the $VV_\perp$ channel, moreover, is particularly small and very sensitive to large power corrections.
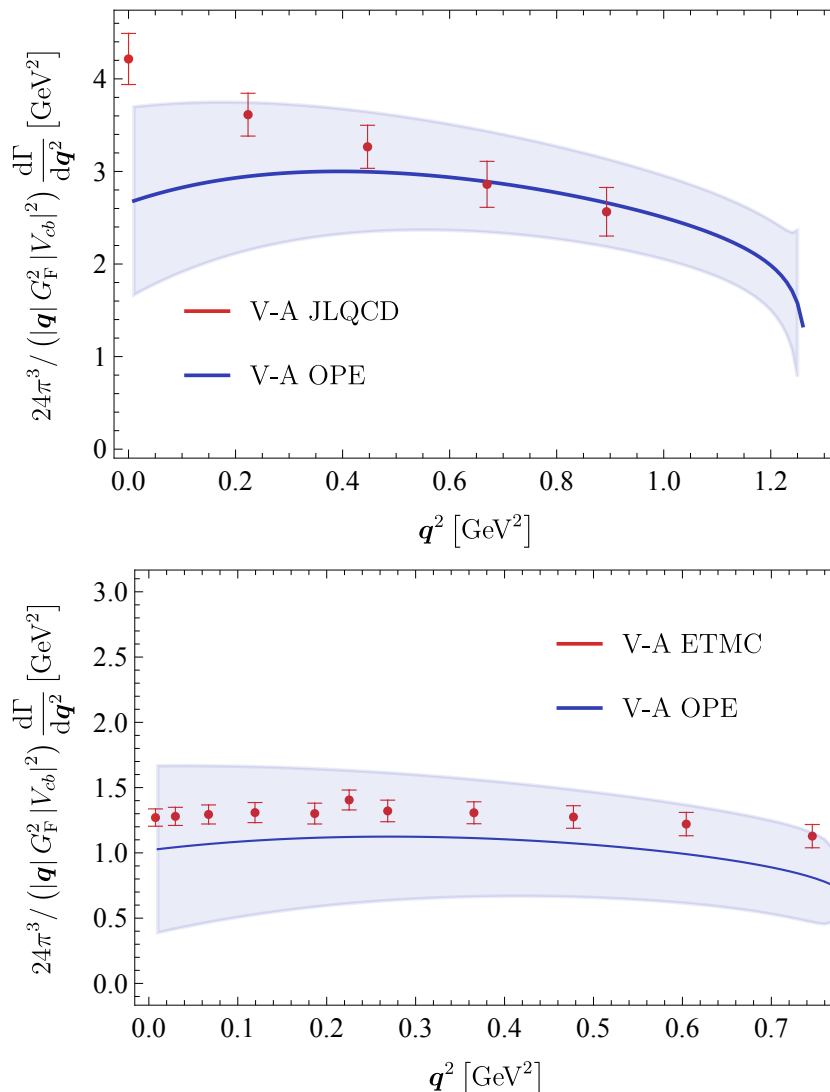
Figure 14. Differential q^2 spectrum, divided by $|\boldsymbol{q}|$, in the SM. Comparison of OPE with JLQCD (top panel) and ETMC (bottom panel) data are shown.

Figures 18 and 19 show $L_1(\boldsymbol{q}^2)$ in the individual channels, for the JLQCD and ETMC cases. In general, we observe good agreement with the lattice data, especially at low $\boldsymbol{q}^2$. However, the expansion in powers of α_s and $1/m_b$ of the denominator is not justified when the lowest-order contribution to the denominator becomes particularly small or has a zero, like in the $VV_\parallel$ and $VV_\perp$ channels. In these cases we also show the unexpanded version of the ratio, whose uncertainty is much larger, but we stress that away from the singularities the expanded form is preferable, and this appears to be confirmed by better agreement with the lattice data.

Figure 20 shows the second central moment computed at different values of $\boldsymbol{q}^2$ in the ETMC case. We do not display the $VV_\perp$ channel, for which the OPE result would have a very large uncertainty. In the case of $L_{2c}(\boldsymbol{q}^2)$ the OPE does not reproduce the lattice results
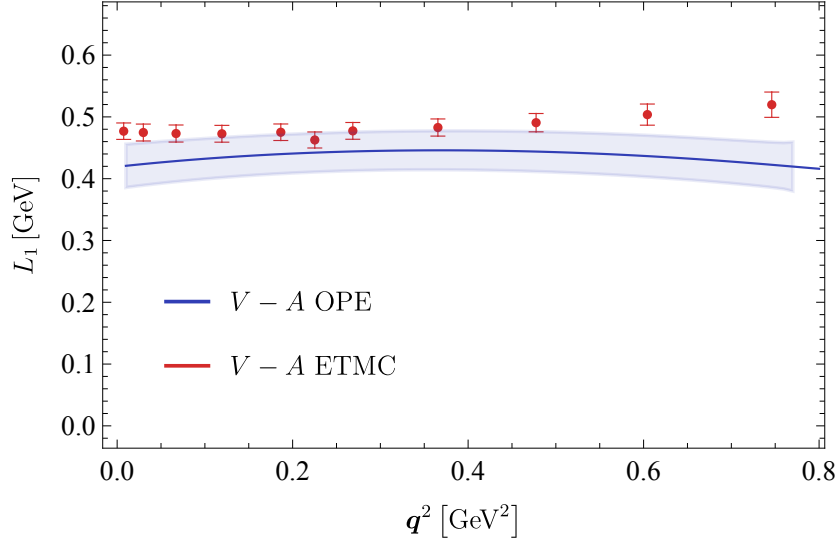
Figure 15. Differential lepton energy mean value, $L_1(\boldsymbol{q}^2)$, in the SM. The comparison of OPE with ETMC data is shown.
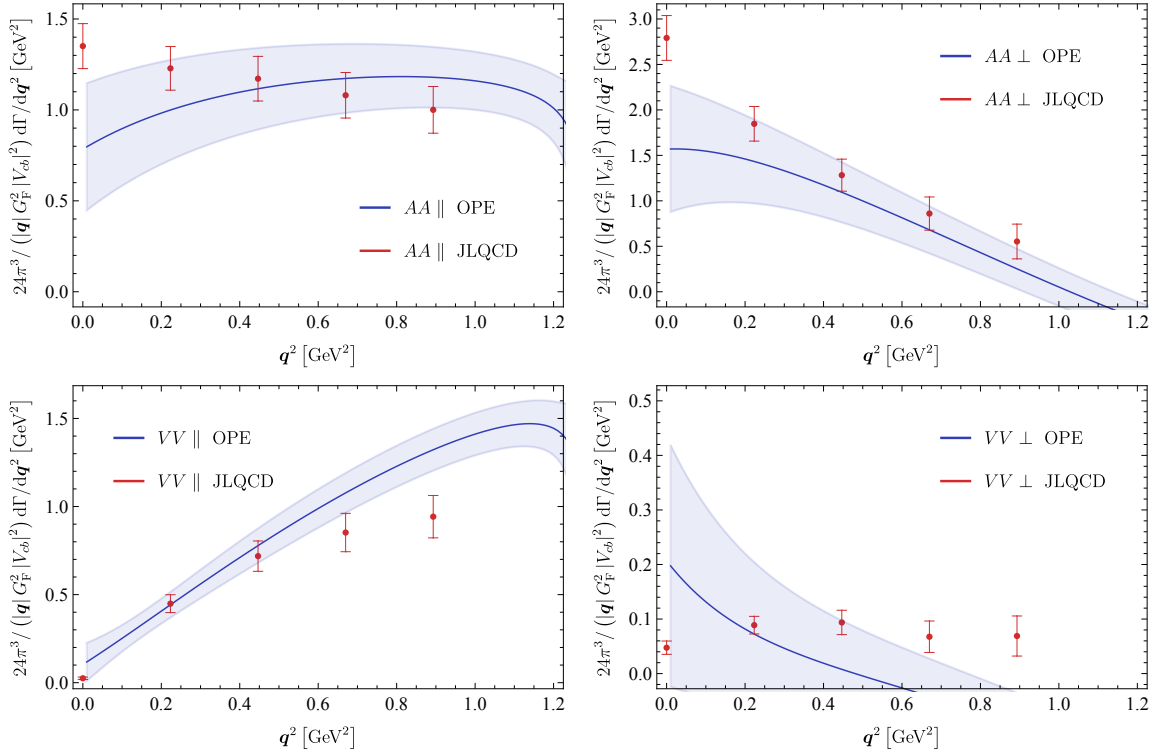
Figure 16. Differential $\boldsymbol{q}^2$ spectrum, divided by $|\boldsymbol{q}|$, in the various channels. The plots show the comparison between OPE and JLQCD data.
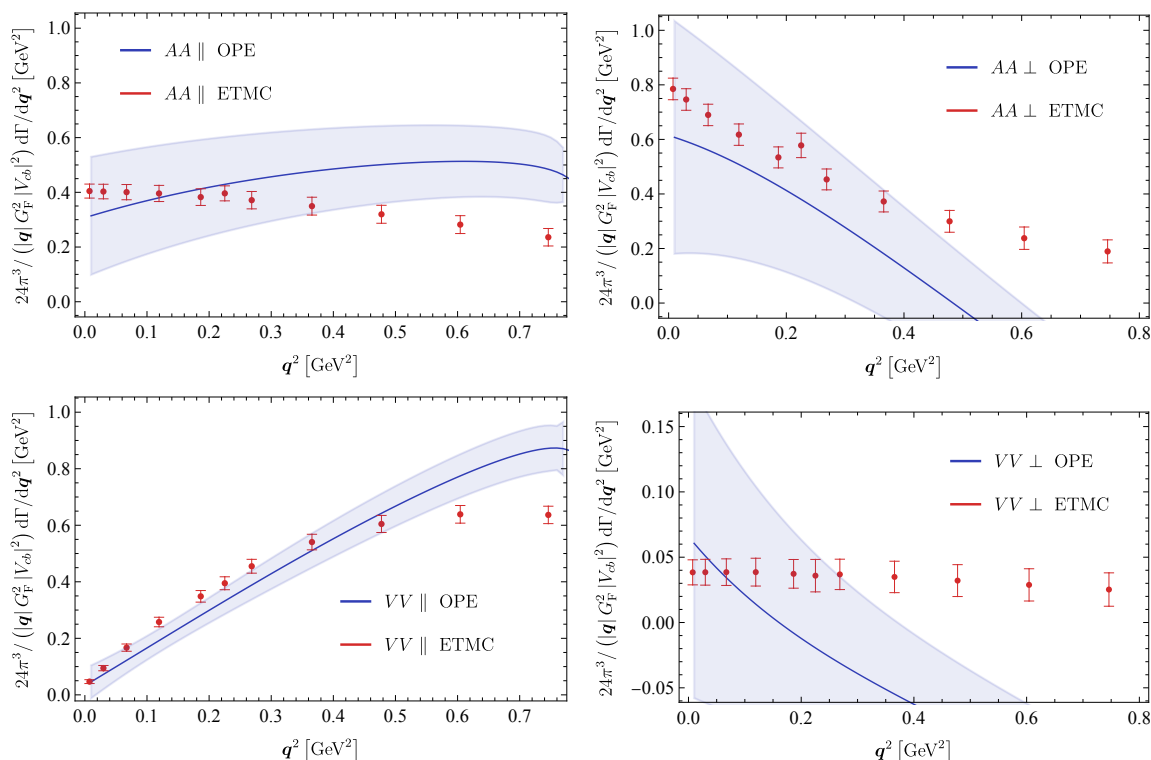
Figure 17. Differential q^2 spectrum, divided by $|\mathbf{q}|$, in the various channels. The plots show the comparison between OPE and ETMC data.

within uncertainties, except for very small q^2. It is certainly possible that our method to estimate the OPE uncertainty fails here as a result of multiple cancellations between large contributions to L_2 and L_1^2 which are not necessarily replicated by higher-order contributions. On the other hand, it has not yet been possible to estimate discretisation and finite volume effects on our lattice results, and the additional systematics could affect this particular quantity in a relevant way. For this quantity we do not display the comparison with the JLQCD data, which agree with the OPE but have very large uncertainties.

