

PREPARED FOR SUBMISSION TO JHEP

Approaches to inclusive semileptonic $B_{(s)}$ -meson decays from Lattice QCD

Alessandro Barone,^{a,b,c} Shoji Hashimoto,^{c,d} Andreas Jüttner,^{a,b,e} Takashi Kaneko,^{c,d,f}
Ryan Kellermann^{c,d}

^a*School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*

^b*STAG Research Center, University of Southampton, Southampton SO17 1BJ, UK*

^c*High Energy Accelerator Research Organization (KEK), Ibaraki 305-0801, Japan*

^d*School of High Energy Accelerator Science, SOKENDAI (The Graduate University for Advanced Studies), Ibaraki 305-0801, Japan*

^e*CERN, Theoretical Physics Department, Geneva, Switzerland*

^f*Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University, Aichi 464-8602, Japan*

E-mail: a.barone@soton.ac.uk, shoji.hashimoto@kek.jp,
andreas.juttner@cern.ch, takashi.kaneko@kek.jp, kelry@post.kek.jp

ABSTRACT: We address the nonperturbative calculation of the inclusive decay rate of semileptonic $B_{(s)}$ -meson decays from lattice QCD. Precise Standard-Model predictions are key ingredients in searches for new physics, and this type of computation may eventually provide new insight into the long-standing tension between the inclusive and exclusive determinations of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements $|V_{cb}|$ and $|V_{ub}|$. We present results from a pilot lattice computation for $B_s \rightarrow X_c l \nu_l$, where the initial b quark described by the relativistic-heavy-quark (RHQ) formalism on the lattice and the other valence quarks discretised with domain-wall fermions are simulated approximately at their physical quark masses. We compare two different methods for computing the decay rate from lattice data of Euclidean n -point functions, namely Chebyshev and Backus-Gilbert approaches. We further study how much the ground-state meson dominates the inclusive decay rate and indicate our strategy towards a computation with a more comprehensive systematic error budget.

ARXIV EPRINT: [2305.14092](https://arxiv.org/abs/2305.14092)

Contents

1	Introduction	2
2	Theoretical framework	3
2.1	The inclusive decay rate	3
2.2	Ground-state limit	5
2.3	Inclusive decays on an Euclidean space-time lattice	6
2.4	Data analysis	9
2.4.1	Chebyshev-polynomial approximation	10
2.4.2	Backus-Gilbert	12
3	Numerical setup	14
4	Results	17
4.1	Polynomial approximation of the kernel	17
4.2	Chebyshev polynomials and Backus-Gilbert in practice	20
4.3	The inclusive decay rate	21
4.4	The inclusive decay rate in the ground-state limit	27
5	Conclusions and outlook	29
A	Chebyshev polynomials	30
A.1	Standard polynomials	30
A.1.1	Expansion in Chebyshev polynomials	31
A.2	Shifted Chebyshev polynomials	31
A.2.1	Expansion in Chebyshev polynomials with exponential map	32
A.2.2	Matrix relations	33
B	Generalised Backus-Gilbert	34
B.1	The method	34
B.2	A different perspective	37
C	Fit strategy	38
C.1	MAP with bounds	38
C.2	MAP with bootstrap	38

1 Introduction

The study of the b -quark sector of particle physics remains an exciting arena of precision physics, in which intriguing tensions between observations and Standard-Model (SM) predictions have been found [1–6]. Scrutinising these findings and better controlling and reducing experimental and theoretical error budgets therefore remain a crucial task. Any such anomaly could be an indicator of new effects: while new particles may be too heavy to be produced with energies achievable by current experimental facilities, quantum effects could leave detectable traces in flavour-physics processes. One of these long-standing tensions involves the measured values of the CKM matrix elements $|V_{cb}|$ and $|V_{ub}|$ between exclusive and inclusive decays. Apart from leptonic decays, these can be determined through the exclusive semileptonic decay of a B into a $D^{(*)}$ (or π), or through the measurement of the inclusive decay rate, respectively. For example, one of the most recent determination of $|V_{cb}|$ finds

$$\begin{aligned} |V_{cb}| &= (42.19 \pm 0.78) \times 10^{-3} \quad \text{inclusive [7, 8]}, \\ |V_{cb}| &= (39.36 \pm 0.68) \times 10^{-3} \quad \text{exclusive [9–13]}. \end{aligned}$$

Lattice computations provide crucial nonperturbative input to the exclusive determination and the required techniques in this case are well established (see reviews [13, 14]). The existing results for the inclusive decay are based on perturbative QCD. First viable theoretical proposals for how to accomplish the computation of the inclusive decay rate on the lattice have appeared only recently [15]. The idea relies on the extraction of a forward-scattering matrix element through analytic continuation of lattice results obtained in an unphysical kinematical region. In [16] it was then proposed to address decay and transition rates of multi-hadron processes through finite-volume Euclidean four-point functions provided that a method to extract the associated spectral function exists.

In this paper, we present work towards an improved understanding of the calculation of the inclusive decay rate by means of a pilot study of semileptonic decays of B_s mesons into charmed particles, namely $B_s \rightarrow X_c l \nu_l$, following [17], where the extraction of the spectral function is bypassed and the decay rate is evaluated directly. Preliminary work has been presented in [18, 19]. In particular, we improve and compare two existing methods, namely Chebyshev [17, 20, 21] and Backus-Gilbert [22, 23] reconstructions. Our work uses the relativistic-heavy-quark action (RHQ) [24–26] to simulate the bottom-valence quark at its physical mass, while the strange- and charm-valence quarks are treated with a domain-wall fermion action [27–30], and their masses are tuned to values close to the ones found in nature.

The structure of this paper is as follows: in Sec. 2 we describe the theoretical framework, extending the formalism introduced in [17]. We also address the ground-state limit and its connection with the corresponding exclusive processes. In Sec. 2.3 we describe some details of the lattice implementation. In Sec. 2.4 we report on our analysis strategies; to keep the discussion fluent we refer to App. A, B and C for technical details. Finally, we discuss the details of the simulation in Sec. 3 and present our results in Sec. 4. We summarise our findings and discuss future prospects in Sec. 5.

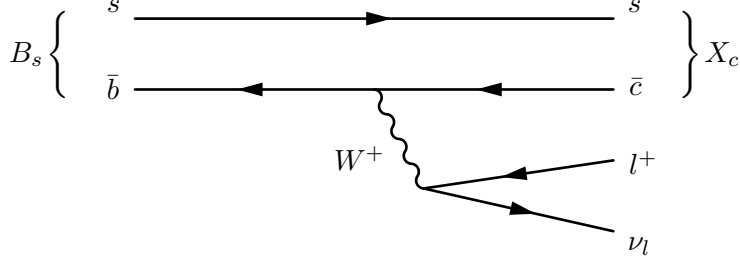


Figure 1. Feynman diagram for $B_s \rightarrow X_c l \nu_l$.

2 Theoretical framework

2.1 The inclusive decay rate

We start by reviewing the formalism to calculate the decay rate of inclusive semileptonic processes [31, 32]. Here, we focus on the decay $B_s \rightarrow X_c l \nu_l$ illustrated in Fig. 1, but the formalism is more generally applicable to other channels such as, e.g., $B \rightarrow X l \nu_l$ or $D_{(s)} \rightarrow X l \nu_l$. The final state X_c represents all possible charmed-meson final states allowed by flavour, spin and parity quantum numbers. The ground-state contribution to X_c in the vector channel is given by the D_s meson. The leading order weak Hamiltonian for the $\bar{b} \rightarrow \bar{c}$ process is given by

$$H_W = \frac{4G_F}{\sqrt{2}} V_{cb} [\bar{b}_L \gamma^\mu c_L] [\bar{\nu}_l \gamma_\mu l_L], \quad (2.1)$$

where G_F is the Fermi constant and V_{cb} is the CKM matrix element for the charged-current flavour-changing quark transition. The electroweak quark current for this process is then $J_\mu = \bar{b}_L \gamma^\mu c_L = \bar{b} \gamma_\mu (1 - \gamma_5) c$, which we can also write as $J_\mu = V_\mu - A_\mu$ with $V_\mu = \bar{b} \gamma_\mu c$ and $A_\mu = \bar{b} \gamma_\mu \gamma_5 c$.

The differential decay rate for the inclusive process depends on three kinematical variables, i.e. one more than the corresponding exclusive decay due to the freedom in the mass of the outgoing hadrons. Neglecting QED corrections it reads

$$\frac{d\Gamma}{dq^2 dq_0 dE_l} = \frac{G_F^2 |V_{cb}|^2}{8\pi^3} L_{\mu\nu} W^{\mu\nu}. \quad (2.2)$$

The lepton contribution is given in terms of the leptonic tensor

$$L^{\mu\nu} = p_l^\mu p_{\nu_l}^\nu + p_l^\nu p_{\nu_l}^\mu - g^{\mu\nu} p_l \cdot p_{\nu_l} - i\epsilon^{\mu\alpha\nu\beta} p_{l,\alpha} p_{\nu_l,\beta}, \quad (2.3)$$

where p_l and p_{ν_l} are the four-momenta of the lepton and the neutrino, respectively. The hadronic tensor $W^{\mu\nu}$ is defined as

$$\begin{aligned} W^{\mu\nu}(p_{B_s}, q) &= \frac{1}{2E_{B_s}} \int d^4x e^{iq \cdot x} \langle B_s(\mathbf{p}_{B_s}) | J^{\mu\dagger}(x) J^\nu(0) | B_s(\mathbf{p}_{B_s}) \rangle \\ &= \frac{1}{2E_{B_s}} \sum_{X_c} (2\pi)^3 \delta^{(4)}(p_{B_s} - q - p_{X_c}) \\ &\quad \times \langle B_s(\mathbf{p}_{B_s}) | J^{\mu\dagger}(0) | X_c(\mathbf{p}_{X_c}) \rangle \langle X_c(\mathbf{p}_{X_c}) | J^\nu(0) | B_s(\mathbf{p}_{B_s}) \rangle, \end{aligned} \quad (2.4)$$

where in the second line we have inserted the sum $\sum_{X_c} |X_c(\mathbf{p}_{X_c})\rangle \langle X_c(\mathbf{p}_{X_c})|$ over a complete set of states, which is understood to include an integration over all possible momenta \mathbf{p}_{X_c} under a Lorentz invariant phase-space integral, and $q = p_{B_s} - p_{X_c} = p_l + p_{\nu_l}$ is the transferred momentum between the initial and final hadronic states. Note that we will consider only the case of the B_s meson at rest, i.e. $\mathbf{p}_{B_s} = (0, 0, 0)$, and will henceforth suppress the corresponding momentum label. The hadronic tensor can be decomposed into five scalar structure functions $W_i \equiv W_i(q^2, v \cdot q)$ as

$$W^{\mu\nu} = -g^{\mu\nu}W_1 + v^\mu v^\nu W_2 - i\epsilon^{\mu\nu\alpha\beta} v_\alpha q_\beta W_3 + q^\mu q^\nu W_4 + (v^\mu q^\nu + v^\nu q^\mu)W_5, \quad (2.5)$$

where $v = p_{B_s}/M_{B_s} = (1, 0, 0, 0)$ is the velocity of the initial B_s meson at rest, and $q = (q_0, \mathbf{q}) = (M_{B_s} - \omega, -\mathbf{p}_{X_c})$. From now on, we will indicate with $\omega = E_{X_c}$ the energy of the final-state hadron. The individual components of the hadronic tensor can be expressed conveniently in terms of the structure functions,

$$W_{00} = -W_1 + W_2 + q_0^2 W_4 + 2q_0 W_5, \quad (2.6)$$

$$W_{ij} = \delta_{ij}W_1 + q_i q_j W_4 - i\epsilon_{ij0k} q^k W_3, \quad (2.7)$$

$$W_{0i} = W_{i0} = q_i(q_0 W_4 + W_5), \quad (2.8)$$

where i, j, k refers to the spatial indices 1, 2, 3. We note that contracting the spatial indices with the three-momentum components q_i , we can invert these relations and find expressions for the structure functions in terms of the hadronic tensor and \mathbf{q} .

Integrating over the lepton energy $E_l = p_{l,0}$ and assuming $m_l = 0$ we obtain the expression for the decay rate

$$\Gamma = \frac{G_F^2 |V_{cb}|^2}{24\pi^3} \int_0^{q_{\max}^2} d\mathbf{q}^2 \sqrt{\mathbf{q}^2} \bar{X}(\mathbf{q}^2), \quad (2.9)$$

where the integration over ω is contained in

$$\bar{X}(\mathbf{q}^2) = \sum_{l=0}^2 \bar{X}^{(l)}(\mathbf{q}^2), \quad \bar{X}^{(l)}(\mathbf{q}^2) \equiv \int_{\omega_{\min}}^{\omega_{\max}} d\omega X^{(l)}(\mathbf{q}^2), \quad (2.10)$$

and where we defined

$$\begin{aligned} X^{(0)}(\mathbf{q}^2) &= \mathbf{q}^2 W_{00} + \sum_i (q_i^2 - \mathbf{q}^2) W_{ii} + \sum_{i \neq j} q^i W_{ij} q^j, \\ X^{(1)}(\mathbf{q}^2) &= -q_0 \sum_i q^i (W_{0i} + W_{i0}), \\ X^{(2)}(\mathbf{q}^2) &= q_0^2 \sum_i W_{ii}. \end{aligned} \quad (2.11)$$

Recalling that the D_s meson is the lightest final state in this inclusive decay process, and imposing four-momentum conservation we obtain $q_{\max}^2 = (M_{B_s}^2 - M_{D_s}^2)^2 / (4M_{B_s}^2)$, $\omega_{\min} = \sqrt{M_{D_s}^2 + \mathbf{q}^2}$ and $\omega_{\max} = M_{B_s} - \sqrt{\mathbf{q}^2}$ for the integral limits. $X^{(l)}$ and $\bar{X}^{(l)}$ depend only on \mathbf{q}^2 and not on individual components of \mathbf{q} , as can be seen after substituting Eqs. (2.6), (2.7) and (2.8) into the expressions (2.11).

Starting from the decomposition of the hadronic tensor $W^{\mu\nu} = W_{VV}^{\mu\nu} + W_{AA}^{\mu\nu} - W_{VA}^{\mu\nu} - W_{AV}^{\mu\nu}$, the $X^{(l)}$ can also be rewritten in a way that exposes the $V - A$ nature of the charged current, namely

$$X^{(l)} = X_{VV}^{(l)} + X_{AA}^{(l)} - X_{VA}^{(l)} - X_{AV}^{(l)}, \quad (2.12)$$

and similarly for $\bar{X}^{(l)}$.

2.2 Ground-state limit

In this section we consider a hypothetical world in which only the lowest-mass final state D_s contributes to the inclusive decay, i.e.,

$$W_{\mu\nu} \rightarrow \delta(\omega - E_{D_s}) \frac{1}{4E_{B_s}E_{D_s}} \langle B_s(\mathbf{p}_{B_s}) | V_\mu^\dagger | D_s(\mathbf{p}_{D_s}) \rangle \langle D_s(\mathbf{p}_{D_s}) | V_\nu | B_s(\mathbf{p}_{B_s}) \rangle. \quad (2.13)$$

In this limit we can reconstruct the inclusive decay rate from lattice simulations of the exclusive decay, allowing us to compute the ground-state contribution. We will also use results in this limit to devise consistency checks of the inclusive-decay setup. The required hadronic form factors $f_+(q^2)$ and $f_-(q^2)$ parametrising the corresponding matrix element

$$\langle D_s(\mathbf{p}_{D_s}) | V_\mu | B_s(\mathbf{p}_{B_s}) \rangle = f_+(q^2)(p_{B_s} + p_{D_s})_\mu + f_-(q^2)(p_{B_s} - p_{D_s})_\mu, \quad (2.14)$$

of the exclusive decay $B_s \rightarrow D_s l \nu_l$ can be computed separately on the lattice using more conventional methods [33–35].

In order to compute the inclusive decay rate in this limit we now establish the relation between the vector form factor $f_+(q^2)$ and $\bar{X}_{VV} = \sum_{l=0}^2 \bar{X}_{VV}^{(l)}$ defined in Eqs. (2.10) and (2.11) using the decomposition in Eq. (2.12). Let us first decompose $\bar{X}_{VV} = \bar{X}_{VV}^\parallel + \bar{X}_{VV}^\perp$ into longitudinal and transverse components in terms of the projectors $\Pi_{\mu\nu}^\perp = g^{\mu\nu} - q^\mu q^\nu / q^2$ and $\Pi_{\mu\nu}^\parallel = q^\mu q^\nu / q^2$, where

$$\begin{aligned} X_{VV}^\parallel &= q^2 W_1 + \mathbf{q}^2 W_2, \\ X_{VV}^\perp &= 2q^2 W_1, \end{aligned} \quad (2.15)$$

which, inverting Eq. (2.6)-(2.8) and considering $\mathbf{q}^2 \neq 0$, can be expanded as

$$X_{VV}^\parallel = \mathbf{q}^2 W_{VV}^{00} - q_0 \sum_i q_i (W_{VV}^{0i} + W_{VV}^{i0}) + \frac{q_0^2}{\mathbf{q}^2} \sum_{i,j} q_i W_{VV}^{ij} q_j, \quad (2.16)$$

$$\begin{aligned} X_{VV}^\perp &= (q_0^2 - \mathbf{q}^2) \sum_i W_{VV}^{ii} + \sum_{i,j} \left(1 - \frac{q_0^2}{\mathbf{q}^2}\right) q_i W_{VV}^{ij} q_j \\ &= \sum_i \left(1 - \frac{q_0^2}{\mathbf{q}^2}\right) (q_i^2 - \mathbf{q}^2) W_{VV}^{ii} + \sum_{i \neq j} \left(1 - \frac{q_0^2}{\mathbf{q}^2}\right) q_i W_{VV}^{ij} q_j. \end{aligned} \quad (2.17)$$

Inserting the expression Eq. (2.13) into Eq. (2.16), we obtain

$$\bar{X}_{VV}^\parallel = \frac{M_{B_s}}{E_{D_s}} \mathbf{q}^2 |f_+(q^2)|^2. \quad (2.18)$$

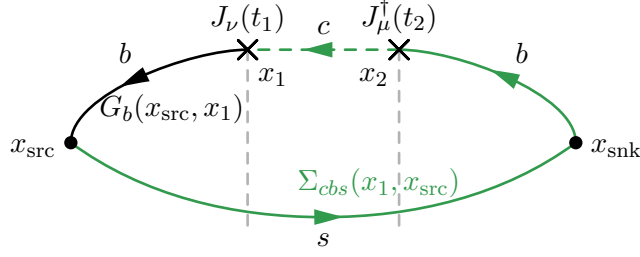


Figure 2. Diagram of the four-point correlator. Two propagators used for the contraction are depicted in the picture. The black one, $G_b(x_{\text{src}}, x_1)$, is a propagator for the b quark from x_1 to x_{src} . The green one, $\Sigma_{cbs}(x_1, x_{\text{src}})$, is a sequential propagator that propagates the s quark from x_{src} to x_{snk} , the b quark from x_{snk} to x_2 and the c quark from x_2 to x_1 .