We also looked at the moments of the hadronic invariant mass. In figure 21 we show the mean hadronic mass $\langle M_X^2 \rangle$ as a function of q^2 computed from JLQCD configurations in comparison with the OPE predictions. Again, we do not display the $VV_\perp$ channel because of the large OPE uncertainty. We observe excellent agreement except at large q^2, but the lattice uncertainty is larger here than in the case of the leptonic moments. Analogous plots for the ETMC calculation are shown in figure 22. In figure 23 we also show $\langle M_X^4 \rangle$ as a function of q^2 with JLQCD data.

4.2.2 Total width and moments

We perform a comparison between OPE predictions and lattice results also in the case of the total semileptonic width and of the global moments introduced in eq. (2.37) and in eq. (2.38). In this case the OPE results are going to be slightly more accurate as we can take advantage of existing two- and even three-loop calculations [62]. We can also
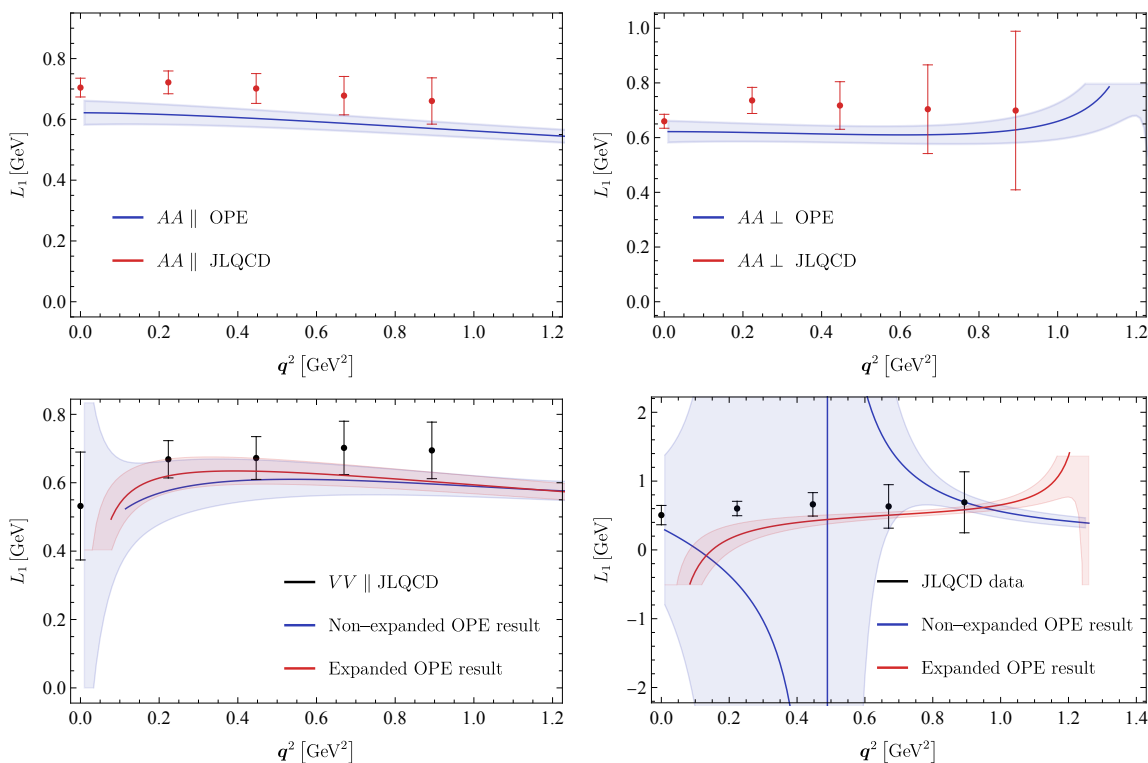
Figure 18. Differential first leptonic moment in the various channels. The plot shows the comparison between OPE and JLQCD data.

test the relevance of the singularity at $q^2_{\max}$. The lattice results for the q^2 spectrum can be interpolated by polynomials or piecewise polynomials, leading to the results shown in table 2 and in table 3. As the q^2-spectrum is peaked near $q^2_{\max}$, see figure 10, the total width is particularly sensitive to that region. In the JLQCD case the limited number of q^2 points makes the extrapolation to the highest q^2 values more uncertain, with clear implications on the estimate of the total width. On the other hand, it is difficult to estimate such uncertainty, hence table 2 shows only the statistical uncertainty.

In the OPE the total width receives large and concurring power and perturbative corrections, which reflect in a $\sim$ 20–40% uncertainty. This is at variance with what happens in the case of the physical b quark, for which a recent estimate of the total uncertainty is about 2% [68]. Indeed, the convergence of the OPE expansion deteriorates rapidly as m_b decreases approaching m_c, even from 2.7 to 2.4 GeV. To illustrate this point we show the various contributions to the semileptonic width in the ETMC case:

$$\frac{\Gamma}{|V_{cb}^2|} = \left[3.03 - 0.32_{\text{pert}} - 0.65_{\mu_G^2} - 0.09_{\mu_\pi^2} - 0.66_{\rho_D^3} - 0.10_{\rho_{LS}^3} + \dots \right] \times 10^{-13} \, \text{GeV} \,, \quad (4.6)$$

where the perturbative contribution includes $O(\alpha_s^3)$ and the non-perturbative contributions include the $O(\alpha_s)$ corrections to the Wilson coefficients. We estimate the perturbative uncertainty by varying the scale of α_s between 1.5 and 3.0 GeV. Notice that more than half of the uncertainty on the width reported in tables 2 and 3 is due to the large uncertainty on the heavy quark masses, in both the JLQCD and ETMC cases.
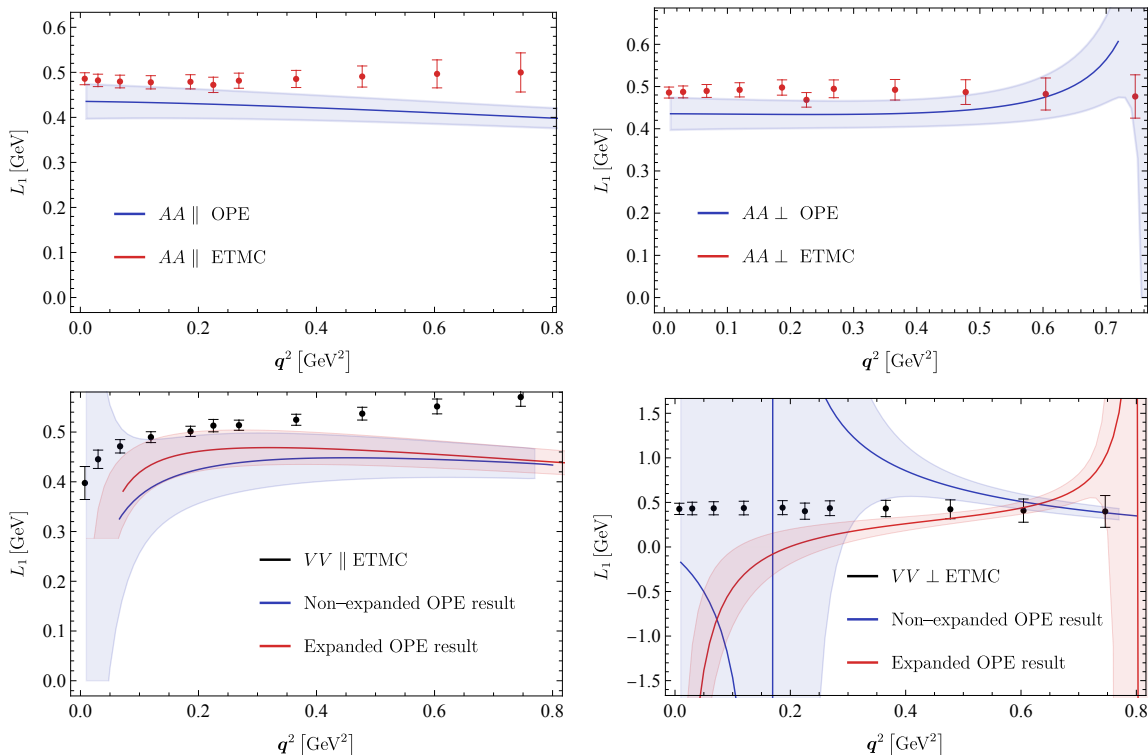
Figure 19. Differential moment $L_1(\boldsymbol{q}^2)$ in the various channels. The plots show the comparison between OPE and ETMC data.

	JLQCD	OPE
$\Gamma/|V_{cb}^2| \times 10^{13}$ (GeV)	4.46(21)	5.7(9)
$\langle E_\ell \rangle$ (GeV)	0.650(40)	0.626(36)
$\langle M_X^2 \rangle$ (GeV^2)	3.75(31)	4.22(30)

Table 2. Total width and moments in the JLQCD case.

	ETMC	OPE
$\Gamma/|V_{cb}^2| \times 10^{13}$ (GeV)	0.987(60)	1.20(46)
$\langle E_\ell \rangle$ (GeV)	0.491(15)	0.441(43)
$\langle E_\ell^2 \rangle$ (GeV^2)	0.263(16)	0.207(49)
$\langle E_\ell^2 \rangle - \langle E_\ell \rangle^2$(GeV^2)	0.022(16)	0.020(8)
$\langle M_X^2 \rangle$ (GeV^2)	3.77(9)	4.32(56)

Table 3. Total width and moments in the ETMC case.