In Sec. 4.4 we will use this relation to devise a cross-check of our method for the computation of the inclusive decay rate by comparing with the exclusive decay to the ground state. Note that because of the Dirac delta in (2.13) the integral over ω just selects the ground-state energy for the D_s meson with a given momentum. This then implies that $\bar{X}^{(l)} = X^{(l)}$ up to $\delta(\omega - E_{D_s})$. Further details on the ground-state limit can also be found in the Appendix of [36].

2.3 Inclusive decays on an Euclidean space-time lattice

We now address the strategy for the computation of the inclusive decay rate on the lattice, which follows [15, 17, 36]. The key quantity is the hadronic tensor in (2.4)

$$W^{\mu\nu}(q) = \frac{1}{2M_{B_s}} \int d^4x e^{iq \cdot x} \langle B_s | J^{\mu\dagger}(x) J^\nu(0) | B_s \rangle. \quad (2.19)$$

The matrix element in Eq. (2.19) can be extracted from the time dependence of the Euclidean four-point function

$$C_{\mu\nu}^{SJJJ}(\mathbf{q}, t_{\text{snk}}, t_2, t_1, t_{\text{src}}) \stackrel{t_2 \geq t_1}{=} \sum_{\mathbf{x}_{\text{snk}}, \mathbf{x}_{\text{src}}} \left\langle \mathcal{O}_{B_s}^S(\mathbf{x}_{\text{snk}}) \tilde{J}_\mu^\dagger(\mathbf{q}, t_2) \tilde{J}_\nu(\mathbf{q}, t_1) \mathcal{O}_{B_s}^{S\dagger}(\mathbf{x}_{\text{src}}) \right\rangle, \quad (2.20)$$

where $\mathcal{O}_{B_s}^S$ is an interpolating operator with quantum numbers of the B_s meson and the currents are projected onto three-momentum by a discrete Fourier transform $\tilde{J}_\nu(\mathbf{q}, t) = \sum_{\mathbf{x}} e^{-i\mathbf{q} \cdot \mathbf{x}} J_\nu(\mathbf{x}, t)$. In this setup the B_s meson is created with zero momentum at source position x_{src} and annihilated at sink position x_{snk} . In Fig. 2 we show the corresponding quark-flow diagram: the black line, $G_b(x_{\text{src}}, x_1)$, is a propagator for the b quark from x_1 to x_{src} whereas the green one, $\Sigma_{cbs}(x_1, x_{\text{src}})$, is a sequential propagator that propagates the s quark from x_{src} to x_{snk} , the b quark from x_{snk} to x_2 and the c quark from x_2 to x_1 .

The matrix element in Eq. (2.19) can be extracted in the window $t_{\text{snk}} - t_2 \gg 0$, $t_1 - t_{\text{src}} \gg 0$ and $t_2 > t_1$, where excited states of the B_s meson have decayed sufficiently. By increasing the overlap of the operator $\mathcal{O}_{B_s}^S$ with the ground-state B_s state the size of this window can be enlarged. This can be achieved by means of operator smearing, to be detailed later. We use a superscript S in case of smearing and L in case of no smearing.

Within the window we expect

$$C_{\mu\nu}^{SJJS}(\mathbf{q}, t_{\text{snk}}, t_2, t_1, t_{\text{src}}) = \frac{1}{4M_{B_s}^2} \langle 0 | \mathcal{O}_{B_s}^S | B_s \rangle \langle B_s | \tilde{J}_\mu^\dagger(\mathbf{q}, t_2) \tilde{J}_\nu(\mathbf{q}, t_1) | B_s \rangle \langle B_s | \mathcal{O}_{B_s}^{S\dagger} | 0 \rangle. \quad (2.21)$$

In order to extract the B_s forward-scattering matrix element in Eq. (2.19) we cancel the smeared B_s wave function factors $\langle B_s | \mathcal{O}_{B_s}^{S\dagger} | 0 \rangle$ and $\langle 0 | \mathcal{O}_{B_s}^S | B_s \rangle$ by constructing suitable ratios with B_s meson two-point functions with zero momentum

$$C^{SL}(t_2, t_1) = \sum_{\mathbf{x}_2, \mathbf{x}_1} \langle \mathcal{O}_{B_s}^S(x_2) \mathcal{O}_{B_s}^{L\dagger}(x_1) \rangle \quad (2.22)$$

$$\stackrel{t_2 - t_1 \gg 0}{\approx} \frac{1}{2M_{B_s}} \langle 0 | \mathcal{O}_{B_s}^S | B_s \rangle \langle B_s | \mathcal{O}_{B_s}^{L\dagger} | 0 \rangle e^{-(t_2 - t_1)M_{B_s}}.$$

Our choice of ratio is

$$\frac{C_{\mu\nu}^{SJJS}(\mathbf{q}, t_{\text{snk}}, t_2, t_1, t_{\text{src}})}{C^{SL}(t_{\text{snk}}, t_2) C^{LS}(t_1, t_{\text{src}})} \longrightarrow \frac{\frac{1}{2M_{B_s}} \langle B_s | \tilde{J}_\mu^\dagger(\mathbf{q}, t_2) \tilde{J}_\nu(\mathbf{q}, t_1) | B_s \rangle}{\frac{1}{2M_{B_s}} |\langle 0 | \mathcal{O}_{B_s}^L | B_s \rangle|^2}, \quad (2.23)$$

where we cancel the residual factor $|\langle 0 | \mathcal{O}_{B_s}^L | B_s \rangle|^2 / 2M_{B_s}$ with its value obtained from fits to the time-dependence of, e.g., the C^{LL} two-point function. This leads us to define the key observable

$$C_{\mu\nu}(\mathbf{q}, t) = \frac{1}{2M_{B_s}} \langle B_s | \tilde{J}_\mu^\dagger(\mathbf{q}, 0) e^{-\hat{H}t} \tilde{J}_\nu(\mathbf{q}, 0) | B_s \rangle, \quad (2.24)$$

where we have used time-translation invariance $t = t_2 - t_1$. It is related to the hadronic tensor defined in Eq. (2.19) through a Laplace transform

$$C_{\mu\nu}(\mathbf{q}, t) = \int_0^\infty d\omega \frac{1}{2M_{B_s}} \langle B_s | \tilde{J}_\mu^\dagger(\mathbf{q}, 0) \delta(\hat{H} - \omega) \tilde{J}_\nu(\mathbf{q}, 0) | B_s \rangle e^{-\omega t} \quad (2.25)$$

$$= \int_0^\infty d\omega W_{\mu\nu}(\mathbf{q}, \omega) e^{-\omega t},$$

where

$$W_{\mu\nu}(\mathbf{q}, \omega) = \frac{1}{2M_{B_s}} \sum_{X_c} \delta(\omega - E_{X_c}) \langle B_s | \tilde{J}_\mu^\dagger(\mathbf{q}, 0) | X_c \rangle \langle X_c | \tilde{J}_\nu(\mathbf{q}, 0) | B_s \rangle \quad (2.26)$$

corresponds to the spectral representation of $C_{\mu\nu}(\mathbf{q}, t)$. By means of Eq. (2.23) we can compute $C_{\mu\nu}$ on the lattice from a combination of meson two- and four-point functions for a finite and discrete set of Euclidean times t . The determination of the hadronic tensor by means of inversion of the integral equation Eq. (2.25) therefore constitutes an ill-posed inverse problem, similar to the extraction of hadronic spectral densities from Euclidean correlators: while the reconstruction of $C_{\mu\nu}$ from $W_{\mu\nu}$ is straightforward, the other way around is a very difficult task.

Fortunately, in order to compute the inclusive decay rate Eq. (2.9), we do not have to compute the hadronic tensor itself, but only integrals $\bar{X}^{(l)}(\mathbf{q}^2)$, where the hadronic tensor is

smearred with the leptonic tensor integrated over the lepton energy, as defined in Eqs. (2.9)-(2.11). In general, we can write

$$\bar{X}^{(l)}(\mathbf{q}^2) = \int_{\omega_{\min}}^{\omega_{\max}} d\omega W^{\mu\nu}(\mathbf{q}, \omega) k_{\mu\nu}^{(l)}(\mathbf{q}, \omega), \quad (2.27)$$

where $k_{\mu\nu}^{(l)}(\mathbf{q}, \omega)$ is a known kinematic factor that depends only on the energy and three-momentum. Introducing a step function $\theta(\omega_{\max} - \omega)$ and extending the limit of integration as $\omega_{\max} \rightarrow \infty$ and $\omega_{\min} \rightarrow \omega_0$, with $\omega_0 \leq \omega_{\min}$ we can rewrite

$$\begin{aligned} \bar{X}^{(l)}(\mathbf{q}^2) &= \int_{\omega_0}^{\infty} d\omega W^{\mu\nu}(\mathbf{q}, \omega) k_{\mu\nu}^{(l)}(\mathbf{q}, \omega) \theta(\omega_{\max} - \omega) \\ &= \int_{\omega_0}^{\infty} d\omega W^{\mu\nu}(\mathbf{q}, \omega) K_{\mu\nu}^{(l)}(\mathbf{q}, \omega), \end{aligned} \quad (2.28)$$

defining the *kernel function* $K_{\mu\nu}^{(l)}(\mathbf{q}, \omega) = k_{\mu\nu}^{(l)}(\mathbf{q}, \omega) \theta(\omega_{\max} - \omega)$. Note that ω_0 can be chosen freely in $0 \leq \omega_0 \leq \omega_{\min}$ as there are no states below the ground state energy ω_{\min} , as seen from (2.26). For instance, for $B_s \rightarrow X_c l \nu_l$ we expect $\omega_{\min} = M_{D_s}$ for the contribution from the vector channel at vanishing transferred momentum \mathbf{q} . We will later exploit this freedom in the choice of ω_0 .

Let us now discuss how to obtain $\bar{X}^{(l)}$ from lattice data for $C_{\mu\nu}(\mathbf{q}, t)$. First we introduce a smoothing of the kernel $K_{\mu\nu}^{(l)}$ by replacing the step function by a sigmoid of the form

$$\theta_{\sigma}(x) = \frac{1}{1 + e^{-x/\sigma}}. \quad (2.29)$$

While we eventually have to take the limit $\sigma \rightarrow 0$ in order to obtain the physical decay rate, smoothing is useful to control and understand the systematic effects involved in the strategy to compute the decay rate. Following [17], we now expand the smoothed kernel $K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \omega)$ as a polynomial of $e^{-a\omega}$ (we will set $a = 1$ for simplicity) up to some order N , i.e.,

$$K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \omega) \simeq c_{\mu\nu, 0}^{(l)}(\mathbf{q}; \sigma) + c_{\mu\nu, 1}^{(l)}(\mathbf{q}; \sigma) e^{-\omega} + \dots + c_{\mu\nu, N}^{(l)}(\mathbf{q}; \sigma) e^{-\omega N}, \quad (2.30)$$

with N coefficients $c_{\mu\nu, k}^{(l)}(\mathbf{q}; \sigma)$. In this way, the target quantity $\bar{X}_{\sigma}^{(l)}(\mathbf{q}^2)$, which now also depends on the smearing parameter σ , can be computed as

$$\begin{aligned} \bar{X}_{\sigma}^{(l)}(\mathbf{q}^2) &= \int_{\omega_0}^{\infty} d\omega W^{\mu\nu}(\mathbf{q}, \omega) e^{-2\omega t_0} K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \omega; t_0) \\ &\simeq c_{\mu\nu, 0}^{(l)} \int_{\omega_0}^{\infty} d\omega W^{\mu\nu}(\mathbf{q}, \omega) e^{-2\omega t_0} + c_{\mu\nu, 1}^{(l)} \int_{\omega_0}^{\infty} d\omega W^{\mu\nu}(\mathbf{q}, \omega) e^{-2\omega t_0} e^{-\omega} + \dots \\ &\quad + c_{\mu\nu, N}^{(l)} \int_{\omega_0}^{\infty} d\omega W^{\mu\nu}(\mathbf{q}, \omega) e^{-2\omega t_0} e^{-\omega N}. \end{aligned} \quad (2.31)$$

The factor $e^{-2\omega t_0}$ has been introduced, and compensated for in $K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \omega; t_0) = e^{2\omega t_0} K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \omega)$, in order to avoid the equal-time matrix element $t_1 = t_2$, see Eq. (2.20), which contains contributions from the opposite time ordering corresponding to unphysical

$\bar{b}sc\bar{b}$ final states. We will discuss suitable choices for the free parameter t_0 together with the discussion of the analysis of actual simulation data. Inserting now Eq. (2.25) we arrive at the compact expression

$$\bar{X}_\sigma^{(l)}(\mathbf{q}^2) = \sum_{k=0}^N c_{\mu\nu,k}^{(l)} C^{\mu\nu}(\mathbf{q}, k + 2t_0), \quad (2.32)$$

which relates $C^{\mu\nu}(\mathbf{q}, t)$, which can be computed on the lattice, to $\bar{X}_\sigma^{(l)}(\mathbf{q}^2)$. The expression is understood to be an approximation of $\bar{X}_\sigma^{(l)}(\mathbf{q}^2)$ due the truncation to a finite value N ; we use the same convention for all similar quantities that we address in the following sections. Note that the order N of the polynomial approximation is now directly related to the separation in Euclidean time of the two charged currents in the four-point function in Eq. (2.20). What remains to be done towards the computation of the decay rate for a given value of σ , is to carry out the phase-space integration in Eq. (2.9).

Before we close this section, let us list the explicit expressions for the kernels $K_{\sigma,\mu\nu}^{(l)}$:

$$K_{\sigma,00}^{(0)}(\mathbf{q}, \omega; t_0) = e^{2\omega t_0} \mathbf{q}^2 \theta_\sigma(\omega_{\max} - \omega), \quad (2.33)$$

$$K_{\sigma,ii}^{(0)}(\mathbf{q}, \omega; t_0) = e^{2\omega t_0} (q_i^2 - \mathbf{q}^2) \theta_\sigma(\omega_{\max} - \omega), \quad (2.34)$$

$$K_{\sigma,ij}^{(0)}(\mathbf{q}, \omega; t_0) \stackrel{i \neq j}{=} e^{2\omega t_0} q_i q_j \theta_\sigma(\omega_{\max} - \omega), \quad (2.35)$$

$$K_{\sigma,0i}^{(1)}(\mathbf{q}, \omega; t_0) = -e^{2\omega t_0} q_i q_0 \theta_\sigma(\omega_{\max} - \omega), \quad (2.36)$$

$$K_{\sigma,ii}^{(2)}(\mathbf{q}, \omega; t_0) = e^{2\omega t_0} q_0^2 \theta_\sigma(\omega_{\max} - \omega). \quad (2.37)$$

For the parallel and perpendicular components at $\mathbf{q}^2 \neq 0$, as defined in Sec. 2.2, we have

$$K_{\sigma,00}^{\parallel}(\mathbf{q}, \omega; t_0) = e^{2\omega t_0} \mathbf{q}^2 \theta_\sigma(\omega_{\max} - \omega), \quad (2.38)$$

$$K_{\sigma,0i}^{\parallel}(\mathbf{q}, \omega; t_0) = -e^{2\omega t_0} q_0 q_i \theta_\sigma(\omega_{\max} - \omega), \quad (2.39)$$

$$K_{\sigma,ij}^{\parallel}(\mathbf{q}, \omega; t_0) = e^{2\omega t_0} \frac{q_0^2}{\mathbf{q}^2} q_i q_j \theta_\sigma(\omega_{\max} - \omega), \quad (2.40)$$

$$K_{\sigma,ii}^{\perp}(\mathbf{q}, \omega; t_0) = e^{2\omega t_0} (q_i^2 - \mathbf{q}^2) \left(1 - \frac{q_0^2}{\mathbf{q}^2}\right) \theta_\sigma(\omega_{\max} - \omega), \quad (2.41)$$

$$K_{\sigma,ij}^{\perp}(\mathbf{q}, \omega; t_0) \stackrel{i \neq j}{=} e^{2\omega t_0} q_i q_j \left(1 - \frac{q_0^2}{\mathbf{q}^2}\right) \theta_\sigma(\omega_{\max} - \omega). \quad (2.42)$$

All other index combinations vanish.

2.4 Data analysis

In the previous section we reduced the problem of computing the inclusive decay rate to that of finding a suitable polynomial approximation for the kernel $K_{\sigma,\mu\nu}^{(l)}(\mathbf{q}, \omega; t_0)$. Here we describe two separate methods that we follow (and later compare in Sec. 4), for determining the expansion coefficients $c_{\mu\nu,k}^{(l)}$ given lattice data for the ratio of correlation functions in Eq. (2.23).

The analysis has to deal with the statistical noise from the data and also systematic errors, e.g. those associated with the polynomial approximation. Here we consider data for

a single lattice spacing and lattice volume, leaving discretisation and finite-volume errors for future studies.

In principle, $\bar{X}_\sigma^{(l)}(\mathbf{q}^2)$ as defined in Eq. (2.32), could be computed straightforwardly from lattice data for $C_{\mu\nu}(\mathbf{q}, t)$. For a given order N , the coefficients $c_{\mu\nu, k}^{(l)}$ in the power series for the analytically known kernel $K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \omega)$ could, for instance, be determined via linear regression, allowing to construct $\bar{X}_\sigma^{(l)}(\mathbf{q}^2)$ from the data for $C_{\mu\nu}(\mathbf{q}, t)$. The order of the expansion is limited by the number of time slices in the window where $C_{\mu\nu}(\mathbf{q}, t)$ can be extracted from the lattice data. Unfortunately, the exponential deterioration of the signal-to-noise ratio with increasing Euclidean time separation t makes a meaningful signal for the decay rate difficult to extract. What is needed is some form of regulator that provides balance between statistical noise and systematic error due to the truncation. We proceed with outlining two methods that achieve this: one based on Chebyshev polynomials and the other based on the modified Backus-Gilbert method.

For the sake of readability we introduce the following notation

$$\begin{aligned}\bar{X}_\sigma^{(l)}(\mathbf{q}^2) &= \int_{\omega_0}^{\infty} d\omega W^{\mu\nu}(\mathbf{q}, \omega) e^{-2\omega t_0} K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \omega; t_0) \\ &= \frac{1}{2M_{B_s}} \int_{\omega_0}^{\infty} d\omega K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \omega; t_0) \langle B_s | \tilde{J}^{\mu\dagger}(\mathbf{q}, 0) e^{-\omega t_0} \delta(\hat{H} - \omega) e^{-\omega t_0} \tilde{J}^\nu(\mathbf{q}, 0) | B_s \rangle \\ &= \langle \psi^\mu(\mathbf{q}) | K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \hat{H}; t_0) | \psi^\nu(\mathbf{q}) \rangle ,\end{aligned}\tag{2.43}$$

where we made use of Eq. (2.26) and defined $|\psi^\nu(\mathbf{q})\rangle = e^{-\hat{H}t_0} \tilde{J}^\nu(\mathbf{q}, 0) | B_s \rangle / \sqrt{2M_{B_s}}$. Note that the kernel has been promoted to an operator, $K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \hat{H}; t_0)$.

2.4.1 Chebyshev-polynomial approximation

Chebyshev polynomials $T_k(\omega)$ defined on $-1 \leq \omega \leq 1$ provide an optimal approximation of functions under the L_∞ -norm. We provide a summary of basic properties in App. A. For the case at hand we define shifted Chebyshev polynomials $\tilde{T}_k(\omega)$, which are defined in the interval $\omega_0 \leq \omega \leq \infty$. Here, $\tilde{T}_k(\omega) = T_k(h(\omega))$, and $h(\omega) = Ae^{-\omega} + B$ is a map $h : [\omega_0, \infty) \rightarrow [-1, 1]$, where expressions for the coefficients A and B can be found in Eq. (A.23). The kernel function from the previous section can then be expanded up to order N as

$$K_{\sigma, \mu\nu}^{(l)}(\mathbf{q}, \omega; t_0) = \frac{1}{2} \tilde{c}_{\mu\nu, 0}^{(l)} \tilde{T}_0(\omega) + \sum_{k=1}^N \tilde{c}_{\mu\nu, k}^{(l)} \tilde{T}_k(\omega),\tag{2.44}$$

where $\tilde{T}_0(\omega) = 1$ by definition, and

$$\tilde{T}_k(\omega) = \sum_{j=0}^k \tilde{t}_j^{(k)} e^{-j\omega},\tag{2.45}$$

with coefficients $\tilde{t}_j^{(k)}$ defined and discussed in App. A.2. Making use of the Chebyshev polynomials' orthogonality properties, the coefficients $\tilde{c}_{\mu\nu, k}^{(l)}$ are defined by projection as in

Eq. (A.21),

$$\tilde{c}_{\mu\nu,k}^{(l)} = \int_{\omega_0}^{\infty} d\omega K_{\sigma,\mu\nu}^{(l)}(\mathbf{q}, \omega; t_0) \tilde{T}_k(\omega) \Omega_h(\omega), \quad (2.46)$$

where the weight function $\Omega_h(x)$ is defined in App. A. In this way, the expectation value of the kernel operator is

$$\langle \psi^\mu | K_{\sigma,\mu\nu}^{(l)}(\mathbf{q}, \hat{H}; t_0) | \psi^\nu \rangle = \frac{1}{2} \tilde{c}_{\mu\nu,0}^{(l)} \langle \psi^\mu | \tilde{T}_0(\hat{H}) | \psi^\nu \rangle + \sum_{k=1}^N \tilde{c}_{\mu\nu,k}^{(l)} \langle \psi^\mu | \tilde{T}_k(\hat{H}) | \psi^\nu \rangle. \quad (2.47)$$

By construction, in particular thanks to the condition of Eq. (A.9), shifted Chebyshev polynomials are bounded, $|\tilde{T}_k(\omega)| \leq 1$. As we will discuss later, this a crucial ingredient in the data analysis: in order to make use of this property, we divide the terms $\langle \psi^\mu | \tilde{T}_k(\hat{H}) | \psi^\nu \rangle$ by a normalisation factor $\langle \psi^\mu | \psi^\nu \rangle = C^{\mu\nu}(2t_0)$. For a more compact notation we define

$$\langle K_{\sigma}^{(l)} \rangle_{\mu\nu} \equiv \frac{\langle \psi_\mu | K_{\sigma,\mu\nu}^{(l)}(\mathbf{q}, \hat{H}; t_0) | \psi_\nu \rangle}{\langle \psi_\mu | \psi_\nu \rangle}, \quad \langle \tilde{T}_k \rangle_{\mu\nu} \equiv \frac{\langle \psi_\mu | \tilde{T}_k(\hat{H}) | \psi_\nu \rangle}{\langle \psi_\mu | \psi_\nu \rangle}, \quad (2.48)$$

such that

$$\langle K_{\sigma}^{(l)} \rangle_{\mu\nu} = \frac{1}{2} \tilde{c}_{\mu\nu,0}^{(l)} \langle \tilde{T}_0 \rangle_{\mu\nu} + \sum_{k=1}^N \tilde{c}_{\mu\nu,k}^{(l)} \langle \tilde{T}_k \rangle_{\mu\nu}, \quad (2.49)$$

where in this case there is no summation on μ, ν . We refer to $\langle \tilde{T}_k \rangle_{\mu\nu}$ as the *Chebyshev matrix elements*, for which, thanks to the normalisation, $|\langle \tilde{T}_k \rangle_{\mu\nu}| \leq 1$. In terms of the Chebyshev expansion the expression for $\bar{X}_{\sigma}^{(l)}(\mathbf{q}^2)$ now reads

$$\bar{X}_{\sigma}^{(l)}(\mathbf{q}^2) = \sum_{\{\mu,\nu\}} \langle \psi_\mu | \psi_\nu \rangle \langle K_{\sigma}^{(l)} \rangle_{\mu\nu}, \quad (2.50)$$

and explicitly

$$\bar{X}_{\sigma}^{(0)} = C_{00}(2t_0) \langle K_{\sigma}^{(0)} \rangle_{00} + \sum_i C_{ii}(2t_0) \langle K_{\sigma}^{(0)} \rangle_{ii} + \sum_{i \neq j} C_{ij}(2t_0) \langle K_{\sigma}^{(0)} \rangle_{ij}, \quad (2.51)$$

$$\bar{X}_{\sigma}^{(1)} = \sum_i \left(C_{0i}(2t_0) \langle K_{\sigma}^{(1)} \rangle_{0i} + C_{i0}(2t_0) \langle K_{\sigma}^{(1)} \rangle_{i0} \right), \quad (2.52)$$

$$\bar{X}_{\sigma}^{(2)} = \sum_i C_{ii}(2t_0) \langle K_{\sigma}^{(2)} \rangle_{ii}. \quad (2.53)$$

The Chebyshev matrix elements can be constructed directly from the lattice data using

$$\frac{\langle \psi_\mu | e^{-\hat{H}t} | \psi_\nu \rangle}{\langle \psi_\mu | \psi_\nu \rangle} = \frac{C_{\mu\nu}(t + 2t_0)}{C_{\mu\nu}(2t_0)} \equiv \bar{C}_{\mu\nu}(t). \quad (2.54)$$

Using the properties of shifted Chebyshev polynomials as detailed in App. A.2, we can directly relate the matrix element $\langle \tilde{T}_k \rangle_{\mu\nu}$ to the correlator $\bar{C}_{\mu\nu}$. In particular,

$$\begin{aligned} \langle \tilde{T}_k \rangle_{\mu\nu} &= \frac{\langle \psi_\mu | \tilde{T}_k(\hat{H}) | \psi_\nu \rangle}{\langle \psi_\mu | \psi_\nu \rangle} = \sum_{X_c} \frac{\langle \psi_\mu | \tilde{T}_k(\hat{H}) | X_c \rangle \langle X_c | \psi_\nu \rangle}{\langle \psi_\mu | \psi_\nu \rangle} \\ &= \sum_{X_c} \sum_{j=0}^k \tilde{t}_j^{(k)} e^{-jE_{X_c}} \frac{\langle \psi_\mu | X_c \rangle \langle X_c | \psi_\nu \rangle}{\langle \psi_\mu | \psi_\nu \rangle} \\ &= \sum_{j=0}^k \tilde{t}_j^{(k)} \bar{C}_{\mu\nu}(j), \end{aligned} \quad (2.55)$$

where we have inserted the identity $I = \sum_{X_c} |X_c\rangle \langle X_c|$ and $\tilde{t}_j^{(k)}$ are defined in (A.15). Overall the full Chebyshev expansion of the kernel reads

$$\begin{aligned} \langle K_\sigma^{(l)} \rangle_{\mu\nu} &= \frac{1}{2} \bar{c}_{\mu\nu,0}^{(l)} \langle \tilde{T}_0 \rangle_{\mu\nu} + \sum_{k=1}^N \bar{c}_{\mu\nu,k}^{(l)} \langle \tilde{T}_k \rangle_{\mu\nu} \\ &= \sum_{k=0}^N \bar{C}_{\mu\nu}(k) \sum_{j=k}^N \bar{c}_{\mu\nu,j}^{(l)} \left(1 - \frac{1}{2} \delta_{0j}\right) \tilde{t}_k^{(j)}, \end{aligned} \quad (2.56)$$

where we emphasise once more that the analytical expressions for the coefficients $\bar{c}_{\mu\nu,j}^{(l)}$ and $\tilde{t}_k^{(j)}$ are known and can be evaluated. Collecting the coefficients into

$$\bar{c}_{\mu\nu,k}^{(l)} \equiv \sum_{j=k}^N \bar{c}_{\mu\nu,j}^{(l)} \tilde{t}_k^{(j)} \left(1 - \frac{1}{2} \delta_{0j}\right), \quad (2.57)$$

we arrive at the compact expression

$$\langle K_\sigma^{(l)} \rangle_{\mu\nu} = \sum_{k=0}^N \bar{c}_{\mu\nu,k}^{(l)} \bar{C}_{\mu\nu}(k). \quad (2.58)$$

While $\bar{c}_{\mu\nu,k}^{(l)}$ is known in terms of solvable analytical expressions, $\bar{C}_{\mu\nu}(k)$ needs to be computed on the lattice using Monte-Carlo methods. The resulting statistical error on $\bar{C}_{\mu\nu}(k)$ can lead to violations of the bound $|\langle \tilde{T}_k \rangle_{\mu\nu}| \leq 1$ when solving the linear system in Eq. (2.55). This can however be avoided in a Bayesian analysis of the correlator data, imposing the bound in terms of priors. One way to impose the constraint is to use a Gaussian prior on some internal parameters $\langle \tilde{T}_k \rangle_{\mu\nu} \sim \mathcal{N}(0, 1)$ and convert it to a flat prior on the interval $[-1, 1]$ using the map $f(x) = \text{erf}(x/\sqrt{2})$ such that $\langle \tilde{T}_k \rangle_{\mu\nu} = f(\langle \tilde{T}_k \rangle_{\mu\nu})$. We refer to App. C for a thorough discussion on the fitting procedure that we adopt.

2.4.2 Backus-Gilbert

A different approach to determine the polynomial approximation of the kernel is given by a variant of the Backus-Gilbert method [37] proposed in [22, 38]. In this work, we consider a more general scenario to allow the use of different polynomial bases following [39].

Note that, although what we propose is mathematically equivalent to the approach in [39], our formulation may have the advantage of avoiding some of the numerical technicalities that arise in the original version. Indeed, while the latter requires the inversion of an ill-conditioned matrix with the help of arbitrary precision arithmetic, our approach relies on the inversion of an equivalent diagonal matrix in the case where an orthogonal polynomial basis is chosen, at least as far as the systematics are concerned. We briefly present the idea below and refer to App. B for a more detailed discussion. Note that we adopt a different notation with respect to the original works (we use F instead of W for the final functional to avoid confusion with the hadronic tensor).

The central idea is to address the reconstruction of the (smeared) kernel $K_{\sigma,\mu\nu}^{(l)}$ of the form

$$K_{\sigma,\mu\nu}^{(l)}(\mathbf{q}, \omega; t_0) = \sum_{k=0}^N g_{\mu\nu,k}^{(l)} \tilde{P}_k(\omega), \quad (2.59)$$

where $\tilde{P}_k(\omega) = \sum_{j=0}^k \tilde{p}_j^{(k)} e^{-j\omega}$ are a basis of functions defined on $[\omega_0, \infty)$, and $g_{\mu\nu,k}^{(l)} \equiv g_{\mu\nu,k}^{(l)}(\mathbf{q}, \sigma; t_0)$ is a set of coefficients to be determined. In order to compute them, the strategy is to minimise the functional

$$F_{\mu\nu,\lambda}^{(l)}[g] = (1 - \lambda) \frac{A_{\mu\nu}^{(l)}[g]}{A_{\mu\nu}^{(l)}[0]} + \lambda B_{\mu\nu}^{(l)}[g], \quad (2.60)$$

where

$$A_{\mu\nu}^{(l)}[g] = \int_{\omega_0}^{\infty} d\omega \Omega(\omega) \left[K_{\sigma,\mu\nu}^{(l)}(\mathbf{q}, \omega; t_0) - \sum_{k=0}^N g_{\mu\nu,k}^{(l)} \tilde{P}_k(\omega) \right]^2 \quad (2.61)$$

is the L_2 -norm of the difference between the target kernel function and its reconstruction, weighted with a smooth function $\Omega(\omega)$, and

$$B_{\mu\nu}^{(l)}[g] = \sum_{j,k=0}^N g_{\mu\nu,j}^{(l)} \text{Cov}[\bar{C}_{\mu\nu}^P(j), \bar{C}_{\mu\nu}^P(k)] g_{\mu\nu,k}^{(l)} \quad (2.62)$$

is the variance of the corresponding channel $\bar{X}_{\mu\nu}^{(l)}$, with $\bar{C}_{\mu\nu}^P(k) = \sum_{j=0}^k \tilde{p}_j^{(k)} \bar{C}_{\mu\nu}(j)$. The functional $F_{\mu\nu,\lambda}^{(l)}$ encodes the information about both systematic and statistical error, whose interplay is controlled by the parameter $\lambda \in [0, 1)$, which in principle can be chosen by hand. The values of the coefficients $g_{\mu\nu,k}^{(l)}(\lambda)$ for each λ are given by the variational principle, i.e.

$$g_{\mu\nu,k}^{(l)}(\lambda) \leftrightarrow \frac{\partial F_{\mu\nu,\lambda}^{(l)}}{\partial g_{\mu\nu,k}^{(l)}} = 0. \quad (2.63)$$

We can now devise a method to find the optimal λ^* . Following [38], we can simply evaluate the functional $F_{\mu\nu,\lambda}^{(l)}$ at its minimum i.e. $F_{\mu\nu}^{(l)}(\lambda) = F_{\mu\nu,\lambda}^{(l)}[g(\lambda)]$, which then becomes a function of λ , and require that λ^* maximises $F_{\mu\nu}^{(l)}(\lambda)$, $\left. \frac{dF_{\mu\nu}^{(l)}(\lambda)}{d\lambda} \right|_{\lambda^*} = 0$. It is clear that this

choice corresponds to $A_{\mu\nu}^{(l)}[g^*]/A_{\mu\nu}^{(l)}[0] = B_{\mu\nu}^{(l)}[g^*]$, i.e. an optimal balance between statistical and systematic errors. This is the prescription we follow and take $g_{\mu\nu,k}^{*(l)} \equiv g_{\mu\nu,k}^{(l)}(\lambda^*)$.

Following the steps for the Chebyshev approach we get for the kernel

$$\langle K_\sigma^{(l)} \rangle_{\mu\nu} = \sum_{k=0}^N g_{\mu\nu,k}^{*(l)} \langle \tilde{P}_k \rangle_{\mu\nu}, \quad (2.64)$$

$$\langle \tilde{P}_k \rangle_{\mu\nu} = \frac{\langle \psi_\mu | \tilde{P}_k(\hat{H}) | \psi_\nu \rangle}{\langle \psi_\mu | \psi_\nu \rangle} = \sum_{j=0}^k \tilde{p}_j^{(k)} \frac{\langle \psi_\mu | e^{-j\hat{H}} | \psi_\nu \rangle}{\langle \psi_\mu | \psi_\nu \rangle} = \bar{C}_{\mu\nu}^P(k). \quad (2.65)$$

In particular, considering the domain $[\omega_0, \infty)$, we focus on two choices:

- *exponential Backus-Gilbert*: $\tilde{P}_k(\omega) = e^{-k\omega}$ and $\Omega(\omega) = 1$ (and set $g_{\mu\nu,0}^{(l)} = 0$ by hand, as in the original proposal [22]);
- *Chebyshev Backus-Gilbert*: $\tilde{P}_k(\omega) = \tilde{T}_k(\omega)$, i.e. the shifted Chebyshev polynomials with $\Omega(\omega) = 1/\sqrt{e^{a(\omega-\omega_0)} - 1}$ being the weight that enters in the definition of the scalar product as in (A.13).

3 Numerical setup

We perform a pilot study using a $24^3 \times 64$ lattice with 2+1-flavour domain-wall fermion (DWF) [40, 41] gauge-field ensembles with the Iwasaki gauge action [42] taken from the RBC/UKQCD Collaboration [43] at lattice spacing $a \simeq 0.11$ fm and pion mass $M_\pi \simeq 330$ MeV. The correlation functions analysed in this paper have been generated with the Grid [44–46] and Hadrons [47] software packages. Part of the fits in the analysis have been performed using lsqfit [48, 49].

We use the same simulation parameter RBC/UKQCD is using in the heavy-light meson projects on exclusive semileptonic $B_{(s)}$ meson decays [35, 50–52]. In particular, the valence-strange quark is simulated using DWF, whereas the valence-charm quark is simulated by using the Möbius DWF action [29, 30]. Their masses are tuned such that mesons containing bottom, charm and strange valence quarks have masses close to the physical ones. The bottom quark has been simulated at its physical mass using the Columbia formulation of the relativistic-heavy-quark (RHQ) action [25, 26], which is based on the Fermilab heavy quark action [24]. In particular, this formulation allows to reduce the b -quark discretisation effects of order $\mathcal{O}((m_0 a)^n)$, $\mathcal{O}(\mathbf{p}a)$ and $\mathcal{O}((\mathbf{p}a)(m_0 a)^n)$ by tuning three nonperturbative parameters, one of them being the bare mass m_0 .