For what concerns the leptonic moments, only the $O(\alpha_s^2)$ corrections have been computed, either numerically for physical values of the heavy-quark masses [60], or analytically in an expansion up to $O(r^7)$ in powers of $r = m_c/m_b$ [58]. Unfortunately, this expansion converges slowly and does not provide reliable results for $r \sim 0.5$, which is the value
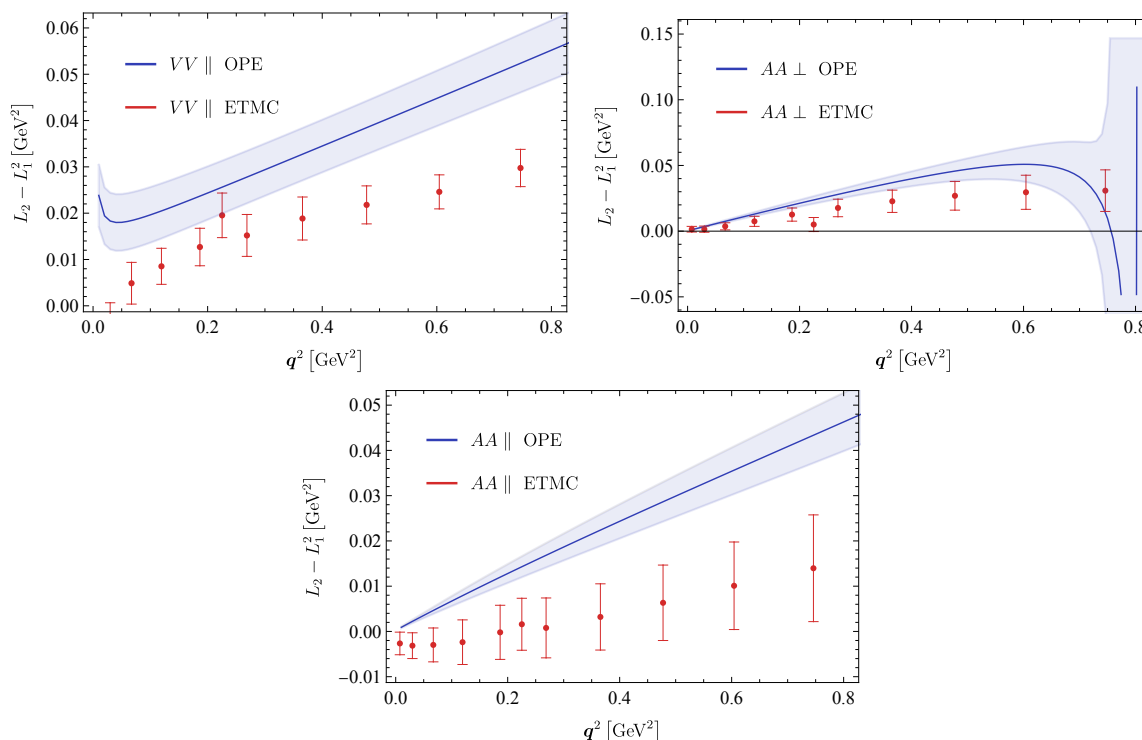
Figure 20. Differential moment $L_{2c} = L_2 - L_1^2$ in the various channels. The plots show the comparison between OPE and ETMC data.

relevant in the ETMC case. We therefore show results computed to $O(\alpha_s)$ and include the $O(\alpha_s \mu_{\pi,G}^2 / m_b^2)$ corrections discussed in ref. [55] as well. The first moment in the ETMC case is given by

$$\langle E_\ell \rangle = \left[0.533 - 0.051_{\mu_G^2} + 0.021_{\mu_\pi^2} - 0.051_{\rho_D^3} - 0.003_{\rho_{LS}^3} - 0.008_{\alpha_s} + \ldots \right] \text{GeV}, \quad (4.7)$$

where both power and perturbative corrections are smaller than in the total width. Similarly, the second central moment $L_{2c} = \langle E_\ell^2 \rangle - \langle E_\ell \rangle^2$ is given by

$$L_{2c} = \left[0.0297 - 0.0057_{\mu_G^2} + 0.0103_{\mu_\pi^2} - 0.0167_{\rho_D^3} + 0.0006_{\rho_{LS}^3} + 0.0021_{\alpha_s} + \ldots \right] \text{GeV}. \quad (4.8)$$

As shown in tables 2 and 3, there is reasonable agreement between OPE and both JLQCD and ETMC data in all cases. As a general comment, we stress that the large contributions of ρ_D^3 are related to a kinematically enhanced Wilson coefficient and do not necessarily imply similarly large higher-power corrections.

Finally, the OPE prediction for the first hadronic mass moment in the JLQCD case is

$$\langle M_X^2 \rangle = \left[3.84 - 0.36_{\mu_\pi^2} + 0.23_{\mu_G^2} + 0.41_{\rho_D^3} + 0.05_{\rho_{LS}^3} + 0.04_{\alpha_s} + \ldots \right] \text{GeV}^2, \quad (4.9)$$

where we do not include the $O(\alpha_s / m_b^2)$ corrections and consequently enlarge the uncertainty slightly. The OPE prediction for the first hadronic moment is in reasonable agreement with both the JLQCD and ETMC values, see table 2 and table 3.
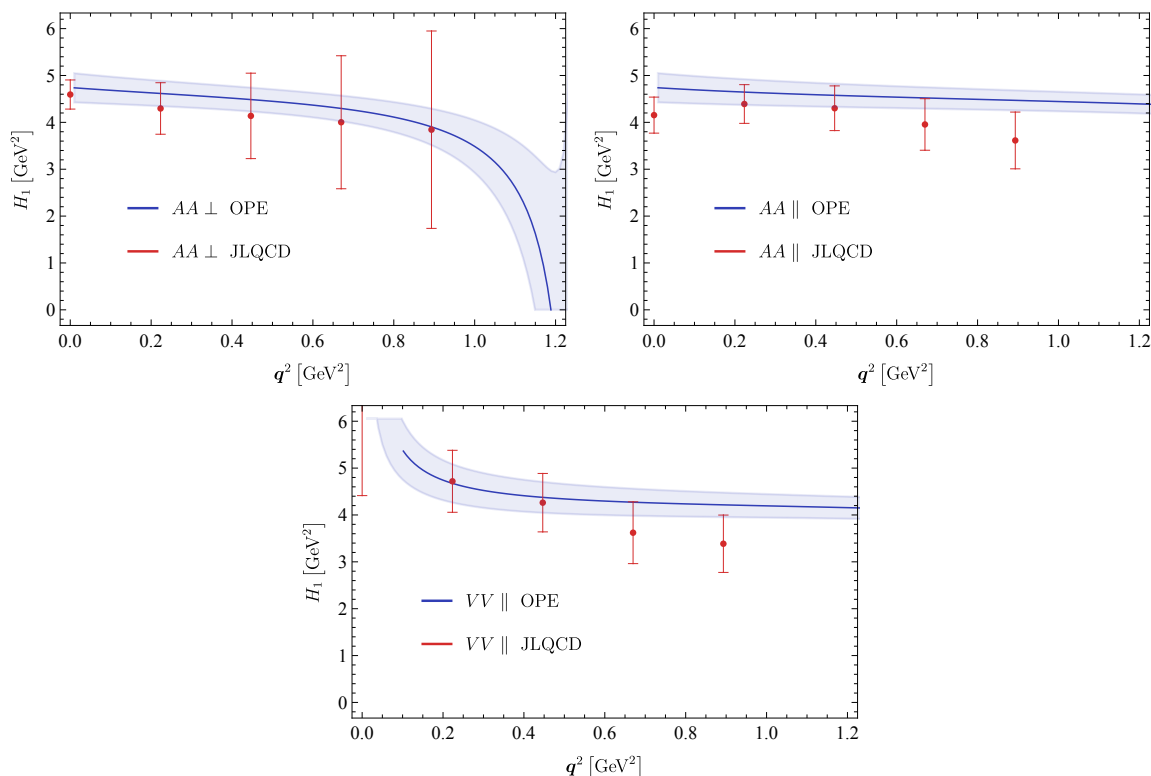
Figure 21. Differential moment $H_1(\boldsymbol{q}^2) = \langle M_X^2 \rangle(\boldsymbol{q}^2)$ in the various channels. The plots show the comparison between OPE and JLQCD data.

4.3 Determination of the OPE parameters

As different physical quantities have a different dependence on the OPE parameters, it is possible to constrain their values using lattice data. The analytic expressions for the power corrections to the differential $\boldsymbol{q}^2$ distribution and for the moments, which encode this dependence, are rather lengthy and are provided in a supplementary Mathematica file.

To illustrate this point, let us consider a few examples using simpler approximate formulas, and focussing on the differential leptonic moments at moderately low $\boldsymbol{q}^2$, where the OPE is more reliable. We choose a $\boldsymbol{q}^2$ value for which we have lattice data, $\boldsymbol{q}_*^2 = 0.1865\,\text{GeV}^2$. In the ETMC setup, the OPE prediction for $L_1(\boldsymbol{q}_*^2)$ can be approximated by

$$L_1^{VV\parallel}(\boldsymbol{q}_*^2) \simeq 0.5597 + \frac{1}{2}\delta_b - 0.47\delta_c + 0.056\mu_G^2 - 0.19\mu_\pi^2 - 0.094\rho_D^3 - 0.057\rho_{LS}^3 \,,$$

$$L_1^{AA\parallel}(\boldsymbol{q}_*^2) \simeq 0.5455 + \frac{1}{2}\delta_b - 0.47\delta_c - 0.141\mu_G^2 - 0.074\mu_\pi^2 - 0.069\rho_D^3 + 0.043\rho_{LS}^3 \,,$$

$$L_1^{AA\perp}(\boldsymbol{q}_*^2) \simeq 0.5448 + \frac{1}{2}\delta_b - 0.47\delta_c - 0.175\mu_G^2 - 0.033\mu_\pi^2 - 0.101\rho_D^3 + 0.039\rho_{LS}^3 \,,$$