For the computation we average over 120 statistically independent gauge configurations, and on each configuration the measurements are performed on 8 different linearly spaced source time planes. We use \mathbb{Z}_2 wall sources [53–55] to improve the signal. We induce 10 different momenta in the four-point functions in Eq. (2.20) using twisted boundary conditions [56, 57] with the same momentum in all three spatial directions. Considering $\mathbf{q} = 2\pi\boldsymbol{\theta}/L$ in lattice units we have $\boldsymbol{\theta} \equiv (\theta, \theta, \theta)$, where θ indicates the twist. We choose them such that all the momenta are linearly spaced in \mathbf{q}^2 : $\theta_k = 1.90 \sqrt{\frac{k}{3}}$ for $k = 0, 1, \dots, 7$,

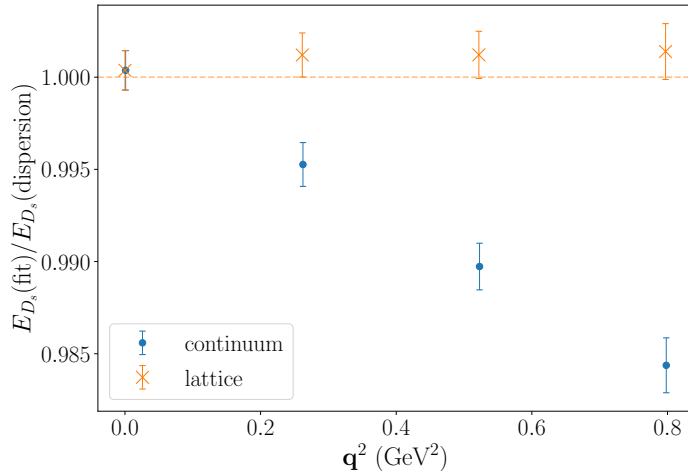


Figure 3. Speed-of-light plot for the D_s meson. The numerator is the energy of the ground state mass for a given momentum as extracted from a fit to the data. The denominator is given by either the lattice dispersion relation or the continuum one, where the D_s mass has been determined from a fit to the data.

where the factor 1.90 is determined by the value of $\mathbf{q}_{\text{max}}^2 = 1.83$ in lattice units. We also take $\theta = 1.90 \sqrt{\frac{1}{9}}$ and $\theta = 1.90 \sqrt{\frac{2}{9}}$ to increase the resolution in \mathbf{q}^2 for small momenta.

We compute two-point functions for both B_s and D_s . As discussed in Sec. 2.3, for B_s we consider three cases at zero momentum $C_{B_s}^{LS}(t, t_{\text{src}})$, $C_{B_s}^{SL}(t, t_{\text{src}})$ and $C_{B_s}^{SS}(t, t_{\text{src}})$ with different smearing combinations, as indicated by the superscripts “L” (local) and “S” (smeared). The smeared-smeared $C_{B_s}^{SS}(t, t_{\text{src}})$ is also used to determine the renormalisation constant together with the three-point functions. The sources are smeared gauge-invariantly using Jacobi iteration [58, 59] using the same parameters as in RBC/UKQCD’s study of exclusive semileptonic decays in [52, 60, 61].

The D_s correlators are relevant mainly for the analysis of the ground-state limit in Sec. 4.4. We consider again three different combinations of smearing at source and sink and we induce momenta for the c quark with the available twists. We show the speed of light from the fitted masses of the D_s for the smallest momenta, comparing with the continuum dispersion relation and the lattice dispersion relation in Fig. 3. The latter shows excellent agreement with the fitted energies.

We also compute three-point correlators for the $B_s \rightarrow D_s l \nu_l$ process

$$C_{B_s D_s, \mu}^{SS}(\mathbf{q}, t_{\text{snk}}, t, t_{\text{src}}) = \sum_{\mathbf{x}_{\text{snk}}, \mathbf{x}} \langle \mathcal{O}_{B_s}^S(\mathbf{x}_{\text{snk}}) V_\mu(\mathbf{x}, t) \mathcal{O}_{D_s}^{S\dagger}(\mathbf{x}_{\text{src}}) \rangle. \quad (3.1)$$

Following the analysis of [35, 50, 51], we extract its form factors and compare with our inclusive results. The momentum is carried by the charm quark through twisted boundary conditions, $\mathbf{q} = 2\pi\boldsymbol{\theta}/L$. We use a source-sink separation of $t_{\text{snk}} - t_{\text{src}} = 20$ in lattice units. The corresponding quark-flow diagram is depicted in Fig. 4.

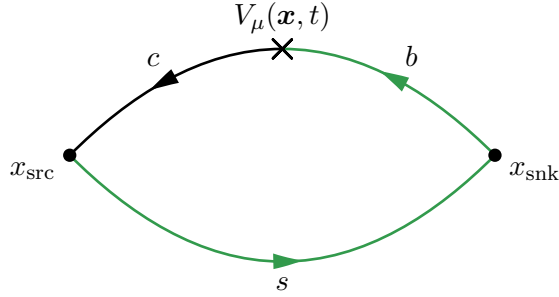


Figure 4. Three-point correlator diagram for the exclusive channel $B_s \rightarrow D_s l \nu_l$.

We now move to the four-point correlators defined in Eq. (2.20), which are the building blocks in the computation of inclusive processes. We use the same source-sink separation as for the three-point functions, i.e., $t_{\text{snk}} - t_{\text{src}} = 20$ in lattice units. The current J_μ^\dagger is fixed at the time slice $t_2 = t_{\text{src}} + 14$, such that the time dependence is enclosed in $0 \leq t \leq 14$ with $t = t_2 - t_1$. For this choice we find ground state saturation at the points where we insert the currents. In practice, referring to Fig. 2, the contractions are performed between a b -quark propagator $G_b(x_1, x_{\text{src}})$ and a sequential propagator $\Sigma_{cbs}(x_1, x_{\text{src}})$. For the latter, we first propagate the s quark to point x_{snk} , starting from a \mathbb{Z}_2 wall source at t_{src} ; we then use it as a sequential source at fixed t_{snk} with zero momentum to propagate the b quark. The b quark is propagated to point x_2 , and it is then used again as a source with a specific choice of gamma matrix corresponding to the current $J_\mu^\dagger(x_2)$ and the momentum insertion to propagate the c quark.

As before, the momentum \mathbf{q} induced through twisted boundary conditions is carried by the c quark. Given that we are dealing with $(V - A)$ currents, we consider all possible combinations of $J_\mu^\dagger(x_2)$ and $J_\nu(x_1)$, i.e. $V_\mu^\dagger V_\nu, V_\mu^\dagger A_\nu, A_\mu^\dagger V_\nu, A_\mu^\dagger A_\nu$. However, in the limit of massless leptons the combinations $A_\mu^\dagger V_\nu$ and $V_\mu^\dagger A_\nu$ do not contribute to the total decay rate. Indeed, these terms are related to the structure function W_3 as $W_{ij}^{AV} + W_{ij}^{VA} = i\epsilon_{ij0k}q^k W_3$, as can be seen analysing parity in Eq. (2.7), which does not contribute to the total decay rate for $m_l = 0$.

The local vector and axial-vector currents used in our lattice calculation receive a finite renormalisation. We use the *almost nonperturbative* prescription of [62], whereby

$$V_\mu = \rho_V^{bc} \sqrt{Z_V^{cc} Z_V^{bb}} (V_\mu)_{\text{bare}} \quad \text{and} \quad A_\mu = \rho_A^{bc} \sqrt{Z_V^{cc} Z_V^{bb}} (A_\mu)_{\text{bare}} . \quad (3.2)$$

The subscript “bare” indicates the bare, unrenormalised heavy-light vector or axial-vector current. Z_V^{cc} is the vector-current renormalisation constant for domain-wall fermions. Due to the approximate chiral symmetry of domain-wall fermions, $Z_V^{cc} = Z_A^{cc}$ up to residual chiral-symmetry-breaking effects. The renormalisation constants Z_V^{bb} and Z_V^{cc} are computed from the *charge* of the heavy-light mesons, and are defined as

$$Z_V^{bb} = \frac{C_{B_s}^{SS}(t_{\text{snk}}, t_{\text{src}})}{C_{B_s B_s, 0}^{SS}(t_{\text{snk}}, t, t_{\text{src}})} \quad \text{and} \quad Z_V^{cc} = \frac{C_{D_s}^{SS}(t_{\text{snk}}, t_{\text{src}})}{C_{D_s D_s, 0}^{SS}(t_{\text{snk}}, t, t_{\text{src}})} , \quad (3.3)$$

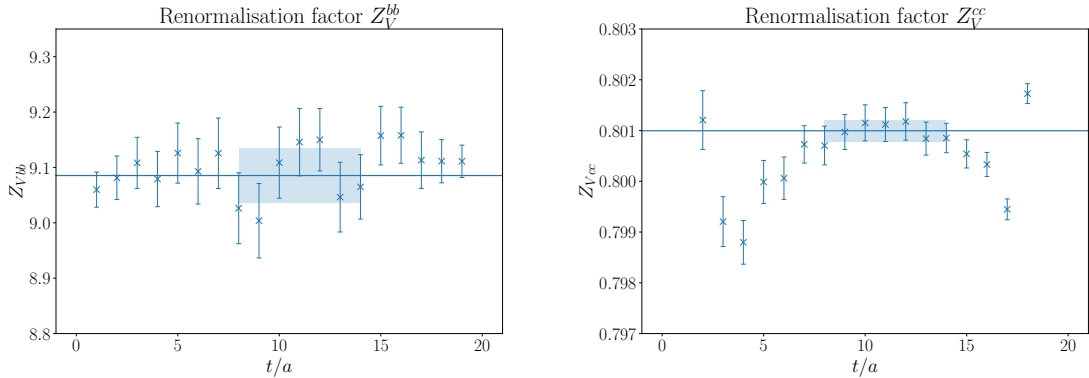


Figure 5. Determination of renormalisation Z_V^{bb} (left) and Z_V^{cc} (right) from the ratio of two- and three-point functions defined in Eq. (3.3).

where both the two- and three-point functions are zero-momentum projected. The results for $Z_V^{bb} = 9.085(50)$ and $Z_V^{cc} = 0.80099(21)$ are reported in Fig. 5. The coefficient $\rho_{V/A}^{bc}$ is expected to be close to unity and can be computed in perturbation theory. Here we set it to its tree-level value, i.e. $\rho_{V/A}^{bc} = 1$. This is sufficient for the qualitative study aimed at here, where no attempt is made at taking the continuum limit.

For all the three-point and four-point functions we always average over the spatial directions given that the momentum is the same in all three directions. Note in particular that for the four-point correlators we have to average separately over $J_i^\dagger J_i$ and $J_i^\dagger J_k$ with $i \neq k$, which can be seen from Eq. (2.7).

4 Results

In this section we present and discuss the main results of our work. We first discuss how well the kernels $K_{\mu\nu,\sigma}^{(l)}$ are approximated by the polynomials and then discuss the reconstruction via Chebyshev and Backus-Gilbert methods. Eventually we combine various analysis steps for a prediction of the inclusive decay rate. Towards the end of this section we compare our results with the ground-state contribution. We emphasise that the work presented here focuses on a qualitative understanding of the methods aiming at developing reliable techniques, which in future work can be used to make phenomenologically relevant predictions.

4.1 Polynomial approximation of the kernel

In this section we discuss the key aspects of the polynomial approximation. The two ingredients to optimise the approximation are the choice of the starting point of the approximation ω_0 , and the value of t_0 in (2.54). In particular, we choose $t_0 = 1/2$ in lattice units, such that the exponential growth of the term $e^{2\omega t_0}$ in the kernels (2.33)-(2.37) is minimal, and the number of data points we can use is maximised. We study two values of ω_0 , i.e. $\omega_0 = 0$ and $\omega_0 = 0.9\omega_{\min}$ for each momentum \mathbf{q}^2 . Note that this section deals purely with the

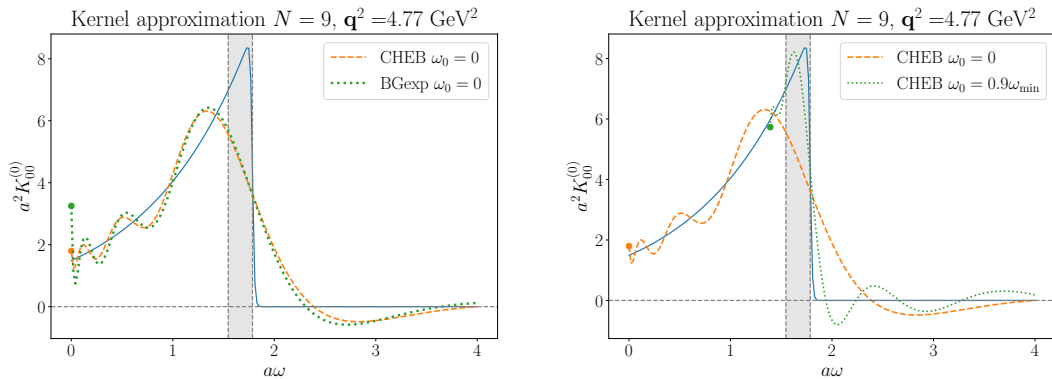


Figure 6. Comparison between Chebyshev and Backus-Gilbert (with exponential basis) approach with $N = 9$ at $\omega_0 = 0$ (left) and comparison between Chebyshev approach with different values of ω_0 (right) for kernel $K_{00}^{(0)}$ at $\mathbf{q}^2 = 4.77 \text{ GeV}^2$. The solid blue line shows the target kernel function with a smearing $\sigma = 0.02$.

approximation of the kernel with no connection to the data; for the Backus-Gilbert method this means that we set $\lambda = 0$.

In Fig. 6 we highlight some of the key features of our approach and in Fig. 7 we show the approximation for different kernels $K_{\mu\nu,\sigma}^{(l)}$ with $l = 0, 1, 2$. The plots are for the smallest and one of the largest \mathbf{q}^2 computed, respectively. Here we illustrate the case of $\sigma = 0.02$, which smooths the step function only mildly. Later we will also discuss the case of larger values of σ .

Some comments are in order. First of all, we point out that with the current data set, the polynomial order $N = 9$ is the maximum value available. This depends on the size of the lattice and the choice of t_{src} , t_2 and t_{snk} in the four-point correlator. In particular, setting $a = 1$, the available time slices are $2t_0 \leq t < t_2 - t_{\text{src}}$, which in our case correspond to $1 \leq t < 14$. On top of that, we need to make sure that $t \ll t_2 - t_{\text{src}}$, i.e. $t_1 - t_{\text{src}} \gg 0$: the choice $N = 9$ corresponds to a separation $t_1 - t_{\text{src}} = 4$. Of course, with an improved data set N could be chosen larger and the differences between the two approaches would reduce further.

We also notice that the kernel with $l = 0$ is the most delicate to treat, as it is the one that shows the sharpest drop to zero at the threshold. Note also that for the case $l = 0$ we plotted only $K_{00}^{(0)}$ as all the other kernels are the same up to a constant factor. Secondly, as shown in Fig. 6 (left) the results for Chebyshev and Backus-Gilbert agree very well and the quality of the approximation seems comparable.

The quality of the approximation varies with ω_0 : as shown in Fig. 6 (right), starting the approximation as close as possible to ω_{\min} gives the best result, as the nodes of the interpolation (the points where the target function and its polynomial reconstruction meet) are denser in the allowed phase space in energy (the grey shaded area). This is most evident in the case of large \mathbf{q}^2 , as ω_{\min} is moved further away from 0. This is then the region where we expect larger deviations for the values of $\bar{X}^{(l)}(\mathbf{q}^2)$ between the two choices of ω_0 . Note also that a value slightly below ω_{\min} (e.g. $0.9\omega_{\min}$) safeguards against statistical fluctuations

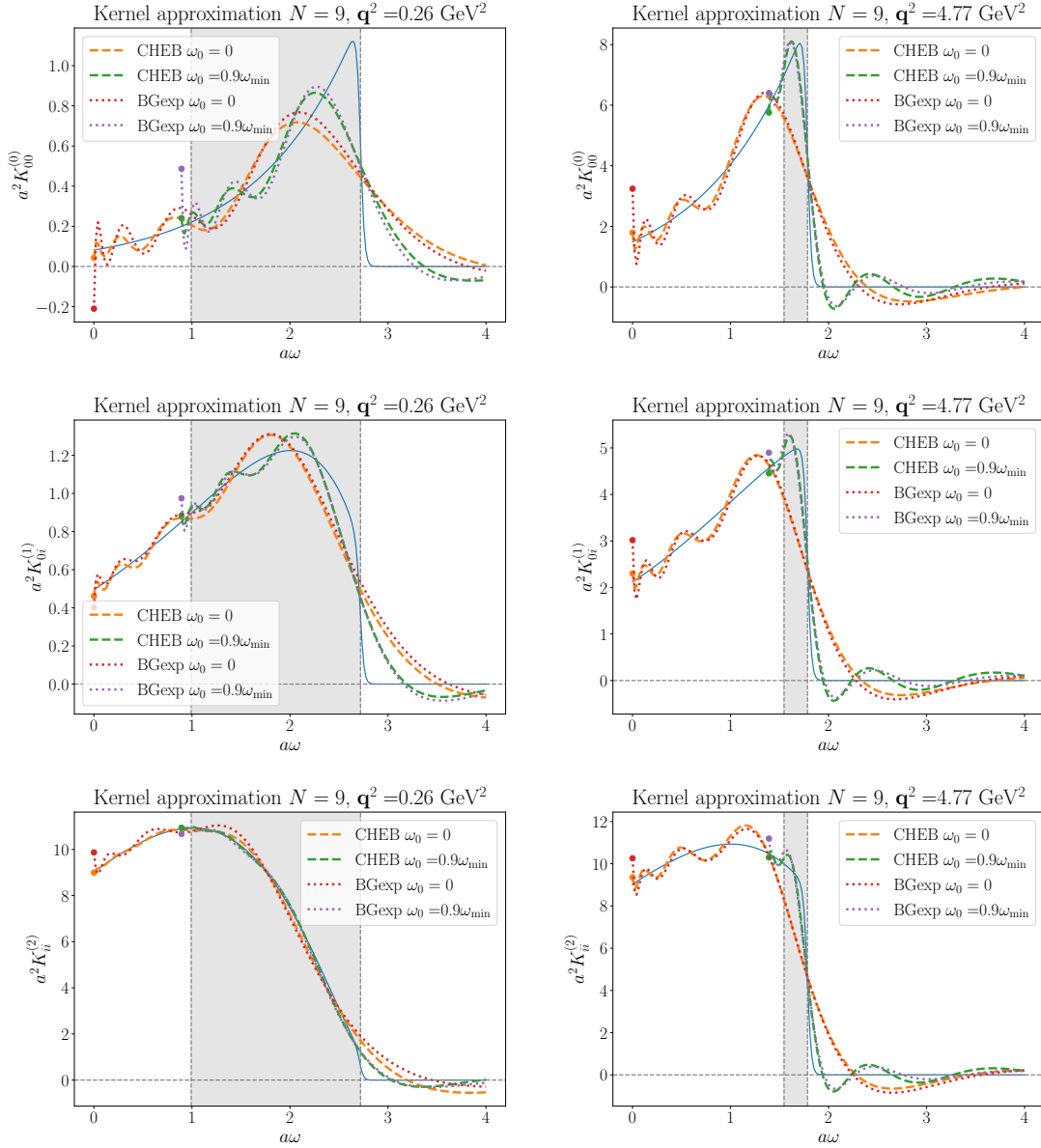


Figure 7. Polynomial approximation at order $N = 9$ of the kernel $K_{\mu\nu,\sigma}^{(l)}(\mathbf{q}, \omega; 2t_0)$, for $l = 0$ (first row), $l = 1$ (second row) and $l = 2$ (third row) with $t_0 = 1/2$ and $\sigma = 0.02$. The left column shows the case of the smallest $\mathbf{q}^2 = 0.26 \text{ GeV}^2$, whereas the right column shows one of the largest momentum $\mathbf{q}^2 = 4.77 \text{ GeV}^2$. The grey area corresponds to the kinematically allowed range $\omega_{\min} \leq \omega \leq \omega_{\max}$ for the given \mathbf{q}^2 . The solid lines show the target function; the dashed lines show the approximation with the Chebyshev approach, whereas the dotted ones show the approximation with Backus-Gilbert with an exponential base and $\lambda = 0$.

in the D_s -meson mass.