where $\delta_b = m_b - 2.39$, $\delta_c = m_c - 1.19$, and all quantities are expressed in GeV to the appropriate power. Notice that the lowest order expression for the differential leptonic moments is universal, namely does not depend on the channel. We do not consider the $VV_\perp$ channel because, as discussed above, the expanded form does not provide a good
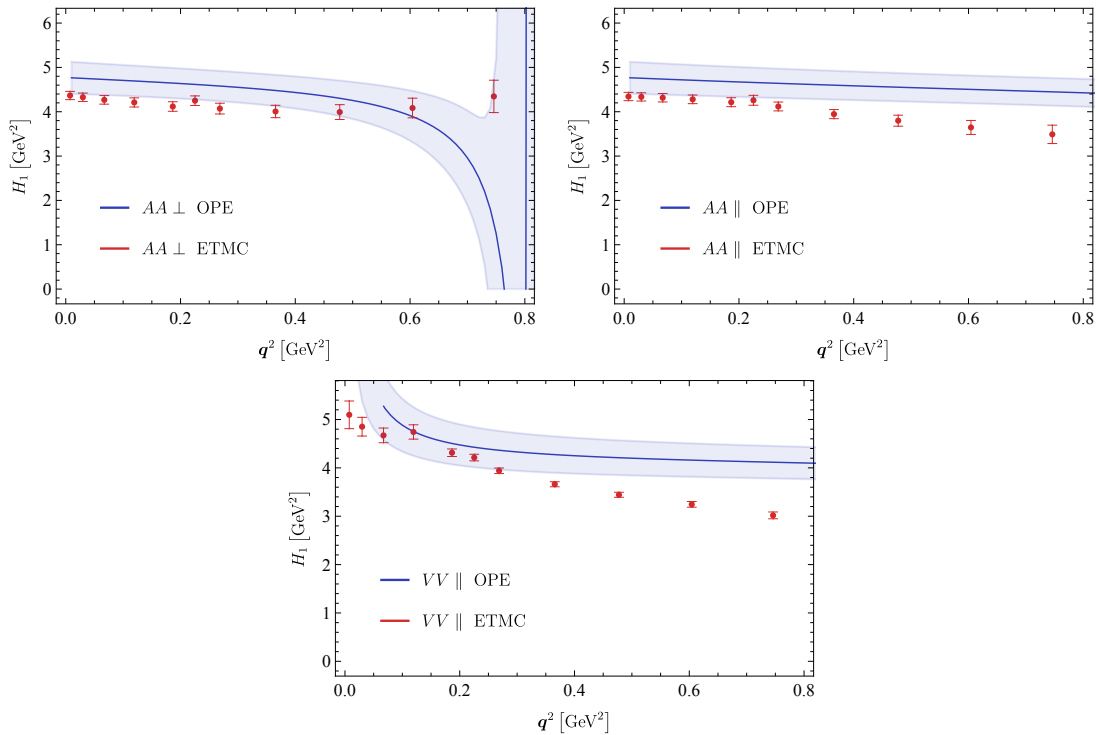
Figure 22. Differential moment $H_1(\boldsymbol{q}^2) = \langle M_X^2 \rangle(\boldsymbol{q}^2)$ in the various channels. The plots show the comparison between OPE and ETMC data.
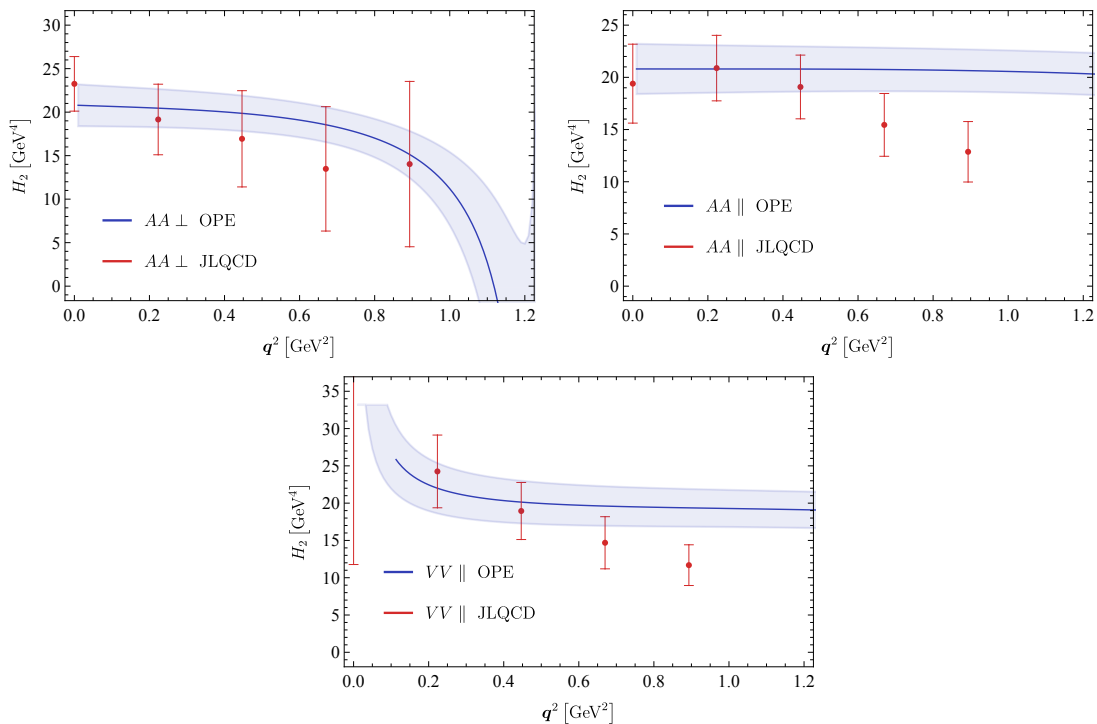
Figure 23. Differential moment $H_2(\boldsymbol{q}^2) = \langle M_X^4 \rangle(\boldsymbol{q}^2)$ in the various channels. The plots show the comparison between OPE and JLQCD data.

approximation. The analogous expressions for the second central moments are

$$L_{2c}^{VV\parallel}(\boldsymbol{q}_*^2) \simeq 0.005 + 0.010\mu_\pi^2 + 0.052\rho_D^3 - 0.015\rho_{LS}^3\,,$$
$$L_{2c}^{AA\parallel}(\boldsymbol{q}_*^2) \simeq 0.009 + 0.010\mu_\pi^2 - 0.058\rho_D^3 + 0.011\rho_{LS}^3\,,$$
$$L_{2c}^{AA\perp}(\boldsymbol{q}_*^2) \simeq 0.019 + 0.010\mu_\pi^2 - 0.026\rho_D^3 - 0.002\rho_{LS}^3\,.$$

Each of these moments has a different dependence on the non-perturbative parameters and they can be used in a fit to the lattice ETMC results to obtain constraints on those parameters. In fact, using only these six inputs with their theoretical uncertainty does not lead to any improvement on the constraints given in table 1. Considering additional $\boldsymbol{q}^2$ points enhances the sensitivity to the non-perturbative parameters, but one has to estimate the correlation among the theoretical uncertainties at adjacent $\boldsymbol{q}^2$ points. One can also include in the fit the data for the $\boldsymbol{q}^2$ distribution in the different channels, as well as additional moments like the hadronic mass moments. A global fit to lattice data is however beyond the scope of this paper, especially because our estimate of the lattice systematic uncertainty is incomplete. We stress that the limiting factor here is not the statistical uncertainty of the present ETMC calculation, but the theoretical uncertainty we attach to the OPE predictions. In this respect the unphysical case we have considered, with the partonic energy release (of the order of $m_b - m_c$) about a factor 2 (JLQCD) or 3 (ETMC) smaller than in reality, is strongly penalising. At the physical point the OPE enjoys a much better convergence and the prospects for constraining the non-perturbative parameters are better than it appears from this exploratory study.

4.4 Computations with a smooth kernel

In sections 2 and 3 we have seen that the reconstruction of the discontinuous kernel is one of the main problems in the calculation of physical quantities. As far as the comparison with the OPE is concerned, however, the kernel does not need to be discontinuous. Indeed, one can compute inclusive (unphysical) quantities in the OPE employing a smooth kernel ($\sigma \neq 0$) and compare them directly with the analogous quantities computed on the lattice. In this way it is possible to check that the level of agreement between the two calculations is not affected by the $\sigma \to 0$ limit, and to extract information on the non-perturbative parameters of the OPE, as well as on the heavy quark masses, from slightly more precise lattice data.

In figure 24 we show the $\boldsymbol{q}^2$ spectrum in the different channels computed on the lattice using the sigmoid approximation $\theta_\sigma^{\mathsf{s}}$ of eq. (2.36) for $\theta(\omega_{\max} - \omega)$ with $\sigma = 0.12m_B$. In the OPE calculation, where the partonic kinematics holds, we replace $\theta(1 - \hat{\eta} - \sqrt{\boldsymbol{q}^2})$ by the sigmoid $\theta_\sigma^{\mathsf{s}}(1 - \hat{\eta} - \sqrt{\boldsymbol{q}^2})$ using $\sigma = 0.12m_B$. At low $\boldsymbol{q}^2$ the agreement between OPE and ETMC data is similar to that in figure 17, while at large $\boldsymbol{q}^2$ there is marginal improvement, as expected because the smearing occurs over a larger ω range. In figure 25 we show the first differential leptonic moment $L_1(\boldsymbol{q}^2)$ in the different channels, excluding $VV_\perp$ because of the large uncertainties in the OPE calculation. With respect to figure 19 we observe a marked improvement of the agreement between OPE and ETMC data at large $\boldsymbol{q}^2$ in the $AA_\parallel$ and $VV_\parallel$ channels, while in the $AA_\perp$ channel the agreement is slightly worse. Finally, in figure 26 we show the $\boldsymbol{q}^2$ spectrum in the different channels computed from the
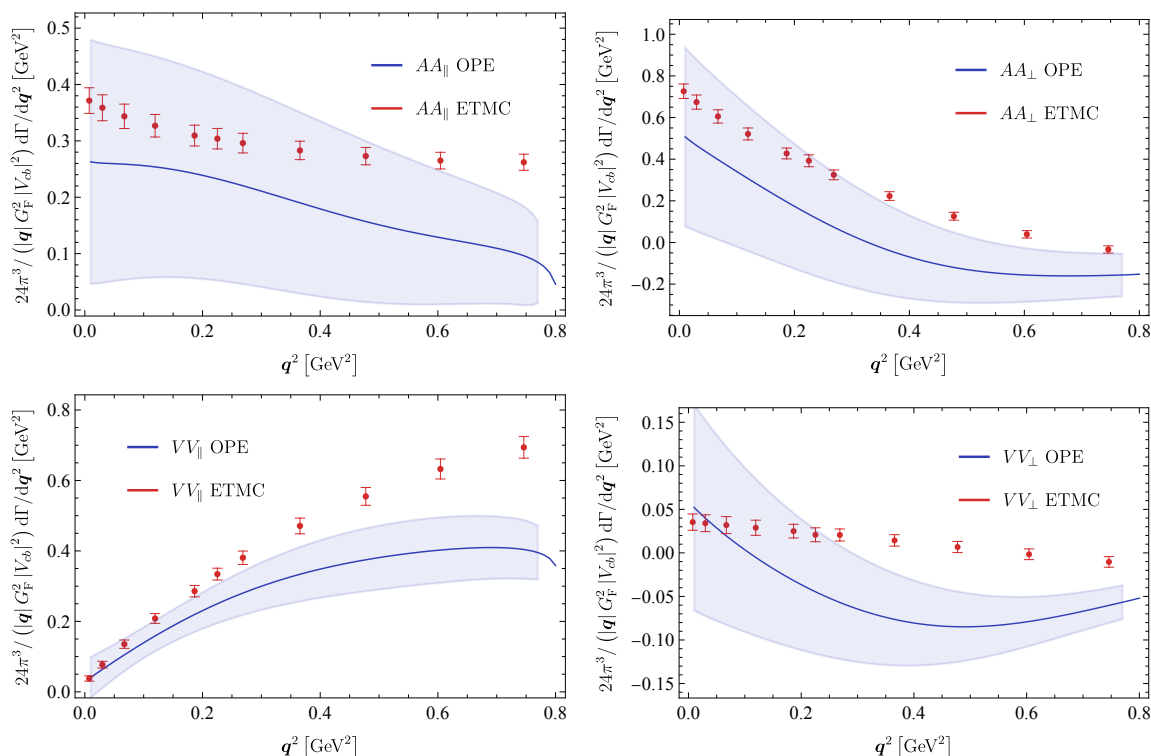
Figure 24. Differential q^2 spectrum computed with a sigmoid approximation to the kernel with $\sigma = 0.12 m_B$. The plots show the comparison between OPE and ETMC data.

JLQCD configurations using the sigmoid approximation $\theta_\sigma^{\mathtt{s}}$ with $\sigma = 0.1/a$. Here the overall agreement between lattice calculations and OPE is similar to figure 16, but now the q^2 dependence of the lattice data is closer to the OPE result, obtained using $\theta_\sigma^{\mathtt{s}}(1 - \hat{\eta} - \sqrt{q^2})$ with $\sigma = 0.1$.

5 Discussion and future prospects

In this article we have presented the first comprehensive investigation of inclusive semileptonic B-meson decays on the lattice. Using the method of ref. [16] we have computed various inclusive observables with gauge-field ensembles generated by the JLQCD and ETM collaborations for unphysically light values of the b quark mass (about $2.7 \, \mathrm{GeV}$ and $2.4 \, \mathrm{GeV}$, respectively) and m_c close to its physical value. In this exploratory study we have not performed the continuum and infinite-volume limits.

An important feature of the method we have adopted is that it requires the approximation of the energy-integral kernel. The kinematics of the inclusive semileptonic decay involves a discontinuity at the boundary of the phase space, for which a reasonable approximation with the Euclidean correlator obtained on the lattice is impractical. The problem can be dealt with using a sequence of smooth kernels, parametrized by a smearing width σ, that converge to the physical phase space in the limit $\sigma \to 0$. As emphasized in section 2, the $\sigma \to 0$ limit does not commute with the infinite-volume limit that has to be taken first.
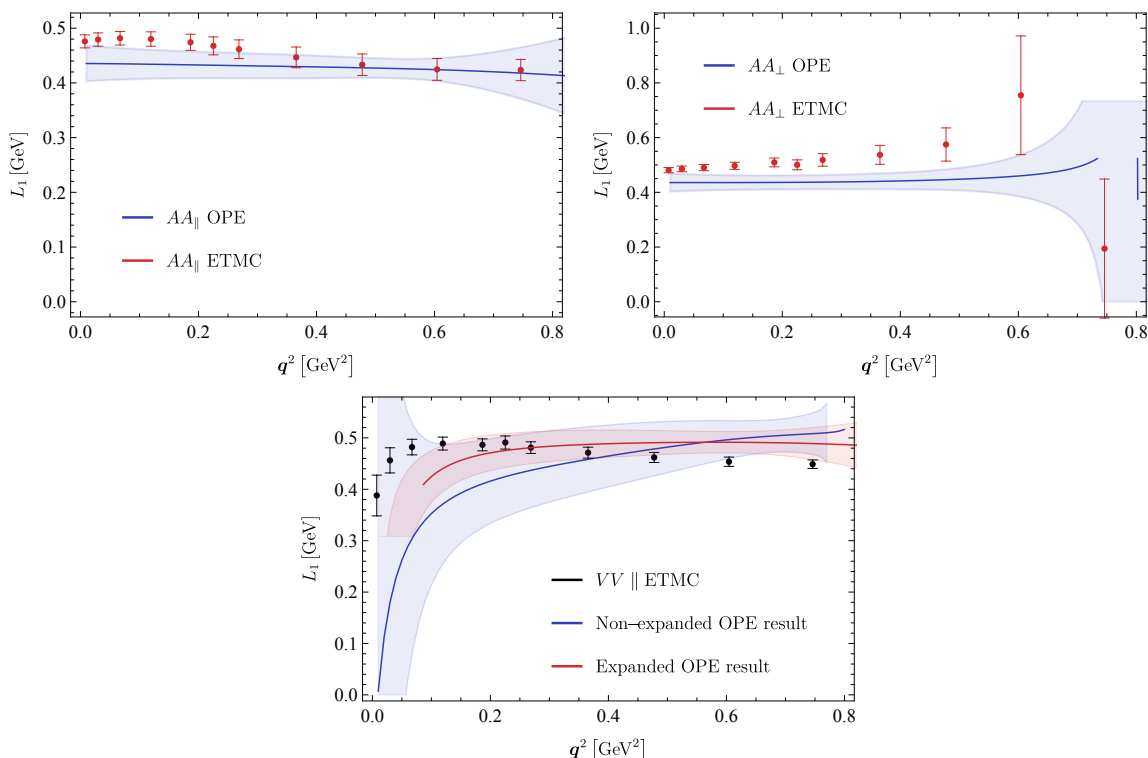
Figure 25. Differential first lepton moments computed with a sigmoid approximation to the kernel with $\sigma = 0.12 m_B$. The plots show the comparison between OPE and ETMC data.

Under the assumption that finite volume effects are negligible with respect to the statistical errors associated with our lattice results, we have studied the $\sigma \to 0$ extrapolation in detail and found that it does not induce a significant uncertainty.

We have compared the JLQCD results with the contributions of the charmed ground states, estimated from a JLQCD calculation of the $B_s \to D_s^{(*)}$ form factors for the same values of the heavy-quark masses (details are given in the appendix A). Due to the proximity between the charm and bottom masses and to the limited phase space available in the decay, the inclusive results are nearly saturated by the ground-state contributions. Although the correlator themselves show the presence of excited states, their contribution to the inclusive rate is relatively small. The ETMC results obtained at even lower b quark mass are also expected to be largely dominated by the ground states.

While the JLQCD and ETMC results cannot be compared directly as they are obtained at different b quark masses, they can be both compared with the expectations from the OPE, assuming that discretisation and finite-volume effects are negligible. When the OPE can be considered reliable, the agreement with both JLQCD and ETMC results is generally good, while we observe possible indications of quark-hadron duality violation at large q^2. The variance of the lepton energy distribution also shows a clear and unexpected deviation, which could be due to underestimated uncertainties in our OPE calculation or to non-negligible lattice systematics. To the best of our knowledge, this is the first time that the onset of quark-hadron duality is studied on the lattice. For $m_b \sim 2.4\,\text{GeV}$, the OPE
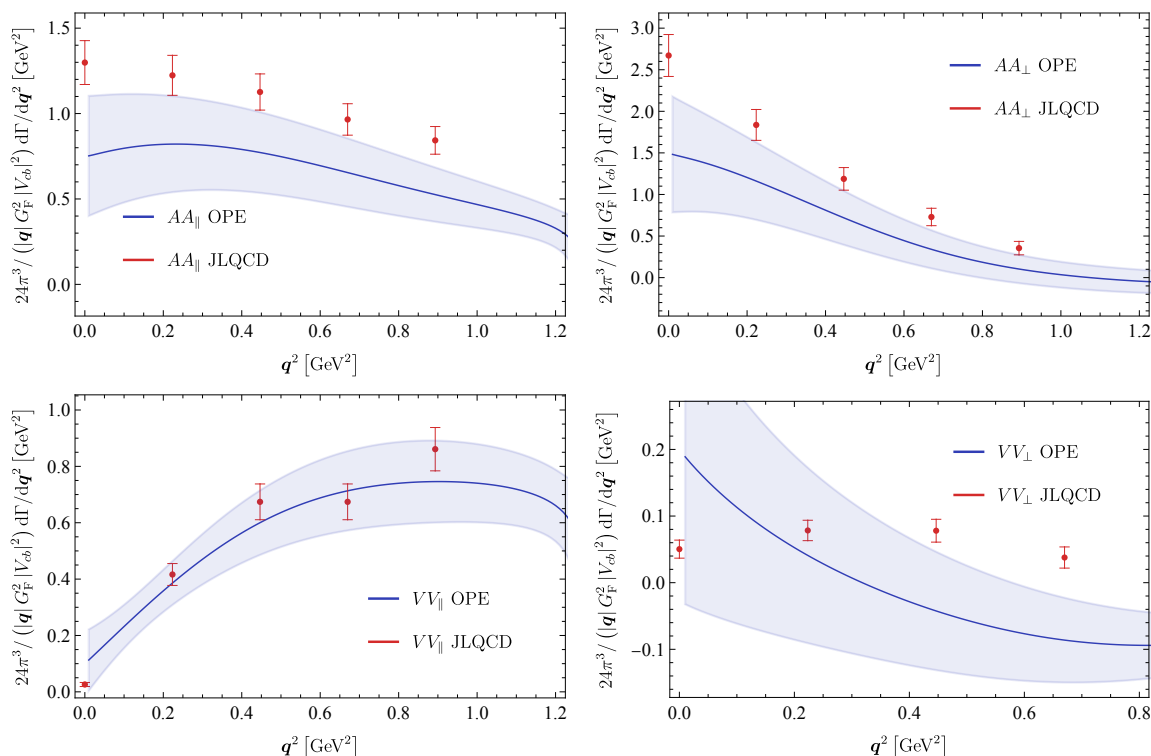
Figure 26. Differential q^2 spectrum computed with a sigmoid approximation to the kernel with $\sigma = 0.1/a$. The plots show the comparison between OPE and JLQCD data.

converges much more slowly than at the physical point, but the normalised moments allow us to perform a relatively clean comparison with the lattice data.

We have found that the calculation of the total width and of other global quantities like the moments of the lepton energy or of the hadronic invariant-mass distribution depends crucially on the number of q^2 points that can be computed on the lattice. In the ETMC calculation the flexibility due to the use of twisted boundary conditions has allowed us to reach an accuracy of 6% on the total width and 3% on the first leptonic moment. These uncertainties do not yet include several lattice systematics that need to be considered, but are dominated by statistical uncertainties and could be improved with a dedicated effort. This is an aspect that will become important for future phenomenological applications, which should also focus on reaching the physical b mass.