4.2 Chebyshev polynomials and Backus-Gilbert in practice

We now discuss the quality of the data analysis as outlined in Sec. 2.4. Focusing first on the Chebyshev-polynomial approach, the correlator data are traded with the fitted Chebyshev matrix elements as

$$\bar{C}_{\mu\nu}^{\text{fit}}(k) = \sum_{j=0}^k \tilde{a}_j^{(k)} \langle \tilde{T}_j \rangle_{\mu\nu}, \quad (4.1)$$

where the coefficients $\tilde{a}_j^{(k)}$ are given by the power representation of the Chebyshev polynomials, see App. A. Following (2.58), the kernel with fitted Chebyshev matrix elements can be written as

$$\langle K_{\sigma}^{(l)} \rangle_{\mu\nu} = \frac{\tilde{c}_{\mu\nu,0}^{(l)}}{2} \langle \tilde{T}_0 \rangle_{\mu\nu} + \sum_{k=1}^N \tilde{c}_{\mu\nu,k}^{(l)} \langle \tilde{T}_k \rangle_{\mu\nu} = \sum_{k=0}^N \tilde{c}_{\mu\nu,k}^{(l)} \bar{C}_{\mu\nu}^{\text{fit}}(k). \quad (4.2)$$

An example of the Chebyshev matrix elements obtained from the fits can be seen in Fig. 8, where we compare two different extractions according to the starting point of the approximation ω_0 . The plots show the distribution of each order of the Chebyshev matrix elements obtained through the fitting procedure described in C: each histogram plots values obtained for all the 1000 bootstrap bins. We show the axial channel $A_i A_i$, as its signal turns out to be particularly clean. In Fig. 9 we show results for the $A_i A_j$ channel, with $i \neq j$, which is found to be the noisiest channel. Here, only few terms can be determined meaningfully by the lattice data. Higher-order terms just follow the flat prior distribution in $[-1, 1]$.

In both cases we observe that a larger number of Chebyshev matrix elements can be determined meaningfully for $\omega_0 = 0$ than for $\omega_0 = 0.9\omega_{\text{min}}$. For example, in the $A_i A_i$ channel the distribution of the former is close to the prior distribution, which is flat between -1 and $+1$, for $N = 9$, whereas the latter start flattening at $N \gtrsim 7$. A possible explanation is as follows: as can be seen from (A.27), $\tilde{a}_j^{(k)}|_{\omega_0=0} = e^{-0.9\omega_{\text{min}}k} \tilde{a}_j^{(k)}|_{\omega_0=0.9\omega_{\text{min}}}$. The additional exponential factor largely cancels the ground-state exponential decay in the correlation function in Eq. (4.1). Hence, the polynomial approximation has less structure to describe and higher-order terms become less relevant. Nevertheless, in both cases the χ^2 of the fits are acceptable and the reconstruction of the data as in Eq. (4.1) gives comparable results.

We now move to the Backus-Gilbert case, for which we have so far only considered the limit $\lambda = 0$. In this limit the coefficients of the polynomial approximation are determined without reference to the data. We then consider the case $\lambda \neq 0$ and, by visual inspection of Fig. 10, find that the polynomial approximation of the kernel function gets worse. The effect of non-zero λ can be understood as a correction to the optimal coefficients, as outlined in Sec. B.2. In particular, if we rewrite the coefficients as $g_{\mu\nu,k}^{*(l)} = \gamma_{\mu\nu,k}^{(l)} + \epsilon_{\mu\nu,k}^{*(l)}$ we have

$$\langle K_{\sigma}^{(l)} \rangle_{\mu\nu} = \sum_{k=0}^N g_{\mu\nu,k}^{*(l)} \bar{C}_{\mu\nu}(k) = \sum_{k=0}^N \gamma_{\mu\nu,k}^{(l)} \bar{C}_{\mu\nu}(k) + \sum_{k=0}^N \epsilon_{\mu\nu,k}^{*(l)} \bar{C}_{\mu\nu}(k), \quad (4.3)$$

where $\gamma_{\mu\nu}^{(l)}$ are the coefficients for $\lambda = 0$ and $\epsilon_{\mu\nu,k}^{*(l)}$ is a correction which takes care of reducing the noise coming from the statistical error.

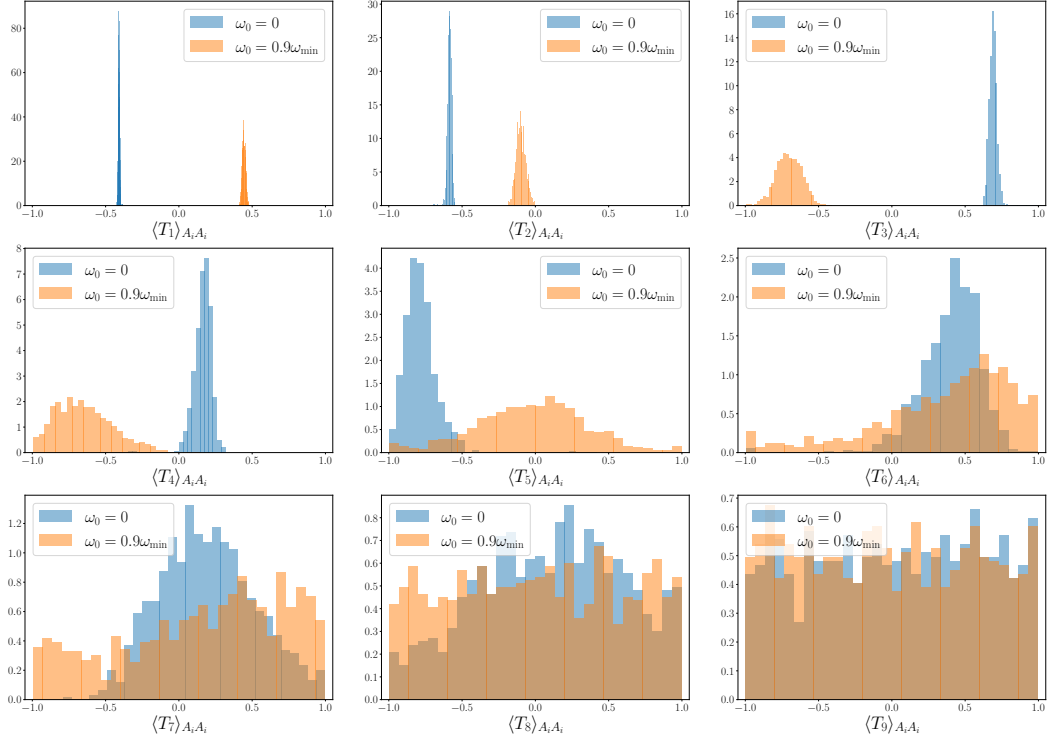


Figure 8. Histogram of the Chebyshev matrix elements $\langle \tilde{T}_k \rangle_{A_i A_i}$ for $k = 1, 2, \dots, N$ with $N = 9$ for two values $\omega_0 = 0$ (blue) and $\omega_0 = 0.9\omega_{\min}$ (orange) at $\mathbf{q}^2 = 0.26 \text{ GeV}^2$. The matrix element $\langle \tilde{T}_0 \rangle_{A_i A_i} = 1$ by definition and is therefore not shown. This channel is one of the most precise: we find that in both cases the fitting procedure is able to determine the matrix elements up to order $N \simeq 7$, after which the distribution of the bootstrap bins remains flat.

4.3 The inclusive decay rate

In this section we present the main results of our work. In Fig. 11 we show the results of $\bar{X}(\mathbf{q}^2)$ for all the simulated values of \mathbf{q}^2 . For each simulation point we show the results of three studied approaches, i.e., Chebyshev polynomials, exponential Backus-Gilbert and Chebyshev Backus-Gilbert, all of them for both $\omega_0 = 0$ and $\omega_0 = 0.9\omega_{\min}$. We find that all sets of three points for a given value of ω_0 agree very well. However, sets with different ω_0 start deviating as we increase the value of \mathbf{q}^2 . As discussed in the previous section, this can be understood in terms of the polynomial approximation of the kernel: as \mathbf{q}^2 increases, the phase space in ω shrinks, and the two approximations start differing increasingly. Our data indicates that the approximation improves as $\omega_0 \rightarrow \omega_{\min}$. In order for the approximations for different ω_0 to be comparable the order of the polynomial needs to be increased for lower ω_0 . It is also conceivable that other systematics like finite-volume or cutoff effects play a role here. These effects are beyond the scope of this work but will have to be addressed in future work.

In the previous section we have seen that the shape of the kernel, and hence, the quality of approximation, varies substantially for different l and \mathbf{q}^2 . The degree to which this impacts the combined result $\bar{X}(\mathbf{q}^2)$ depends on the magnitude of each contribution, as

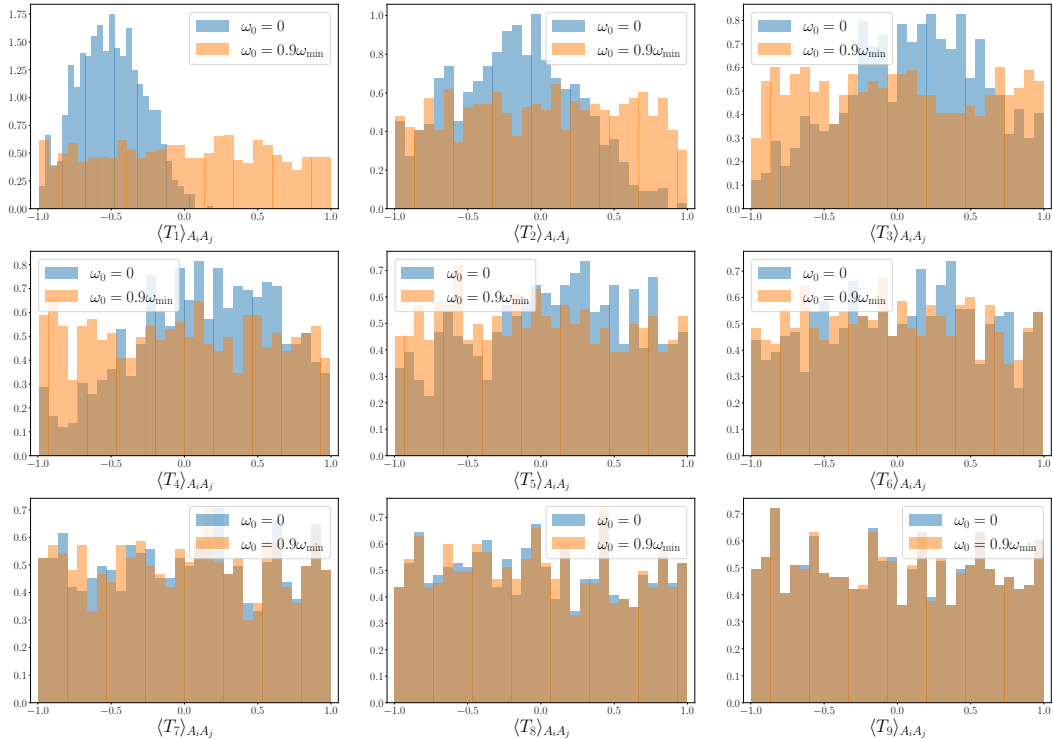


Figure 9. Histogram of the Chebyshev matrix elements $\langle \tilde{T}_k \rangle_{A_i A_j}$ with $i \neq j$ for $k = 1, 2, \dots, N$ with $N = 9$ for two values $\omega_0 = 0$ (blue) and $\omega_0 = 0.9\omega_{\min}$ (orange) at $\mathbf{q}^2 = 0.26 \text{ GeV}^2$. The results for $\langle \tilde{T}_k \rangle_{A_i A_j}$ are less well constrained than the ones for $A_i A_i$ shown in Fig. 8. The minimum of the χ^2 is determined almost entirely by the uniform priors.

illustrated in Fig. 12. The plots indicate that the largest contribution originates from the channel with $l = 2$. The underlying kernel is, at least for smaller values of \mathbf{q}^2 , relatively smooth (Fig. 7). We therefore expect less sensitivity to the systematics of the polynomial approximation in this kinematical region but more care is needed for larger \mathbf{q}^2 .

We now address the stability against the order of the polynomial N . Starting from the Chebyshev approach, we study the saturation in Fig. 13. We start from the fit with $N = 9$. The plot shows the result where the first k Chebyshev matrix elements (cf. legend) are taken from the fit, and the remaining $N - k$ are replaced by a flat distribution $-1 \leq \langle \tilde{T}_j \rangle_{\mu\nu} \leq 1$ with $j = k + 1, \dots, N$. We can see that the signal is dominated by small orders; for $\omega_0 = 0$, the signal is saturated at around $N \simeq 5$, whereas for $\omega_0 = 0.9\omega_{\min}$ saturation starts at $N \simeq 3$. This is also compatible with the previous discussion on the fit of the Chebyshev matrix elements, cf. with Fig. 8 and Fig. 9.

In order to estimate higher-order contributions, which are not constrained by our data, we study how the results change after adding more terms in the Chebyshev distributions on top of the $N = 9$ available. In this way we obtain an estimate of the approximation up to $N = 50$, as in Fig. 14. We show in particular the case of distributions with random values in $\mathbb{Z}_2 = \{-1, 1\}$ for $\langle \tilde{T}_k \rangle_{\mu\nu}$ beyond $k = 9$; the case with uniform distribution with values in $[-1, +1]$ gives similar results with slightly smaller errors. In both cases, the extra

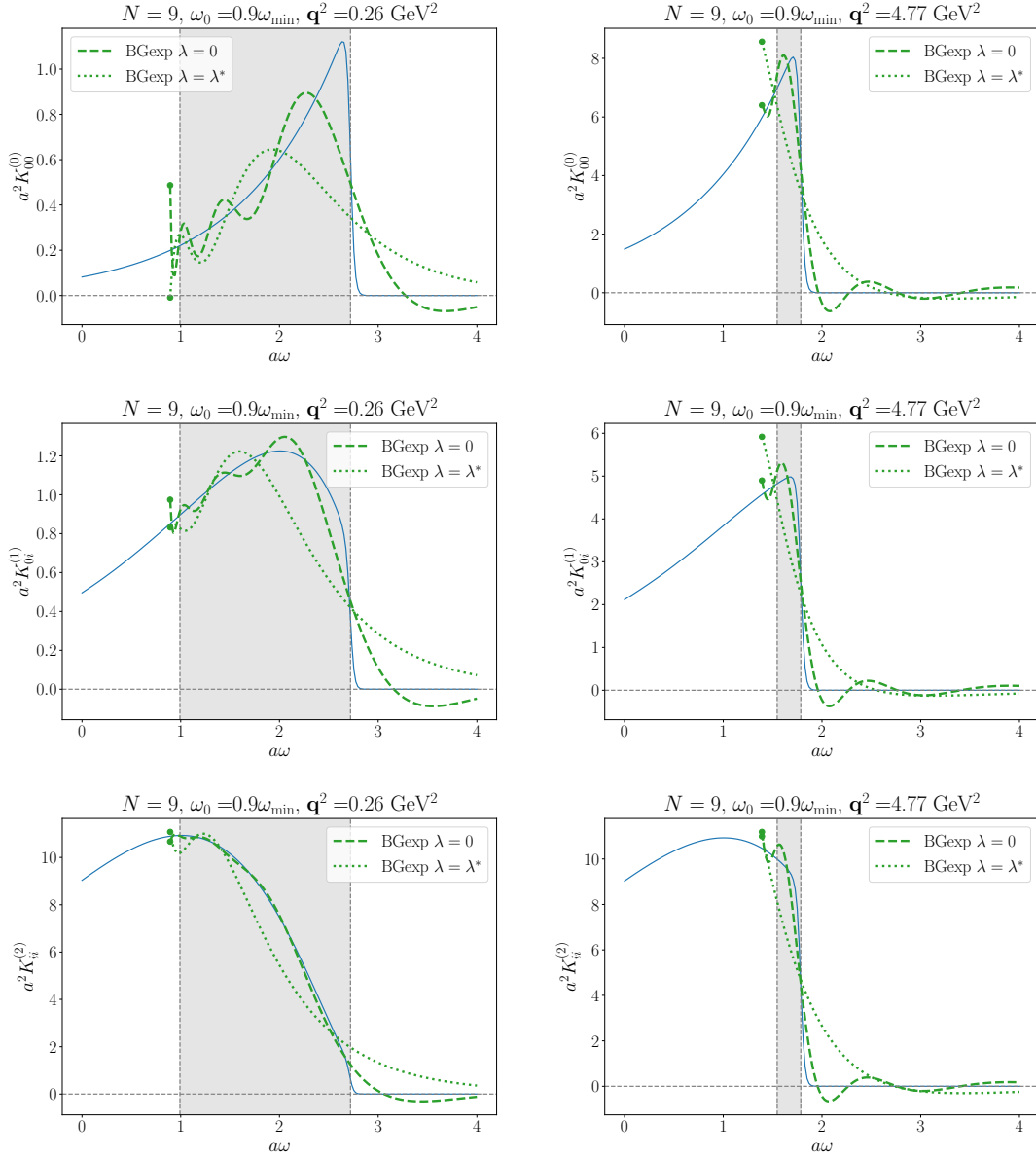


Figure 10. Polynomial approximation of the kernel $K_{\mu\nu,\sigma}^{(l)}(\mathbf{q}, \omega; 2t_0)$, for $l = 0$ (first row), $l = 1$ (second row) and $l = 2$ (third row) with $t_0 = 1/2$ and $\sigma = 0.02$ in the case of Backus-Gilbert with exponential basis and $\lambda \neq 0$. The value of λ has been chosen to be λ^* for each plot, which gives equal weight to the statistical and systematic errors.

terms contribute to the final error only mildly: these observations suggest that the results obtained do not suffer from huge systematic error from the polynomial approximation. A more complete study is however required for a reliable estimate of the underlying systematic effects.