Finally, we have shown that one can constrain the non-perturbative parameters in the OPE analysis from our results. We have not attempted a fit to the lattice data in the unphysical setup we have considered, as this is penalised by large uncertainties from higher-dimensional operators. With larger b-quark masses these uncertainties will be reduced and the data obtained at different values of m_b will provide an additional handle on the non-local matrix elements that appear in eq. (4.5).

There are certainly many issues to be improved or investigated in order to get results of direct phenomenological relevance. First, we have not yet studied the continuum and infinite-volume limits. Although we have presented a rather detailed discussion of the

systematics associated with the reconstruction of the smearing kernels, including the required extrapolation at vanishing smearing radius, this last step is only permitted after having checked the onset of the infinite-volume limit. The continuum and infinite-volume limits can only be taken by performing calculations at different values of the lattice spacing and at different physical volumes, a task that is beyond the exploratory nature of this study and that we postpone to future work on this subject.

Second, the calculation has to be performed at the physical b and light quark masses. Simulations with physical pion masses are nowadays possible and, for instance, a collection of $N_f = 2 + 1 + 1$ ensembles with physical light, strange and charm quark masses has been produced by the ETM collaboration at different values of the lattice spacing and with different physical volumes. Although it is not possible to simulate directly a physical b quark on these ensembles (because of potentially dangerous cutoff effects), the problem can nevertheless be approached by using well-established techniques such as the ETMC ratio method [71], based on ratios of the observable of interest computed at nearby heavy-quark masses. The ratio method has been already applied to determine the mass of the b quark, the leptonic decay constants, the bag parameters of $B_{(s)}$ mesons and the matrix elements of dimension-four and dimension-five operators appearing in the Heavy Quark Expansion of pseudoscalar and vector meson masses [65, 66, 72–74]. Its main advantages can be summarised as follows: i) B-physics computations can be carried out using the same relativistic action setup with which the lighter-quark computations are performed; ii) an extra simulation at the static point limit is not necessary, while the exact information about it is automatically incorporated in the construction of the ratios of the observable; iii) the use of ratios greatly helps in reducing the discretisation effects. However, there is an important subtlety. In order to apply the ratio method (or any other method based on extrapolations in the b-quark mass) in the case of the inclusive decay rates one has to cope with the fact that at unphysical (lighter) values of the b-quark mass the phase-space available to the decay shrinks. This implies that some of the hadronic channels that are open at the physical value of m_b are *totally* excluded from the phase-space integral at $m_h < m_b$. The important point to be noticed here is that this happens when the integration limits are imposed sharply, i.e. by using the exact Heaviside functions that implement the phase-space constraint. The problem is totally analogous to the ordered double-limit required in order to deal with the finite-volume distortion of the hadronic spectral density. Indeed, we envisage applying the ratio method to $\Gamma \equiv \Gamma(m_b)$ *before* taking the $\sigma \to 0$ extrapolation: while $\Gamma(m_h)$ is (at least in principle) a distribution in m_h, $\Gamma_\sigma(m_h)$ is certainly a smooth function that can safely be extrapolated at the physical value of m_b. Moreover, we already have simulations with $m_h \sim 0.8 m_b$, and it can be reasonably argued that for such large masses the missing (mostly continuum) states scale with m_h.

Although we have compared the lattice results with the OPE, a more direct and effective validation of our method would come from a comparison with experimental data, such as those for the branching ratio and for the electron energy spectrum in inclusive semileptonic decays of the D or D_s mesons [75, 76]. Here the challenge is to get accurate results at physical light-quark masses, while the charm quark can be simulated directly on present lattices. Beside validating the method without extrapolations in the heavy-

quark mass, a calculation of charm decays might shed light on the following two open and phenomenologically relevant questions. *i*) To what extent is the OPE applicable to charm decays? *ii*) What is the role played by weak annihilation (WA) contributions? The first question refers to the onset of quark-hadron duality, and a detailed study of charm decays in connection with their OPE description may yield an insight on this conceptual issue. Answering the second question may help us quantifying the role played by WA contributions in charmless semileptonic B decays, hence improving the inclusive determination of $|V_{ub}|$. If one could reproduce the lepton energy spectrum of the D_s inclusive semileptonic decays that is measured experimentally, a more ambitious future application would be a direct calculation of $B \to X_u \ell \nu$.

Finally, one may wonder whether the foreseeable precision will be sufficient for a precision determination of $|V_{cb}|$ and for interesting phenomenology. Indeed, present experimental errors for $B \to X_c \ell \nu$ are 1.4% on the branching ratio and a few per mille on the first few moments of the lepton energy distribution. The lattice precision is unlikely to get close to that, at least initially. On the other hand, on a relatively short time-scale lattice calculations of inclusive semileptonic decays might be able to enhance the predictive power of the OPE by accessing other quantities that are inaccurate or beyond the reach of current experiments and are highly sensitive to the non-perturbative parameters, allowing us to validate and improve the results of the semileptonic fits on which the OPE predictions are based.

Acknowledgments

A Contributions from the ground states

Among the complete set of states inserted in eq. (2.4), we consider the contribution of the lowest-lying states, which are the S-wave states, i.e. D and D^* mesons. (Here and in the following, we omit the subscript s for brevity.) The corresponding matrix elements can be
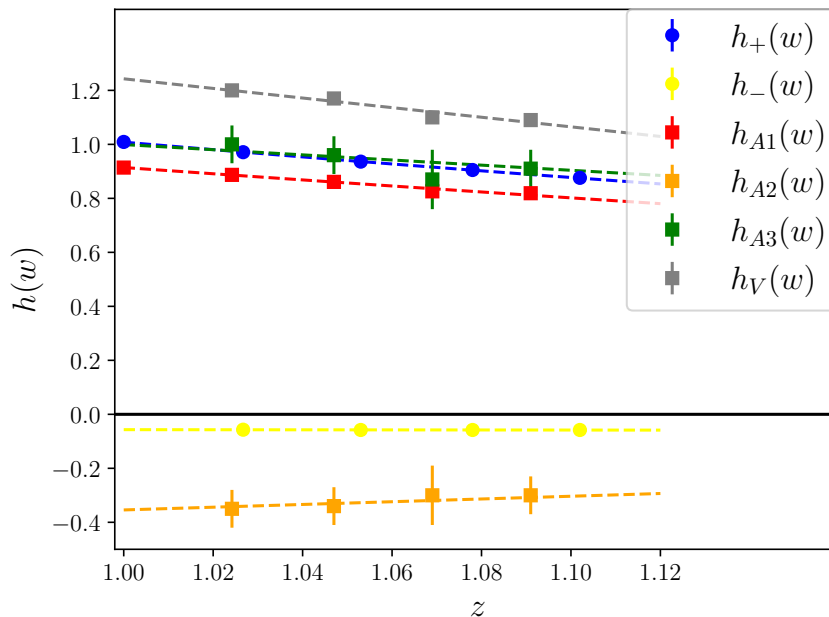
Figure 27. Form factors computed from three-point functions.

parametrized by the form factors as

$$\langle D(v')|V^\mu|B(v)\rangle = h_+(w)(v+v')^\mu + h_-(w)(v-v')^\mu \,, \tag{A.1}$$

$$\langle D^*(v',\epsilon)|V^\mu|B(v)\rangle = -h_V(w)\epsilon^{\mu\nu\rho\sigma}v_\nu v'_\rho\epsilon^*_\sigma \,, \tag{A.2}$$

$$\langle D^*(v',\epsilon)|A^\mu|B(v)\rangle = -ih_{A_1}(w)(1+w)\epsilon^{*\mu}$$
$$+ i\left[h_{A_2}(w)v^\mu + h_{A_3}(w)v'^\mu\right]\epsilon^* \cdot v \,, \tag{A.3}$$

where ϵ^* denotes the polarization vector of the vector D^* meson. We use the HQET definition of the meson states, so that the kinematics is parametrized by the velocities v and v' (with $p = m_B v$ and $p' = m_{D^{(*)}}v'$) and $w = v \cdot v'$. In the rest frame of the B meson $\vec{v}' = -\vec{q}/m_{D^{(*)}}$.

From a separate calculation of the $B \to D^{(*)}$ from factors on the lattice with the same setup as we use for the inclusive decays, we numerically obtain the form factors of the form

$$h_X(w) = c_X^{(0)} + c_X^{(1)}(w-1) + c_X^{(2)}(w-1)^2 \tag{A.4}$$

after fitting the lattice data. The fit is shown in figure 27, and the numerical coefficients $c_X^{(0)}$, $c_X^{(1)}$, $c_X^{(2)}$ are listed in table 4.

Now, we insert the parametrizations given in eqs. (A.1), (A.2), and (A.3) into eq. (2.4) and perform the ω integral, which merely picks the ground state through $\delta(p_0 - q_0 - E_{D^{(*)}}) = \delta(\omega - E_{D^{(*)}})$. For $\bar{X} \equiv \sum_{l=0}^2 X^{(l)}$, we obtain

$$\bar{X}^{VV\|} = \frac{\boldsymbol{q}^2}{4m_D E_D}\left((m_B + m_D)h_+ - (m_B - m_D)h_-\right) \tag{A.5}$$

X	$c_X^{(0)}$	$c_X^{(1)}$	$c_X^{(2)}$
$+$	1.0082(26)	$-1.40(12)$	1.0(1.2)
$-$	$-0.057(11)$	$-0.01(17)$	
A_1	0.9143(34)	$-1.17(15)$	0.4(1.6)
A_2	$-0.354(75)$	0.5(1.2)	
A_3	0.999(75)	$-1.0(1.2)$	
V	1.243(13)	$-1.78(20)$	

Table 4. Numerical coefficients $c_X^{(i)}$ to parametrize the form factors of $B \to D$ ($X = +$ and $-$) and $B \to D^*$ (A_1, A_2, A_3 and V) decays. A polynomial expansion of the form $h_X(w) = c_X^{(0)} + c_X^{(1)}(w-1) + c_X^{(2)}(w-1)^2$ is introduced.

for the D meson contribution, which corresponds to the partial decay rate

$$
\Gamma^{B \to D} = \frac{G_F^2 |V_{cb}|^2}{8\pi^3} \int d\boldsymbol{q}^2 \frac{|\boldsymbol{q}|}{3} \cdot \frac{\boldsymbol{q}^2 (m_B + m_D)^2}{4 m_D E_D} \left[h_+ - \frac{m_B - m_D}{m_B + m_D} h_- \right]^2
$$
$$
= \frac{G_F^2 |V_{cb}|^2 m_B^5}{48\pi^3} \int dw \, (w^2 - 1)^{3/2} r^3 (1 + r)^2 \left[h_+ - \frac{1 - r}{1 + r} h_- \right]^2, \quad \text{(A.6)}
$$

where $w = v \cdot v' = \sqrt{1 + \boldsymbol{q}^2/m_D^2} = E_D/m_D$. The last line is a well-known formula for the $B \to D\ell\nu$ decay rate.

The vector meson D^* contributes in three channels: $AA_\parallel$, $AA_\perp$, $VV_\perp$. The contributions are

$$
\bar{X}^{AA\parallel} = \frac{1}{4 m_{D^*} E_{D^*}} \Big[(m_B - E_{D^*}) E_{D^*} h_{A1}(1 + w)
$$
$$
+ \boldsymbol{q}^2 \left(h_{A1}(1 + w) - h_{A2} - \frac{m_B}{m_{D^*}} h_{A3} \right) \Big]^2, \quad \text{(A.7)}
$$

$$
\bar{X}^{AA\perp} = \left[(m_B - m_{D^*})^2 - 2 m_B (E_{D^*} - m_{D^*}) \right] \frac{(1 + w)^2}{2w} h_{A1}^2, \quad \text{(A.8)}
$$

$$
\bar{X}^{VV\perp} = \left[(m_B - m_{D^*})^2 - 2 m_B (E_{D^*} - m_{D^*}) \right] \frac{\boldsymbol{q}^2}{2 m_{D^*} E_{D^*}} h_V^2. \quad \text{(A.9)}
$$

Adding them together, we obtain

$$
\Gamma^{B \to D^*} = \frac{G_F^2 |V_{cb}|^2}{8\pi^3} \int d\boldsymbol{q}^2 \frac{|\boldsymbol{q}|}{3} \left\{ (q_0^2 - \boldsymbol{q}^2) \left[\frac{(1 + w)^2}{2w} h_{A1}^2 + \frac{\boldsymbol{q}^2}{2 m_{D^*} E_{D^*}} h_V^2 \right] \right.
$$
$$
+ \frac{1}{4 m_{D^*} E_{D^*}} \Big[(m_B - E_{D^*}) E_{D^*} h_{A1}(1 + w)
$$
$$
\left. + \boldsymbol{q}^2 \left(h_{A1}(1 + w) - h_{A2} - \frac{m_B}{m_{D^*}} h_{A3} \right) \Big]^2 \right\}, \quad \text{(A.10)}
$$

where $r = m_{D^*}/m_B$ and $E_{D^*} = m_{D^*} w = m_B r w$, while $\boldsymbol{q}^2 = m_{D^*}^2 (w^2 - 1) = m_B^2 r^2 (w^2 - 1)$, and $q_0^2 - \boldsymbol{q}^2 = (m_B - m_{D^*})^2 - 2 m_B (E_{D^*} - m_{D^*}) = m_B[(1 - r)^2 - 2r(w - 2)]$. Eq. (A.10)

can then be rewritten as

$$\Gamma^{B \to D^*} = \frac{G_F^2 |V_{cb}|^2 m_B^5}{48\pi^3} \int dw \, (w^2 - 1)^{1/2} r^3 (1-r)^2 (w+1)^2 |h_{A1}|^2$$
$$\times \left\{ 2 \frac{r^2 - 2rw + 1}{(1-r)^2} \left[1 + \frac{w-1}{w+1} R_1^2 \right] + \left[1 + \frac{w-1}{1-r}(1 - R_2) \right]^2 \right\}, \qquad \text{(A.11)}$$

with $R_1 \equiv h_V/h_{A1}$ and $R_2 \equiv (h_{A3} + r h_{A2})/h_{A1}$, which confirms a well-known formula.

From this analysis, the contributions of the S-wave ground states, D and D^*, to the integrands $\bar{X}^{VV\parallel}$, $\bar{X}^{VV\perp}$, $\bar{X}^{AA\perp}$, and $\bar{X}^{AA\parallel}$ can be identified.

The contribution of the VA and AV insertions vanishes for the total decay rate as well as for the hadronic mass moments, but it is non-zero for the lepton energy moments. In the SM the contribution of the AV interference from the ground state $B \to D^*$ to the first leptonic moment can be written as

$$\bar{X}^{AV}_{n_\ell=1} = - \left[(m_B - E_{D^*})^2 - \boldsymbol{q}^2 \right] \frac{\boldsymbol{q}^2}{4 E_{D^*}} (1+w) h_V h_{A1} . \qquad \text{(A.12)}$$

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](#)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited. SCOAP³ supports the goals of the International Year of Basic Sciences for Sustainable Development.

References

[1] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531 [INSPIRE].

[2] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652 [INSPIRE].

[3] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01 [INSPIRE].

[4] HFLAV collaboration, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*, *Eur. Phys. J. C* **81** (2021) 226 [arXiv:1909.12524] [INSPIRE].

[5] P. Gambino, M. Jung and S. Schacht, *The V_{cb} puzzle: An update*, *Phys. Lett. B* **795** (2019) 386 [arXiv:1905.08209] [INSPIRE].

[6] P. Gambino et al., *Challenges in semileptonic B decays*, *Eur. Phys. J. C* **80** (2020) 966 [arXiv:2006.07287] [INSPIRE].

[7] K.G. Wilson, *Confinement of Quarks*, *Phys. Rev. D* **10** (1974) 2445 [INSPIRE].

[8] FLAVOUR LATTICE AVERAGING GROUP collaboration, *FLAG Review 2019: Flavour Lattice Averaging Group (FLAG)*, *Eur. Phys. J. C* **80** (2020) 113 [arXiv:1902.08191] [INSPIRE].

[9] S. Hashimoto, *Inclusive semi-leptonic B meson decay structure functions from lattice QCD*, *PTEP* **2017** (2017) 053B03 [arXiv:1703.01881] [INSPIRE].

[10] A.V. Manohar and M.B. Wise, *Inclusive semileptonic B and polarized Λ_b decays from QCD*, *Phys. Rev. D* **49** (1994) 1310 [hep-ph/9308246] [INSPIRE].

[11] B. Blok, L. Koyrakh, M.A. Shifman and A.I. Vainshtein, *Differential distributions in semileptonic decays of the heavy flavors in QCD*, *Phys. Rev. D* **49** (1994) 3356 [*Erratum ibid.* **50** (1994) 3572] [hep-ph/9307247] [ɪɴSPIRE].

[12] M.T. Hansen, H.B. Meyer and D. Robaina, *From deep inelastic scattering to heavy-flavor semileptonic decays: Total rates into multihadron final states from lattice QCD*, *Phys. Rev. D* **96** (2017) 094513 [arXiv:1704.08993] [ɪɴSPIRE].

[13] M. Hansen, A. Lupo and N. Tantalo, *Extraction of spectral densities from lattice correlators*, *Phys. Rev. D* **99** (2019) 094508 [arXiv:1903.06476] [ɪɴSPIRE].

[14] J. Bulava and M.T. Hansen, *Scattering amplitudes from finite-volume spectral functions*, *Phys. Rev. D* **100** (2019) 034521 [arXiv:1903.11735] [ɪɴSPIRE].

[15] J. Bulava, M.T. Hansen, M.W. Hansen, A. Patella and N. Tantalo, *Inclusive rates from smeared spectral densities in the two-dimensional O(3) non-linear σ-model*, *JHEP* **07** (2022) 034 [arXiv:2111.12774] [ɪɴSPIRE].

[16] P. Gambino and S. Hashimoto, *Inclusive Semileptonic Decays from Lattice QCD*, *Phys. Rev. Lett.* **125** (2020) 032001 [arXiv:2005.13730] [ɪɴSPIRE].

[17] K.G. Wilson, *Nonlagrangian models of current algebra*, *Phys. Rev.* **179** (1969) 1499 [ɪɴSPIRE].

[18] L.P. Kadanoff, *Operator Algebra and the Determination of Critical Indices*, *Phys. Rev. Lett.* **23** (1969) 1430 [ɪɴSPIRE].

[19] I.I.Y. Bigi, N.G. Uraltsev and A.I. Vainshtein, *Nonperturbative corrections to inclusive beauty and charm decays: QCD versus phenomenological models*, *Phys. Lett. B* **293** (1992) 430 [*Erratum ibid.* **297** (1992) 477] [hep-ph/9207214] [ɪɴSPIRE].

[20] I.I.Y. Bigi, M.A. Shifman, N.G. Uraltsev and A.I. Vainshtein, *QCD predictions for lepton spectra in inclusive heavy flavor decays*, *Phys. Rev. Lett.* **71** (1993) 496 [hep-ph/9304225] [ɪɴSPIRE].

[21] G. Bailas, S. Hashimoto and T. Ishikawa, *Reconstruction of smeared spectral function from Euclidean correlation functions*, *PTEP* **2020** (2020) 043B07 [arXiv:2001.11779] [ɪɴSPIRE].

[22] JLQCD collaboration, *Form factors of $B \to \pi\ell\nu$ and a determination of $|V_{ub}|$ with Möbius domain-wall-fermions*, arXiv:2203.04938 [ɪɴSPIRE].

[23] JLQCD collaboration, *Determination of charm quark mass from temporal moments of charmonium correlator with Mobius domain-wall fermion*, *PoS* **LATTICE2016** (2017) 192 [arXiv:1702.01498] [ɪɴSPIRE].

[24] A.K. Leibovich, Z. Ligeti, I.W. Stewart and M.B. Wise, *Semileptonic B decays to excited charmed mesons*, *Phys. Rev. D* **57** (1998) 308 [hep-ph/9705467] [ɪɴSPIRE].

[25] R. Baron et al., *Light hadrons from lattice QCD with light (u,d), strange and charm dynamical quarks*, *JHEP* **06** (2010) 111 [arXiv:1004.5284] [ɪɴSPIRE].

[26] ETM collaboration, *Light hadrons from Nf=2+1+1 dynamical twisted mass fermions*, *PoS* **LATTICE2010** (2010) 123 [arXiv:1101.0518] [ɪɴSPIRE].

[27] Y. Iwasaki, *Renormalization Group Analysis of Lattice Theories and Improved Lattice Action: Two-Dimensional Nonlinear O(N) Sigma Model*, *Nucl. Phys. B* **258** (1985) 141 [ɪɴSPIRE].

[28] ALPHA collaboration, *Lattice QCD with a chirally twisted mass term*, *JHEP* **08** (2001) 058 [hep-lat/0101001] [ɪɴSPIRE].

[29] R. Frezzotti and G.C. Rossi, *Twisted mass lattice QCD with mass nondegenerate quarks*, *Nucl. Phys. B Proc. Suppl.* **128** (2004) 193 [hep-lat/0311008] [INSPIRE].