Concerning the Backus-Gilbert method, we investigate the stability around the chosen value of λ^* , obtained with the prescription of Sec. 2.4.2. We focus in particular on the

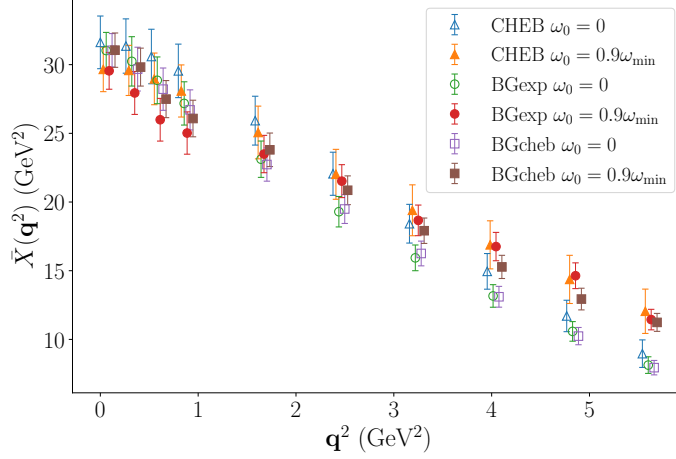


Figure 11. Estimate of $\bar{X}(q^2)$ with the two different strategies for 10 different q^2 with $N = 9$ and $q_{\max}^2 = 5.86 \text{ GeV}^2$.

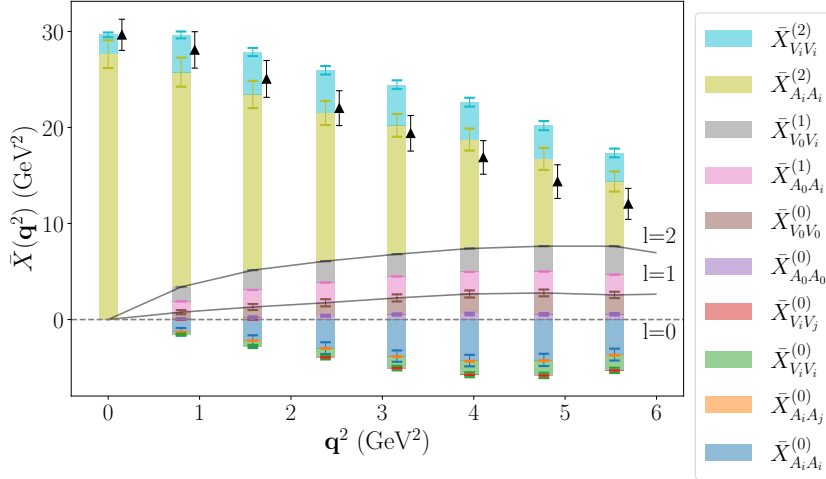


Figure 12. Contributions to $\bar{X}(q)$ from the Chebyshev-polynomial approach at $N = 9$ and $\omega_0 = 0.9\omega_{\min}$ with associated error bars. The black triangles correspond to the final value $\bar{X}(q^2) = \sum_{l=0}^2 \sum_{\{\mu,\nu\}} \bar{X}_{\mu\nu}^{(l)}(q^2)$. The solid black lines separate the contributions from $l = 0$ (bottom), $l = 1$ (middle) and $l = 2$ (top).

channel $\bar{X}_{A_i A_i}^{(2)}$ as it is the one responsible for the largest contribution. The plot is shown in Fig. 15. We can see that for small q^2 the value of $\bar{X}(q)$ is stable, which implies that statistical and systematic errors are well balanced. For larger q^2 the situation is more delicate: this can be understood in terms of the reduced phase space in ω , as shown for example in Fig. 10. A first attempt at mitigating the induced systematic effect could be to identify the region where the two Backus-Gilbert approaches with different bases are consistent, to identify (where possible) a plateau, and to estimate a value inside such region.

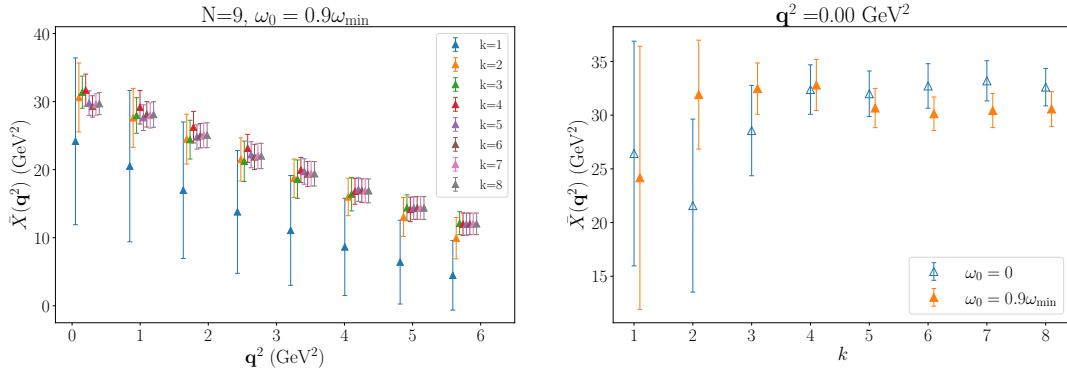


Figure 13. Saturation of Chebyshev polynomial approach for different q^2 and $\omega_0 = 0.9\omega_{\min}$ (left) and for case $q^2 = 0$ for both values of ω_0 as a function of k (right), where k is the number of Chebyshev matrix elements taken from the fit.

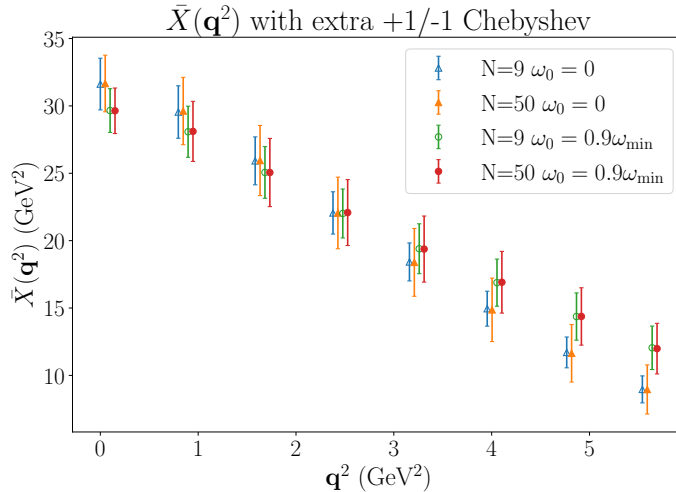


Figure 14. Saturation of Chebyshev polynomial approach, where $N = 9$ is the reference case, and for $N = 50$ higher-order terms are sampled from a \mathbb{Z}_2 distribution.

In the r.h.s. plot of Fig. 15 we see, however, that this is not always the case: there is no clear plateau region for λ . Interestingly, the statistical error of the Chebyshev approach turns out more conservative in this case, and compatible with the result one would obtain from Backus-Gilbert. More generally, apart from the absence of a plateau region in some cases, both choices of polynomial basis are consistent between themselves and with the Chebyshev-polynomial approach.

Coming back to the decay rate, to extract the final result we perform a polynomial fit of degree two on $\bar{X}^{(l)}(q^2)/(\sqrt{q^2})^{2-l}$. The final result is then obtained integrating these results in the physical range in q . Since this is a qualitative study, we don't report any final number; however, the result obtained here seems to be in the right ballpark if compared with the B_s meson decay rate. Furthermore, all the approaches give compatible results,

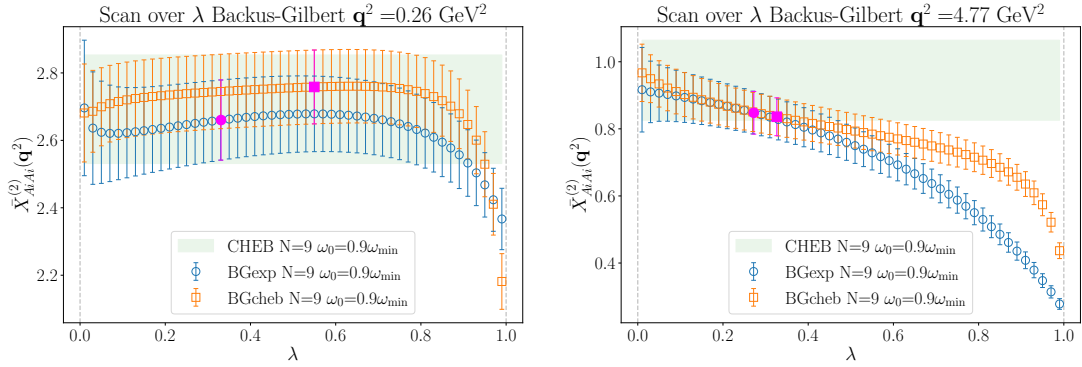


Figure 15. Scan over λ for $q^2 = 0.26 \text{ GeV}^2$ (left) and $q^2 = 4.77 \text{ GeV}^2$ (right) for the Backus-Gilbert method with exponential and Chebyshev basis for $\bar{X}_{A_i A_i}^{(2)}$ with $\omega_0 = 0.9 \omega_{\min}$. The green shaded line is the reference from the Chebyshev; the magenta points correspond to the choice of λ^* . Note that the points $\lambda = 0$ and $\lambda = 1$ (vertical grey dashed lines) are not included in this plot.

and the final statistical error is of order 5%.

We now address similarities and differences between the two approaches. The calculation of $\bar{X}(q^2)$ aims to improve accuracy by combining the naive polynomial approximation with a correction term $\delta\bar{X}(q^2)$ that accounts for variance reduction, i.e.,

$$\bar{X}(q^2) = \bar{X}^{\text{naive}}(q^2) + \delta\bar{X}(q^2), \quad (4.4)$$

where $\bar{X}^{\text{naive}}(q^2)$ would correspond to (2.32). The correction term is specific to the adopted strategy and is given by:

- $\delta\bar{X}^{\text{CHEB}}(q^2) = C_{\mu\nu}(2t_0) \sum_{k=0}^N \bar{c}_{\mu\nu,k} \delta\bar{C}_{\mu\nu}(k)$, for the Chebyshev polynomials technique, where $\delta\bar{C}_{\mu\nu}(k) = \bar{C}_{\mu\nu}(k) - \bar{C}_{\mu\nu}^{\text{fit}}(k)$;
- $\delta\bar{X}^{\text{BG}}(q^2) = C_{\mu\nu}(2t_0) \sum_{k=0}^N \epsilon_{\mu\nu,k}^* \bar{C}_{\mu\nu}(k)$, for the Backus-Gilbert method, which corrects the coefficients of the polynomial approximation as in (4.3).

In both cases, $\delta\bar{X}(q^2)$ can be interpreted as a noisy zero that does not impact the naive calculation but helps with variance reduction. This is represented in Fig. 16, which shows the statistical error on \bar{X} with and without the correction term. The reduction in statistical error is substantial. Additionally, the magnitude of the correction varies depending on ω_0 , where larger values result in a greater increase in $|\delta\bar{X}(q^2)|$ as q^2 increases.

To conclude this section we discuss some of the aspects we neglected for the purpose of this study. In particular, all the results presented here have been obtained with kernels smeared by a sigmoid with a fixed $\sigma = 0.02$. Eventually however, one will first have to first take the infinite-volume and continuum limits, followed by an extrapolation to $\sigma \rightarrow 0$. Exemplarily though, we show the σ dependence at finite lattice spacing and volume in Fig. 17. There, one sees that for our setup and statistical precision the dependence on σ is mild. There is an indication that it might be more pronounced for larger q^2 . We argue that here the extrapolation in σ is quite delicate and could lead to misleading results.

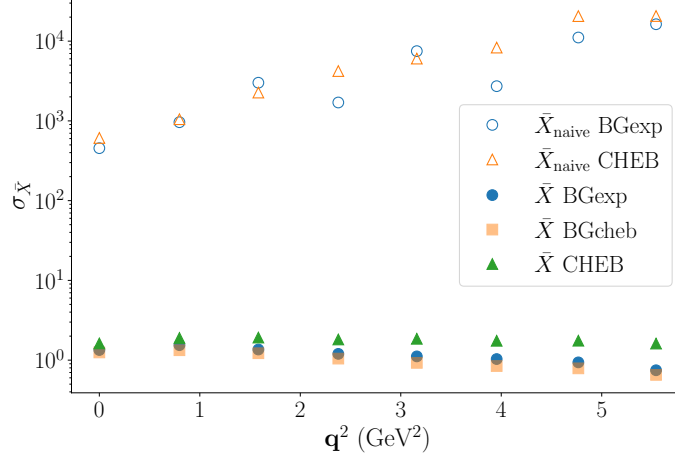


Figure 16. Effect of the variance reduction to $\bar{X}_{\text{naive}}(\mathbf{q}^2)$ from the correction $\delta\bar{X}(\mathbf{q}^2)$ as in Eq. (4.4) for $\omega_0 = 0.9\omega_{\text{min}}$ and $N = 9$. The y axis shows the standard deviation $\sigma_{\bar{X}}$ for $\bar{X}_{\text{naive}}(\mathbf{q}^2)$ (empty symbols) and $\bar{X}(\mathbf{q}^2) = \bar{X}_{\text{naive}}(\mathbf{q}^2) + \delta\bar{X}(\mathbf{q}^2)$ (filled symbols).

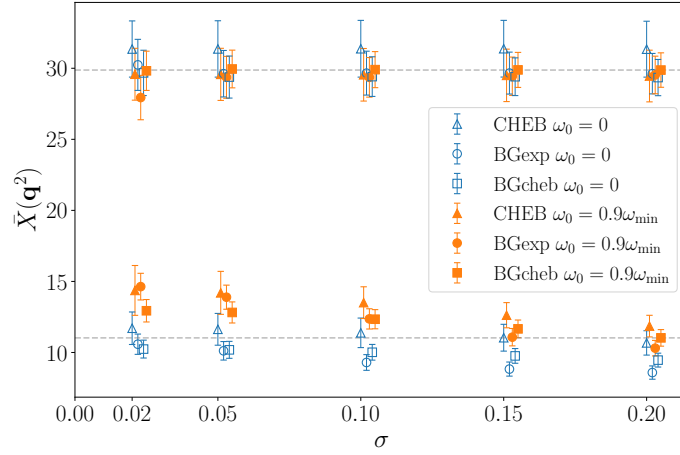


Figure 17. Dependence on the smearing parameter σ of $\bar{X}(\mathbf{q}^2)$ for all the approaches at $N = 9$ in the case $\mathbf{q}^2 = 0.26 \text{ GeV}^2$ (top) and $\mathbf{q}^2 = 4.77 \text{ GeV}^2$ (bottom). The horizontal dashed lines correspond to the central values for the choice $\sigma = 0.02$ with $\omega_0 = 0.9\omega_{\text{min}}$.

Indeed, increasing values of sigma would result in kernel functions quite different from the target ones; on the other side, differences in small values of σ will not be captured by a polynomial approximation with small value of N , as small deviations would be noticeable only for higher degrees of approximations.

4.4 The inclusive decay rate in the ground-state limit

We now study the ground-state limit of the inclusive approach as discussed in Sec. 2.2, which provides for a cross-check of the inclusive-decay analysis strategies. The four-point

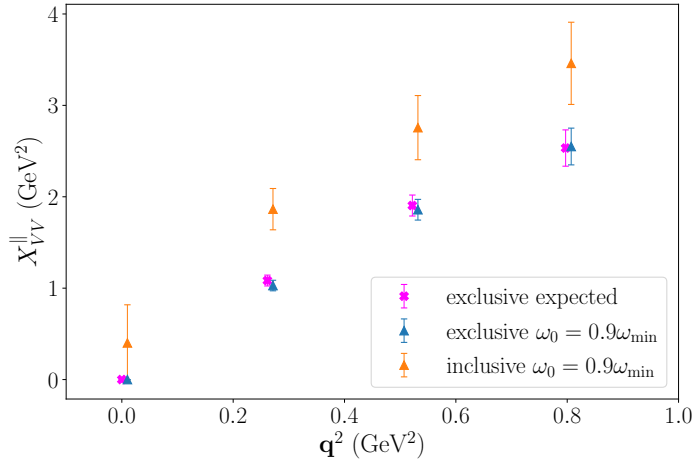


Figure 18. Ground-state limit. The “exclusive” labels refer to the data built from the three-point correlators as in Eq. (4.5), whereas the “inclusive” label refers to the full inclusive data analysis starting from the four-point correlation functions. The analysis has been performed using the Chebyshev approach.

function representing the ground state can be constructed with input from lattice data for the exclusive decay $B_s \rightarrow D_s l \nu_l$. In particular, restricting the discussion to the vector channel VV , the ground-state correlator

$$C_{\mu\nu}^G(t) = \frac{1}{4M_{B_s}E_{D_s}} \langle B_s | V_\mu^\dagger | D_s \rangle \langle D_s | V_\nu | B_s \rangle e^{-E_{D_s}t}, \quad (4.5)$$

can be constructed from lattice data for the ratio of three-point and two-point functions

$$R_{B_s D_s, \mu}(t; \mathbf{q}) = \sqrt{4M_{B_s}E_{D_s}} \sqrt{\frac{C_{B_s D_s, \mu}^{SS}(\mathbf{q}, t_{\text{snk}}, t, t_{\text{src}}) C_{D_s B_s, \mu}^{SS}(\mathbf{q}, t_{\text{snk}}, t, t_{\text{src}})}{C_{B_s}^{SS}(t_{\text{snk}}, t_{\text{src}}) C_{D_s}^{SS}(\mathbf{q}, t_{\text{snk}}, t_{\text{src}})}}, \quad (4.6)$$

which converges to $\mathcal{M}_\mu \equiv \langle D_s | V_\mu | B_s \rangle$ for $t \gg t_{\text{src}}$ and $t < t_{\text{snk}}$. The matrix element can be decomposed into form factors

$$\mathcal{M}_\mu = f_+(q^2)(p_{B_s} + p_{D_s})_\mu + f_-(q^2)(p_{B_s} - p_{D_s})_\mu. \quad (4.7)$$

Recalling that we assume $\mathbf{p}_{B_s} = \mathbf{0}$, we then extract $f_+(q^2)$ from a constant fit to the combination

$$R_{f_+}(t; \mathbf{q}) \stackrel{q \neq 0}{=} \frac{1}{2M_{B_s}} \left(R_{B_s D_s, 0}(t; \mathbf{q}) + (M_{B_s} - E_{D_s}) \frac{\sum_{i=1}^3 R_{B_s D_s, i}(t; \mathbf{q})}{\sum_{i=1}^3 q_i} \right), \quad (4.8)$$

which converges to $f_+(q^2)$ as $R_{B_s D_s, \mu}(t; \mathbf{q}) \rightarrow \mathcal{M}_\mu$. We consider only the three smaller momenta to test the approach, as the signal-to-noise deteriorates rapidly with larger q^2 .

The result of the inclusive analysis for the channel \bar{X}_{VV}^{\parallel} is reported in Fig. 18. In particular, we compare the expected value (2.18) from the extracted values of $f_+(q^2)$ with the inclusive analysis performed using the mock data $C_{\mu\nu}^G$ and the real data $C_{\mu\nu}$. Note

that for the mock data the normalised correlator corresponds simply to $\bar{C}_{\mu\nu}^G(t) = e^{-E_{D_s}t}$ by construction.

We find excellent agreement between the results from the conventional analysis for exclusive decay on the one side, and the one based on ground-state saturation, but using the full analysis chain adopted for the inclusive decay, on the other side. This provides a strong test of the analysis method for inclusive decay discussed in this paper. The results for the full inclusive decay on the other hand differ significantly from the exclusive case: while future studies will have to establish to which extent this could be down to systematics like finite-volume or cutoff effects, the magnitude of the effect makes appear likely to be to a large part due to contributions from the tower of finite states contributing to the inclusive decay. In particular, the deviation is expected to be larger for smaller q^2 , as the available phase space in ω is larger and may include more excited states.

5 Conclusions and outlook

In this work, we have presented a full and flexible setup for studying inclusive semileptonic decays in lattice QCD, focusing in particular on $B_{(s)}$ mesons. We incorporate and compare Chebyshev polynomials and the Backus-Gilbert method, both of which enable efficient and accurate calculations of the total decay rate. In particular, we improved the Chebyshev polynomial technique through the use of a generic set of shifted polynomials in $e^{-\omega}$, and we refined the statistical analysis with a bootstrap method, fully accounting for the bounds $[-1, 1]$. We also showed how the result depends on the number of Chebyshev matrix elements and presented a possible way to take the limit $N \rightarrow \infty$ to address the systematics associated with the polynomial approximation. On the Backus-Gilbert side, we introduced a generalisation of the method of [22] to allow for the use of arbitrary bases of polynomials.

The two methods have been shown to be compatible, and the final results for the decay rate are in agreement. We compared how the two techniques deal with the variance reduction of the final observable: the Chebyshev polynomials' approach relies on trading the data with Chebyshev matrix elements that fully account for the bounds, whereas the Backus-Gilbert method achieves the same goal by modifying the coefficients of the polynomial approximation to reduce the statistical error. We also studied the ground-state limit, which offered a cross-check of the inclusive analysis technique and outlined the effect of excited states in the inclusive decay with respect to the corresponding exclusive process $B_s \rightarrow D_s l\nu_l$.

Overall, our work provides a solid foundation for future studies with these techniques. However, there are still several areas that require further investigation, including systematic errors associated with the polynomial approximation, finite-volume effects, discretisation errors, and the continuum limit. We intend to address these issues in future works, repeating the computations on more ensembles and also addressing similar processes involving D_s mesons, which offer a more controlled environment. Additionally, we plan to explore alternative observables such as hadronic and lepton moments to compare with experimental data and to gain a deeper understanding of the ground state limit, which may provide useful insight on the physics contributing in such processes.

Acknowledgments

This work used the DiRAC Extreme Scaling service at the University of Edinburgh, operated by the Edinburgh Parallel Computing Centre on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BEIS capital funding via STFC capital grant ST/R00238X/1 and STFC DiRAC Operations grant ST/R001006/1. DiRAC is part of the National e-Infrastructure. A.B. is a JSPS International Research Fellows and received funding from the "JSPS Postdoctoral Fellowship for Research in Japan (Short-term)" and is supported by the Mayflower scholarship in the School of Physics and Astronomy of the University of Southampton. The work of S.H. and T.K. is supported in part by JSPS KAKENHI Grant Number 22H00138 and 21H01085 respectively and by the Post-K and Fugaku supercomputer project through the Joint Institute for Computational Fundamental Science (JICFuS).

A Chebyshev polynomials

We summarise here important properties of the standard Chebyshev polynomials relevant for this work and in particular the generalisation for the shifted version extensively used in the analysis. We refer to other sources [63] for more details.

A.1 Standard polynomials

The standard Chebyshev polynomials of the first kind are defined as

$$T_k : [-1, 1] \rightarrow [-1, 1], \quad T_k(x) = \cos(k \cos^{-1}(x)), \quad k \in \mathbb{N}. \quad (\text{A.1})$$

They are orthogonal with respect the scalar product

$$\int_{-1}^1 T_r(x) T_s(x) \Omega(x) dx = \frac{\pi}{2} \delta_{rs} (1 + \delta_{r0}), \quad (\text{A.2})$$

where $\Omega(x) = 1/\sqrt{1-x^2}$ is a weight function. Their polynomial expansion in x^k is given by

$$T_n(x) = \sum_{k=0}^n t_k^{(n)} x^k, \quad (\text{A.3})$$

with

$$\begin{cases} t_0^{(n)} = (-1)^{n/2} & \text{if } n \text{ even} \\ t_k^{(n)} = 0 & \text{if } n - k \text{ odd} \\ t_k^{(n)} = (-1)^{(n-k)/2} 2^{k-1} \frac{n}{n+k} \binom{\frac{n+k}{2}}{\frac{n-k}{2}} & \text{if } k \neq 0 \text{ and } n - k \text{ even} \end{cases}. \quad (\text{A.4})$$

A useful property involves the representation of x^n in terms of the standard Chebyshev polynomial

$$p_n(x) \equiv x^n = 2^{1-n} \sum_{\substack{k=0 \\ n-k \text{ even}}}^n \binom{n}{\frac{n-k}{2}} T_k(x), \quad (\text{A.5})$$

where the prime indicates that the first term is halved.

A.1.1 Expansion in Chebyshev polynomials

Chebyshev polynomials provide the best approximation of the function $f : [-1, 1] \rightarrow \mathbb{R}$ to any given order N in terms of the L_∞ -norm. In other words, the *minmax* error, i.e. the maximum difference between the target function and the reconstructed one, is minimised. In particular, for the functions considered in this work, it is guaranteed that the Chebyshev approximation converges when $N \rightarrow \infty$. The polynomial approximation reads

$$f(x) \simeq \frac{1}{2}c_0T_0(x) + \sum_{k=1}^N c_kT_k(x), \quad c_k = \frac{2}{\pi} \int_{-1}^1 dx f(x)T_k(x)\Omega(x), \quad (\text{A.6})$$

where we recall that $T_0(x) = 1$ by definition. The coefficients are given by the projection of the target function f on the basis of Chebyshev polynomials.

A.2 Shifted Chebyshev polynomials

In general, for the purpose of this work we consider generic functions $f(x)$ defined in an interval $[a, b]$, which we want to approximate with Chebyshev polynomials in e^{-x} . To this end we can define shifted polynomials $\tilde{T}_n(x)$ with $x \in [a, b]$, such that their domain matches the one of the target function. The relation to the standard polynomials is given by

$$\tilde{T}_k(x) = T_k(h(x)), \quad (\text{A.7})$$

where $h : [a, b] \rightarrow [-1, 1]$ is an invertible function that maps the new domain into the domain of the standard Chebyshev polynomials,

$$h(x) = Ae^{-x} + B. \quad (\text{A.8})$$

The coefficients A and B can be determined by imposing $h(a) = -1$ and $h(b) = +1$, for which one obtains

$$A = -\frac{2}{e^{-a} - e^{-b}}, \quad B = \frac{e^{-a} - e^{-b}}{e^{-a} + e^{-b}}. \quad (\text{A.9})$$

The orthogonality relation for the shifted polynomials reads

$$\int_a^b dx \tilde{T}_r(x)\tilde{T}_s(x)\Omega_h(x) = \int_a^b dx T_r(h(x))T_s(h(x))\Omega_h(x) \quad (\text{A.10})$$

where $\Omega_h(x)$ is the new weight for the shifted \tilde{T}_k , which depends on the map h . To show that this recover the original integral in Eq. (A.2), we set $x = h^{-1}(y)$ and $dx = \frac{1}{h'(h^{-1}(y))}dy$ and get

$$\int_{h(a)}^{h(b)} dy T_r(y)T_s(y) \frac{\Omega_h(h^{-1}(y))}{h'(h^{-1}(y))}; \quad (\text{A.11})$$

choosing then

$$\Omega_h(x) = \Omega(h(x))|h'(x)|, \quad (\text{A.12})$$

we finally obtain

$$\int_a^b dx \tilde{T}_r(x) \tilde{T}_s(x) \Omega_h(x) = \int_{-1}^1 dy T_r(y) T_s(y) \Omega(y). \quad (\text{A.13})$$

We can also generalise the polynomial expressions and their properties. The polynomial representation reads

$$\tilde{T}_n(x) = \sum_{j=0}^n t_j^{(n)} h(x)^j = \sum_{j=0}^n t_j^{(n)} (Ae^{-x} + B)^j = \sum_{j=0}^n t_j^{(n)} \sum_{k=0}^j \binom{j}{k} A^k B^{j-k} e^{-kx}. \quad (\text{A.14})$$

We can expand this sum explicitly and re-sum it in order to isolate the coefficients of e^{-kx} . We obtain

$$\tilde{T}_n(x) = \sum_{k=0}^n \tilde{t}_k^{(n)} e^{-kx}, \quad \tilde{t}_k^{(n)} = A^k \sum_{j=k}^n \binom{j}{k} t_j^{(n)} B^{j-k} = \left(\frac{A}{B}\right)^k \sum_{j=k}^n \binom{j}{k} t_j^{(n)} B^j. \quad (\text{A.15})$$

In a similar way we can generalise the power representation as

$$\tilde{p}_n(x) \equiv h(x)^n = 2^{1-n} \sum_{\substack{j=0 \\ n-j \text{ even}}}^n \binom{n}{\frac{n-j}{2}} \tilde{T}_j(x), \quad x \in [a, b]. \quad (\text{A.16})$$

Using

$$\tilde{p}_n(x) = (Ae^{-x} + B)^n = \sum_{k=0}^n \binom{n}{k} A^k B^{n-k} e^{-kx}, \quad (\text{A.17})$$

and starting from $\tilde{p}_0 = 1$ we can work out iteratively the general expression for e^{-nx} as

$$e^{-nx} = \frac{1}{A^n} \left[\tilde{p}_n(x) - \sum_{k=0}^{n-1} \binom{n}{k} A^k B^{n-k} e^{-kx} \right]. \quad (\text{A.18})$$

We can finally collect the numerical coefficients and rewrite everything in terms of the shifted Chebyshev polynomials as

$$e^{-nx} = \sum_{j=0}^n \tilde{a}_j^{(n)} \tilde{T}_j(x). \quad (\text{A.19})$$

The set of coefficients $\{\tilde{a}_0^{(n)}, \tilde{a}_1^{(n)}, \dots, \tilde{a}_n^{(n)}\}$ can be easily found numerically for each value of n .

A.2.1 Expansion in Chebyshev polynomials with exponential map

We have now all the elements necessary to proceed with the polynomial approximation of a generic function in e^{-x} . For the purpose of this work we will restrict ourselves to the case $f : [x_0, \infty) \rightarrow \mathbb{R}$. In particular, the approximation is now

$$f(x) = \frac{1}{2} \tilde{c}_0 \tilde{T}_0(x) + \sum_{k=1}^N \tilde{c}_k \tilde{T}_k(x), \quad \tilde{c}_k = \frac{2}{\pi} \int_{\omega_0}^{\infty} dx f(x) \tilde{T}_k(x) \Omega_h(x). \quad (\text{A.20})$$

The coefficients can be rewritten more explicitly as

$$\tilde{c}_k = \frac{2}{\pi} \int_0^\pi d\theta f(h^{-1}(\cos \theta)) \cos(k\theta) = \frac{2}{\pi} \int_0^\pi d\theta f\left(-\ln\left(\frac{\cos \theta - B}{A}\right)\right) \cos(k\theta). \quad (\text{A.21})$$

The last equality follows from setting $y = h(x)$ and inverting

$$x = h^{-1}(y) = -\log\left(\frac{y - B}{A}\right). \quad (\text{A.22})$$

In this case, the coefficients A and B are given by

$$A = -2e^{x_0}, \quad B = 1. \quad (\text{A.23})$$

A.2.2 Matrix relations

In this subsection we illustrate some useful properties that arise when setting $x_0 \neq 0$, assuming the domain of the target function in $[x_0, \infty)$. We then explicitly consider $B = 1$ and $A = -2e^{-x_0}$ to simplify the treatment, but the following discussion can be generalized trivially. We start expressing (A.15) in matrix notation

$$\begin{pmatrix} \tilde{T}_0(x) \\ \tilde{T}_1(x) \\ \vdots \\ \vdots \\ \tilde{T}_n(x) \end{pmatrix} = \begin{pmatrix} \tilde{t}_0^{(0)} & 0 & \cdots & \cdots & 0 \\ \tilde{t}_0^{(1)} & \tilde{t}_1^{(1)} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ \tilde{t}_0^{(n)} & \tilde{t}_1^{(n)} & \cdots & \cdots & \tilde{t}_n^{(n)} \end{pmatrix} \begin{pmatrix} 1 \\ e^{-x} \\ \vdots \\ \vdots \\ e^{-nx} \end{pmatrix}, \quad (\text{A.24})$$

and (A.19) as

$$\begin{pmatrix} 1 \\ e^{-x} \\ \vdots \\ \vdots \\ e^{-nx} \end{pmatrix} = \begin{pmatrix} \tilde{a}_0^{(0)} & 0 & \cdots & \cdots & 0 \\ \tilde{a}_0^{(1)} & \tilde{a}_1^{(1)} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ \tilde{a}_0^{(n)} & \tilde{a}_1^{(n)} & \cdots & \cdots & \tilde{a}_n^{(n)} \end{pmatrix} \begin{pmatrix} \tilde{T}_0(x) \\ \tilde{T}_1(x) \\ \vdots \\ \vdots \\ \tilde{T}_n(x) \end{pmatrix}. \quad (\text{A.25})$$

It is clear that these $(n+1) \times (n+1)$ matrices $\tilde{\mathbf{t}}$ with $(\tilde{\mathbf{t}})_{ij} = \tilde{t}_j^{(i)}$ and $\tilde{\mathbf{a}}$ with $(\tilde{\mathbf{a}})_{ij} = \tilde{a}_j^{(i)}$ are one the inverse of the other, i.e. $\tilde{\mathbf{a}} = (\tilde{\mathbf{t}})^{-1}$ and vice versa. From (A.15) we can further decompose $\tilde{\mathbf{t}}^{(n)}$ as

$$\tilde{\mathbf{t}} = \mathbf{A} \mathbf{P} \mathbf{t}, \quad (\text{A.26})$$

where $\mathbf{A}_{kk} = A^k = (-2e^{x_0})^k$ is a diagonal matrix, $\mathbf{P}_{jk} = \binom{j}{k}$ is the lower triangular Pascal matrix and the matrix \mathbf{t} follows from (A.4). This expression makes it easy to see the effect of x_0 : considering $\mathbf{A}_{kk} \Big|_{x_0 \neq 0} = e^{x_0 k} \mathbf{A}_{kk} \Big|_{x_0 = 0}$ it follows that

$$\left. (\tilde{\mathbf{t}})_{nk} \right|_{x_0 \neq 0} = \left. \tilde{t}_k^{(n)} \right|_{x_0 \neq 0} = e^{x_0 n} \left. \tilde{t}_k^{(n)} \right|_{x_0 = 0}, \quad \left. (\tilde{\mathbf{a}})_{nk} \right|_{x_0 \neq 0} = \left. \tilde{a}_k^{(n)} \right|_{x_0 \neq 0} = e^{-x_0 n} \left. \tilde{a}_k^{(n)} \right|_{x_0 = 0}. \quad (\text{A.27})$$

B Generalised Backus-Gilbert

In this appendix we reformulate and generalise the modified Backus-Gilbert approach proposed in [22, 38, 39]. The idea is to provide a more general framework which allows for the use of an arbitrary basis and to explore the properties and numerical advantages of different choices.

B.1 The method

The problem we want to address is the evaluation of a generic observable O of the form

$$O = \int_a^b d\omega \rho(\omega) K(\omega), \quad (\text{B.1})$$

where $K(\omega)$ is a function we will refer to as *kernel* and $\rho(\omega)$ is the spectral function related to a given correlation function

$$C(t) = \int_a^b d\omega \rho(\omega) e^{-\omega t}. \quad (\text{B.2})$$

While typically the range of integration is $a = 0$ and $b = \infty$, here we chose to leave it generic to keep the discussion general. The idea to address the computation is to approximate the kernel in polynomial up to some degree N , i.e. $K(\omega) = \sum_{j=0}^N g_j e^{-\omega j}$, such that the target observable can be estimated as

$$O \simeq \sum_{j=0}^N g_j \int_a^b d\omega \rho(\omega) e^{-\omega j} = \sum_{j=0}^N g_j C(j). \quad (\text{B.3})$$

For example, a typical problem consists in the extraction of the spectral density of a correlator, in which case one would consider the kernel to be a smoothed Dirac delta $K(\omega) = \delta_\sigma(\omega)$ with a finite width σ , as for example a Gaussian.

The approach consists of weighting the two functionals $A[g]$ and $B[g]$ against each other, where the first one provides a measure for the systematic effects coming from the polynomial approximation, and the second one provides a measure for the variance σ_O^2 of the observable O , in particular, $B[g] = \sigma_O^2 = g_i \sigma_{ij} g_j$, where we defined $\sigma_{ij} = \text{Cov}[C(i), C(j)]$. This is equivalent to solving a minimisation problem with constraints. We can then define a new functional F_θ as

$$F_\theta[g] = A[g] + \theta^2 B[g], \quad (\text{B.4})$$

and determine the coefficients by variational principle $\frac{\partial F_\theta[g]}{\partial g_j} = 0$ at different values of θ^2 . The value $\theta^2 = 0$ corresponds to addressing exclusively the polynomial approximation, as prescribed by the choice of $A[g]$, whereas the choices $\theta^2 \rightarrow \infty$ would correspond to dealing purely with the variance minimisation and would result in $g_j = 0$. Note that we can map $\theta^2 = \frac{\lambda}{1-\lambda}$ for simplicity, such that $\lambda \in [0, 1)$ and $\theta^2 \rightarrow \infty$ for $\lambda \rightarrow 1$. Furthermore, any relative normalisation term between the two functionals can be reabsorbed into θ^2 . Depending on the choice of the basis, the coefficients g_j may grow over different orders of

magnitude and numerical instabilities may appear. This can be addressed in practice by using arbitrary precision arithmetic.