[30] R. Frezzotti and G.C. Rossi, *Chirally improving Wilson fermions. 1. $O(a)$ improvement*, *JHEP* **08** (2004) 007 [hep-lat/0306014] [INSPIRE].

[31] EUROPEAN TWISTED MASS collaboration, *Up, down, strange and charm quark masses with $N_f = 2+1+1$ twisted mass lattice QCD*, *Nucl. Phys. B* **887** (2014) 19 [arXiv:1403.4504] [INSPIRE].

[32] R. Frezzotti and G.C. Rossi, *Chirally improving Wilson fermions. II. Four-quark operators*, *JHEP* **10** (2004) 070 [hep-lat/0407002] [INSPIRE].

[33] K. Osterwalder and E. Seiler, *Gauge Field Theories on the Lattice*, *Annals Phys.* **110** (1978) 440 [INSPIRE].

[34] R. Frezzotti, G. Martinelli, M. Papinutto and G.C. Rossi, *Reducing cutoff effects in maximally twisted lattice QCD close to the chiral limit*, *JHEP* **04** (2006) 038 [hep-lat/0503034] [INSPIRE].

[35] P. Dimopoulos, R. Frezzotti, C. Michael, G.C. Rossi and C. Urbach, *$O(a^2)$ cutoff effects in lattice Wilson fermion simulations*, *Phys. Rev. D* **81** (2010) 034509 [arXiv:0908.0451] [INSPIRE].

[36] UKQCD collaboration, *Quark mass dependence of hadron masses from lattice QCD*, *Phys. Rev. D* **59** (1999) 074503 [hep-lat/9810021] [INSPIRE].

[37] UKQCD collaboration, *Decay width of light quark hybrid meson from the lattice*, *Phys. Rev. D* **73** (2006) 074506 [hep-lat/0603007] [INSPIRE].

[38] S. Gusken, *A Study of smearing techniques for hadron correlation functions*, *Nucl. Phys. B Proc. Suppl.* **17** (1990) 361 [INSPIRE].

[39] APE collaboration, *Glueball Masses and String Tension in Lattice QCD*, *Phys. Lett. B* **192** (1987) 163 [INSPIRE].

[40] P.F. Bedaque, *Aharonov-Bohm effect and nucleon nucleon phase shifts on the lattice*, *Phys. Lett. B* **593** (2004) 82 [nucl-th/0402051] [INSPIRE].

[41] G.M. de Divitiis, R. Petronzio and N. Tantalo, *On the discretization of physical momenta in lattice QCD*, *Phys. Lett. B* **595** (2004) 408 [hep-lat/0405002] [INSPIRE].

[42] D. Guadagnoli, F. Mescia and S. Simula, *Lattice study of semileptonic form-factors with twisted boundary conditions*, *Phys. Rev. D* **73** (2006) 114504 [hep-lat/0512020] [INSPIRE].

[43] C.T. Sachrajda and G. Villadoro, *Twisted boundary conditions in lattice simulations*, *Phys. Lett. B* **609** (2005) 73 [hep-lat/0411033] [INSPIRE].

[44] P.F. Bedaque and J.-W. Chen, *Twisted valence quarks and hadron interactions on the lattice*, *Phys. Lett. B* **616** (2005) 208 [hep-lat/0412023] [INSPIRE].

[45] J. Chay, H. Georgi and B. Grinstein, *Lepton energy distributions in heavy meson decays from QCD*, *Phys. Lett. B* **247** (1990) 399 [INSPIRE].

[46] I.I.Y. Bigi, M.A. Shifman, N. Uraltsev and A.I. Vainshtein, *High power n of m_b in beauty widths and $n = 5 \to \infty$ limit*, *Phys. Rev. D* **56** (1997) 4017 [hep-ph/9704245] [INSPIRE].

[47] A. Czarnecki, K. Melnikov and N. Uraltsev, *NonAbelian dipole radiation and the heavy quark expansion*, *Phys. Rev. Lett.* **80** (1998) 3189 [hep-ph/9708372] [INSPIRE].

[48] M. Fael, K. Schönwald and M. Steinhauser, *Relation between the* $\overline{\text{MS}}$ *and the kinetic mass of heavy quarks*, *Phys. Rev. D* **103** (2021) 014005 [arXiv:2011.11655] [INSPIRE].

[49] P. Gambino and N. Uraltsev, *Moments of semileptonic B decay distributions in the* $1/m_b$ *expansion*, *Eur. Phys. J. C* **34** (2004) 181 [hep-ph/0401063] [INSPIRE].

[50] V. Aquila, P. Gambino, G. Ridolfi and N. Uraltsev, *Perturbative corrections to semileptonic b decay distributions*, *Nucl. Phys. B* **719** (2005) 77 [hep-ph/0503083] [INSPIRE].

[51] A. Alberti, T. Ewerth, P. Gambino and S. Nandi, *Kinetic operator effects in* $\bar{B} \to X_c l \nu$ *at* $O(\alpha_s)$, *Nucl. Phys. B* **870** (2013) 16 [arXiv:1212.5082] [INSPIRE].

[52] M. Gremm and A. Kapustin, *Order* $1/m_b^3$ *corrections to* $B \to X_c l \bar{\nu}$ *decay and their implication for the measurement of* $\bar{\Lambda}$ *and* λ_1, *Phys. Rev. D* **55** (1997) 6924 [hep-ph/9603448] [INSPIRE].

[53] T. Mannel, S. Turczyk and N. Uraltsev, *Higher Order Power Corrections in Inclusive B Decays*, *JHEP* **11** (2010) 109 [arXiv:1009.4622] [INSPIRE].

[54] P. Gambino, K.J. Healey and S. Turczyk, *Taming the higher power corrections in semileptonic B decays*, *Phys. Lett. B* **763** (2016) 60 [arXiv:1606.06174] [INSPIRE].

[55] A. Alberti, P. Gambino and S. Nandi, *Perturbative corrections to power suppressed effects in semileptonic B decays*, *JHEP* **01** (2014) 147 [arXiv:1311.7381] [INSPIRE].

[56] T. Mannel, A.A. Pivovarov and D. Rosenthal, *Inclusive weak decays of heavy hadrons with power suppressed terms at NLO*, *Phys. Rev. D* **92** (2015) 054025 [arXiv:1506.08167] [INSPIRE].

[57] T. Mannel and A.A. Pivovarov, *QCD corrections to inclusive heavy hadron weak decays at* $\Lambda_{\text{QCD}}^3/m_Q^3$, *Phys. Rev. D* **100** (2019) 093001 [arXiv:1907.09187] [INSPIRE].

[58] A. Pak and A. Czarnecki, *Heavy-to-heavy quark decays at NNLO*, *Phys. Rev. D* **78** (2008) 114015 [arXiv:0808.3509] [INSPIRE].

[59] A. Pak and A. Czarnecki, *Mass effects in muon and semileptonic* $b \to c$ *decays*, *Phys. Rev. Lett.* **100** (2008) 241807 [arXiv:0803.0960] [INSPIRE].

[60] K. Melnikov, $O(\alpha_s^2)$ *corrections to semileptonic decay* $b \to c l \bar{\nu}_l$, *Phys. Lett. B* **666** (2008) 336 [arXiv:0803.0951] [INSPIRE].

[61] S. Biswas and K. Melnikov, *Second order QCD corrections to inclusive semileptonic* $b \to X_c l \bar{\nu}_l$ *decays with massless and massive lepton*, *JHEP* **02** (2010) 089 [arXiv:0911.4142] [INSPIRE].

[62] M. Fael, K. Schönwald and M. Steinhauser, *Third order corrections to the semileptonic* $b \to c$ *and the muon decays*, *Phys. Rev. D* **104** (2021) 016003 [arXiv:2011.13654] [INSPIRE].

[63] A. Alberti, *Higher order corrections to inclusive semileptonic B decays*, Ph.D. Thesis, Turin University, Turin, Italy (2015).

[64] P. Colangelo, F. De Fazio and F. Loparco, *Inclusive semileptonic* Λ_b *decays in the Standard Model and beyond*, *JHEP* **11** (2020) 032 [arXiv:2006.13759] [INSPIRE].

[65] P. Gambino, A. Melis and S. Simula, *Extraction of heavy-quark-expansion parameters from unquenched lattice data on pseudoscalar and vector heavy-light meson masses*, *Phys. Rev. D* **96** (2017) 014511 [arXiv:1704.06105] [INSPIRE].

[66] P. Gambino, V. Lubicz, A. Melis and S. Simula, *Masses, decay constants and HQE matrix elements of pseudoscalar and vector heavy-light mesons in LQCD*, *J. Phys. Conf. Ser.* **1137** (2019) 012005 [INSPIRE].

[67] F. Herren and M. Steinhauser, *Version 3 of RunDec and CRunDec*, *Comput. Phys. Commun.* **224** (2018) 333 [arXiv:1703.03751] [InSPIRE].

[68] M. Bordone, B. Capdevila and P. Gambino, *Three loop calculations and inclusive V_{cb}*, *Phys. Lett. B* **822** (2021) 136679 [arXiv:2107.00604] [InSPIRE].

[69] I.I. Bigi, T. Mannel and N. Uraltsev, *Semileptonic width ratios among beauty hadrons*, *JHEP* **09** (2011) 012 [arXiv:1105.4574] [InSPIRE].

[70] M. Bordone and P. Gambino, *The semileptonic B_s and Λ_b widths*, in *11th International Workshop on the CKM Unitarity Triangle*, (2022) [arXiv:2203.13107] [InSPIRE].

[71] ETM collaboration, *A Proposal for B-physics on current lattices*, *JHEP* **04** (2010) 049 [arXiv:0909.3187] [InSPIRE].

[72] ETM collaboration, *Lattice QCD determination of m_b, f_B and f_{Bs} with twisted mass Wilson fermions*, *JHEP* **01** (2012) 046 [arXiv:1107.1441] [InSPIRE].

[73] ETM collaboration, *B-physics from $N_f = 2$ tmQCD: the Standard Model and beyond*, *JHEP* **03** (2014) 016 [arXiv:1308.1851] [InSPIRE].

[74] ETM collaboration, *Mass of the b quark and B-meson decay constants from $N_f = 2 + 1 + 1$ twisted-mass lattice QCD*, *Phys. Rev. D* **93** (2016) 114505 [arXiv:1603.04306] [InSPIRE].

[75] CLEO collaboration, *Measurement of absolute branching fractions of inclusive semileptonic decays of charm and charmed-strange mesons*, *Phys. Rev. D* **81** (2010) 052007 [arXiv:0912.4232] [InSPIRE].

[76] BESIII collaboration, *Measurement of the absolute branching fraction of inclusive semielectronic D_s^+ decays*, *Phys. Rev. D* **104** (2021) 012003 [arXiv:2104.07311] [InSPIRE].