We now discuss in detail how to generalise the modified Backus-Gilbert [22] for a generic basis of functions, starting from the construction of $A[g]$. Following the original paper we can generalise the L_2 -norm of the difference between the target function and the polynomial reconstruction using an arbitrary family of basis function $P_k(x) = \sum_{j=0}^k p_j^{(k)} x^j$ defined in an interval $x \in [p_-, p_+]$. As for the Chebyshev, we will deal in general with a shifted version of this family of polynomials in e^{-x} defined in a generic interval $[a, b]$

$$\tilde{P}_k(x) = \sum_{j=0}^k \tilde{p}_j^{(k)} e^{-jx}, \quad x \in [a, b], \quad (\text{B.5})$$

where $\tilde{P}_k(x) = P_k(h(x))$ and $h(x) = Ae^{-x} + B$ is an invertible map that satisfies $h(a) = p_-$ and $h(b) = p_+$. The interval $[a, b]$ has to match the range of integration of the observable O in (B.1). The functional $A[g]$ now reads

$$A[g] = \int_a^b d\omega \Omega(\omega) \left[K(\omega) - \sum_{j=0}^N g_j \tilde{P}_j(\omega) \right]^2. \quad (\text{B.6})$$

With respect to the original version we now have introduced a generic weight $\Omega(\omega)$; note that we start the approximation at $\tilde{P}_0(\omega)$ (as long as $\Omega(\omega)$ can be integrated in $[a, b]$).

If we consider only the $A[g]$ term, the solution of the system by variational principle is given by

$$\mathbf{A} \cdot \mathbf{g} = \mathbf{K} \quad \longleftrightarrow \quad \mathbf{g} = \mathbf{A}^{-1} \cdot \mathbf{K} \quad (\text{B.7})$$

where

$$A_{ij} = \int_a^b d\omega \Omega(\omega) \tilde{P}_i(\omega) \tilde{P}_j(\omega), \quad (\text{B.8})$$

$$K_i = \int_a^b d\omega \Omega(\omega) \tilde{P}_i(\omega) K(\omega), \quad (\text{B.9})$$

and \mathbf{g} is a vector of parameters.

With this setup, the convenient choice consists in picking a set of (shifted) orthogonal polynomials

$$\langle \tilde{P}_i, \tilde{P}_j \rangle = \int_a^b dx \Omega(x) \tilde{P}_i(x) \tilde{P}_j(x) \propto \delta_{ij}, \quad (\text{B.10})$$

with Ω being the actual weight that defines the scalar product. The advantage is immediately clear, as the matrix \mathbf{A} becomes

$$A_{ij} = \langle \tilde{P}_i, \tilde{P}_j \rangle, \quad (\text{B.11})$$

and the coefficients are given by

$$g_i = \frac{1}{\langle \tilde{P}_i, \tilde{P}_i \rangle} \int_a^b d\omega \Omega(\omega) \tilde{P}_i(\omega) K(\omega). \quad (\text{B.12})$$

Since the matrix \mathbf{A} is now diagonal, the inverse required to compute Eq. (B.7) is analytically known. Furthermore, the solution is now equivalent to the projection on the polynomial basis.

We can now include the B term, i.e. the covariance matrix of the data. Note that in general we now need to consider a linear combination of the correlator at different time slices according to the polynomial basis, i.e.

$$C^P(k) = \int_a^b d\omega \rho(\omega) \tilde{P}_k(\omega) = \int_a^b d\omega \rho(\omega) \sum_{j=0}^k \tilde{p}_j^{(k)} e^{-j\omega} = \sum_{j=0}^k \tilde{p}_j^{(k)} C(j), \quad (\text{B.13})$$

such that

$$B[g] = \sum_{i,j} g_i \sigma_{ij}^P g_j, \quad \sigma_{ij}^P = \text{Cov}[C^P(i), C^P(j)]. \quad (\text{B.14})$$

The full functional is then

$$F_\theta[g] = A[g] + \theta^2 B[g] \quad (\text{B.15})$$

and the final solution is

$$\mathbf{g}_\theta = \mathbf{F}_\theta^{-1} \cdot \mathbf{K} \quad (\text{B.16})$$

with

$$\mathbf{F}_\theta = \mathbf{A} + \theta^2 \mathbf{B}, \quad (\text{B.17})$$

where $B_{ij} = \sigma_{ij}^P$. If \mathbf{A} is diagonal (and possibly proportional to the identity), the inversion of the matrix \mathbf{F}_θ may be better conditioned and possible numerical instabilities arising from an ill-conditioned matrix \mathbf{A} may be avoided.

On top of that we could also implement some constraints that our approximation has to fulfil. In particular, following what was done for the spectral function in [22, 38], we can require that the polynomial approximation preserves the (weighted) area of the target function, i.e.

$$\int_a^b d\omega \Omega(\omega) \sum_{k=0}^N g_k \tilde{P}_k(\omega) = \int_a^b d\omega \Omega(\omega) K(\omega). \quad (\text{B.18})$$

This can be expressed as

$$\mathbf{R}^T \cdot \mathbf{g}_\theta = r, \quad (\text{B.19})$$

where

$$R_k = \int_a^b d\omega \Omega(\omega) \tilde{P}_k(\omega), \quad r = \int_a^b d\omega \Omega(\omega) K(\omega). \quad (\text{B.20})$$

Taking into account these constraints, the solution becomes

$$\mathbf{g}_\theta = \mathbf{F}_\theta^{-1} \cdot \mathbf{K} + \mathbf{F}_\theta^{-1} \cdot \mathbf{R} \frac{r - \mathbf{R}^T \cdot \mathbf{F}_\theta^{-1} \cdot \mathbf{K}}{\mathbf{R}^T \cdot \mathbf{F}_\theta^{-1} \cdot \mathbf{R}}. \quad (\text{B.21})$$

The final observable then reads

$$O_\theta \simeq \sum_{j=0}^N g_{\theta,j} C^P(j), \quad (\text{B.22})$$

for a given value of θ . The choice of θ is in principle arbitrary. A common choice is to take the value θ^* that gives equal weight to the A and B functional, $A[g_{\theta^*}] = B[g_{\theta^*}]$, i.e. an equal weight to statistical and systematic error. For a given choice of θ , it is important to make sure that the value of the final observable is stable for small changes in θ , in order to make sure that the procedure did not introduce any bias.

To conclude, note that this recovers the method first proposed in [22] if we consider the following substitutions

$$\begin{aligned} \tilde{P}_j(\omega) &\rightarrow e^{-(j+1)\omega}, \\ \Omega(\omega) &\rightarrow 1, \\ \theta^2 &\rightarrow \lambda/(1-\lambda), \\ F[g] &\rightarrow (1-\lambda)F[g]. \end{aligned}$$

B.2 A different perspective

The previous reformulation in Sec. B.1 allows us to rely on arbitrary polynomials for the approximation. In this general picture it is useful to consider a different perspective to the method: we can reduce the problem to finding a suitable correction to the optimal coefficients, i.e.

$$g_j = \gamma_j + \epsilon_j,$$

where γ_j are the coefficients of the polynomial approximation coming purely from the functional $A[\gamma]$, i.e. $\boldsymbol{\gamma} = \mathbf{A}^{-1}\mathbf{K}$ as in Eq. (B.7), and ϵ_j a correction that takes into account the data. We can then rewrite the functional as

$$F_\theta[g] = F_\theta[\boldsymbol{\gamma} + \boldsymbol{\epsilon}] = F_\theta[\boldsymbol{\gamma}] + \delta F_\theta[\boldsymbol{\epsilon}], \quad (\text{B.23})$$

and explicitly

$$\delta F_\theta[\boldsymbol{\epsilon}] = \int_a^b d\omega \Omega(\omega) \left[\sum_{k=0}^N \epsilon_k \tilde{P}_k(\omega) \right]^2 + \theta^2 (2\gamma_i \sigma_{ij}^P \epsilon_j + \epsilon_i \sigma_{ij}^P \epsilon_j). \quad (\text{B.24})$$

The minimisation of $\delta F_\theta[\boldsymbol{\epsilon}]$ gives

$$\boldsymbol{\epsilon}_\theta = -\theta^2 (\mathbf{A} + \theta^2 \boldsymbol{\sigma}^P)^{-1} \boldsymbol{\sigma}^P \boldsymbol{\gamma}, \quad (\text{B.25})$$

which is equivalent to the previous approach. It is then clear that ϵ_j are by construction coefficients that should not modify the quality of the polynomial approximations but take care of the reduction of the statistical noise. In practice, this will of course depend on the choice of θ .

C Fit strategy

We discuss the general strategy for the Bayesian fit used in the analysis. We consider only linear fits, as these are the ones directly relevant for this work. To keep the discussion very general we consider a linear model in the form

$$y(\mathbf{p}, x) = \sum_{\alpha=1}^M p_{\alpha} X_{\alpha}(x), \quad \mathbf{p} = (p_1, p_2, \dots, p_M), \quad (\text{C.1})$$

where $X_{\alpha}(x)$ are known coefficients (which in principle can depend on x) and p_{α} are M parameters we want to determine.

C.1 MAP with bounds

We address the fits using Bayesian statistics, in particular using a *maximum a posteriori* (MAP) probability estimate, which relies on an *augmented* χ^2 with Gaussian priors. On top of that, we implement generic bounds on the parameters. The way we address this is by “wrapping” the parameters in a function $p_{\alpha} = f(\pi_{\alpha})$ which encodes the desired bounds. In this case, the fit is performed on the new parameters π_{α} , and the prior is introduced accordingly. The augmented χ^2 reads

$$\chi_{\text{aug}}^2 = \sum_{i,j=1}^N \left(y_i - \sum_{\alpha=1}^M f(\pi_{\alpha}) X_{\alpha}(x_i) \right) \text{Cov}_{ij}^{-1} \left(y_j - \sum_{\alpha=1}^M f(\pi_{\alpha}) X_{\alpha}(x_j) \right) + \sum_{\alpha=1}^M \frac{(\pi_{\alpha} - \bar{\pi}_{\alpha})^2}{\bar{\sigma}_{\alpha}^2}. \quad (\text{C.2})$$

Note that the prior distributions refer to the internal parameters π_{α} and are assumed to be Gaussians. This allows to deal with a more generic distribution for the parameters p_{α} , depending on the shape of the wrapping function f . The parameters are found as usual by imposing $\frac{\partial \chi_{\text{aug}}^2}{\partial \pi_{\alpha}} = 0$; note that in this case the problem is no more linear due to the presence of f .

C.2 MAP with bootstrap

As outlined in the sections above, the presence of a “wrapping” function on the parameters p_{α} implies that their distribution is in general non Gaussian. This is obvious from the fact that we assume the internal parameters π_{α} to be Gaussian and that the wrapping function implements some bounds, therefore limiting the domain of p_{α} . Instead of fitting the central value of the data and estimating their error from the inverse of the curvature matrix (the Hessian of the χ^2 with respect to the parameters), it is then more convenient to adopt a bootstrap approach, such that the procedure automatically takes into account any deviation from Gaussianity. In practice, one would then fit all the bootstrap bins and reconstruct the distribution of the parameters, treating the error accordingly.

The approach we adopt consists in assuming a normal distribution for the internal parameters $\pi \sim \mathcal{N}(\mu, \sigma)$ such that $p_{\alpha} = f(\pi_{\alpha})$ is distributed according to our prior knowledge of the parameters. In practice, considering a set of N_b bootstrap bins with corresponding data y_i^b , we perform N_b fits to the data where each time we use a different prior value

$\bar{\pi}_\alpha^b$ sampled from the normal distribution $\mathcal{N}(\mu, \sigma)$. This ensures that the correct prior is assumed for p_α . For example, in the case where the data contain little information and $\min(\chi_{\text{aug}}^2) \simeq \min(\chi_{\text{prior}}^2)$, the fit gives back the prior information we encoded by hand.

References

- [1] LHCb collaboration, *Measurement of Form-Factor-Independent Observables in the Decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$* , *Phys. Rev. Lett.* **111** (2013) 191801 [[1308.1707](#)].
- [2] LHCb collaboration, *Test of lepton universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays*, *Phys. Rev. Lett.* **113** (2014) 151601 [[1406.6482](#)].
- [3] BABAR collaboration, *Evidence for an excess of $\bar{B} \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ decays*, *Phys. Rev. Lett.* **109** (2012) 101802 [[1205.5442](#)].
- [4] BABAR collaboration, *Measurement of an Excess of $\bar{B} \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, *Phys. Rev. D* **88** (2013) 072012 [[1303.0571](#)].
- [5] BELLE collaboration, *Measurement of the branching ratio of $\bar{B} \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ relative to $\bar{B} \rightarrow D^{(*)} \ell^- \bar{\nu}_\ell$ decays with hadronic tagging at Belle*, *Phys. Rev. D* **92** (2015) 072014 [[1507.03233](#)].
- [6] LHCb collaboration, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau) / \mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)$* , *Phys. Rev. Lett.* **115** (2015) 111803 [[1506.08614](#)].
- [7] HEAVY FLAVOR AVERAGING GROUP, HFLAV collaboration, *Averages of b -hadron, c -hadron, and τ -lepton properties as of 2021*, *Phys. Rev. D* **107** (2023) 052008 [[2206.07501](#)].
- [8] P. Gambino, K.J. Healey and S. Turczyk, *Taming the higher power corrections in semileptonic B decays*, *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics* **763** (2016) 60 [[1606.06174](#)].
- [9] BABAR collaboration, *Measurement of $|V(cb)|$ and the Form-Factor Slope in $\bar{B} \rightarrow D \ell^- \bar{\nu}_\ell$ Decays in Events Tagged by a Fully Reconstructed B Meson*, *Phys. Rev. Lett.* **104** (2010) 011802 [[0904.4063](#)].
- [10] BELLE collaboration, *Measurement of the decay $B \rightarrow D \ell \nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$* , *Phys. Rev.* **D93** (2016) 032006 [[1510.03657](#)].
- [11] MILC collaboration, *$B \rightarrow D \ell \nu$ form factors at nonzero recoil and $|V_{cb}|$ from 2+1-flavor lattice QCD*, *Phys. Rev. D* **92** (2015) 034506 [[1503.07237](#)].
- [12] HPQCD collaboration, *$B \rightarrow D \ell \nu$ form factors at nonzero recoil and extraction of $|V_{cb}|$* , *Phys. Rev. D* **92** (2015) 054510 [[1505.03925](#)].
- [13] FLAVOUR LATTICE AVERAGING GROUP (FLAG) collaboration, *FLAG Review 2021*, *Eur. Phys. J. C* **82** (2022) 869 [[2111.09849](#)].
- [14] T. Kaneko, *Heavy flavor physics from lattice QCD*, *PoS LATTICE2022* (2023) 238 [[2304.01618](#)].
- [15] S. Hashimoto, *Inclusive semi-leptonic B meson decay structure functions from lattice QCD*, *Progress of Theoretical and Experimental Physics* **2017** (2017) 53 [[1703.01881](#)].
- [16] 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 [[1704.08993](#)].

- [17] P. Gambino and S. Hashimoto, *Inclusive Semileptonic Decays from Lattice QCD*, *PHYSICAL REVIEW LETTERS* **125** (2020) 32001 [[2005.13730](#)].
- [18] A. Barone, A. Jüttner, S. Hashimoto, T. Kaneko and R. Kellermann, *Inclusive semi-leptonic $B_{(s)}$ mesons decay at the physical b quark mass*, *PoS LATTICE2022* (2023) 403 [[2211.15623](#)].
- [19] R. Kellermann, A. Barone, S. Hashimoto, A. Jüttner and T. Kaneko, *Inclusive semi-leptonic decays of charmed mesons with Möbius domain wall fermions*, *PoS LATTICE2022* (2023) 414 [[2211.16830](#)].
- [20] J.C.A. Barata and K. Fredenhagen, *Particle scattering in Euclidean lattice field theories*, *Commun. Math. Phys.* **138** (1991) 507.
- [21] G. Bailas, S. Hashimoto and T. Ishikawa, *Reconstruction of smeared spectral functions from Euclidean correlation functions*, *Progress of Theoretical and Experimental Physics* **2020** (2020) 43 [[2001.11779](#)].
- [22] M. Hansen, A. Lupo and N. Tantalo, *Extraction of spectral densities from lattice correlators*, *Physical Review D* **99** (2019) [[1903.06476](#)].
- [23] P. Gambino, S. Hashimoto, S. Mächler, M. Panero, F. Sanfilippo, S. Simula et al., *Lattice QCD study of inclusive semileptonic decays of heavy mesons*, *JHEP* **07** (2022) 083 [[2203.11762](#)].
- [24] A.X. El-Khadra, A.S. Kronfeld and P.B. Mackenzie, *Massive fermions in lattice gauge theory*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **55** (1997) 3933 [[hep-lat/9604004](#)].
- [25] N.H. Christ, M. Li and H.-W. Lin, *Relativistic Heavy Quark Effective Action*, *Phys. Rev. D* **76** (2007) 074505 [[hep-lat/0608006](#)].
- [26] H.W. Lin and N. Christ, *Nonperturbatively determined relativistic heavy quark action*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **76** (2007) [[hep-lat/0608005](#)].
- [27] Y. Shamir, *Chiral Fermions from Lattice Boundaries*, *Nuclear Physics, Section B* **406** (1993) 90 [[hep-lat/9303005v1](#)].
- [28] V. Furman and Y. Shamir, *Axial symmetries in lattice QCD with Kaplan fermions*, *Nuclear Physics, Section B* **439** (1994) 54 [[hep-lat/9405004v2](#)].
- [29] R.C. Brower, H. Neff and K. Orginos, *The Möbius domain wall fermion algorithm*, *Comput. Phys. Commun.* **220** (2017) 1 [[1206.5214](#)].
- [30] Y.-G. Cho, S. Hashimoto, A. Jüttner, T. Kaneko, M. Marinkovic, J.-I. Noaki et al., *Improved lattice fermion action for heavy quarks*, *JHEP* **05** (2015) 072 [[1504.01630](#)].
- [31] A.V. Manohar and M.B. Wise, *Heavy quark physics*, vol. 10, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol (2000).
- [32] 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 [[hep-ph/9307247](#)].
- [33] E. McLean, C.T.H. Davies, J. Koponen and A.T. Lytle, *$B_s \rightarrow D_s \ell \nu$ Form Factors for the full q^2 range from Lattice QCD with non-perturbatively normalized currents*, *Phys. Rev. D* **101** (2020) 074513 [[1906.00701](#)].

- [34] B. Blossier, P.-H. Cahue, J. Heitger, S. La Cesa, J. Neuendorf and S. Zafeiropoulos, *Extraction of $B_s \rightarrow D_s^{(*)}$ form factors from $N_f=2$ lattice QCD*, *Phys. Rev. D* **105** (2022) 054515 [[2110.10061](#)].
- [35] J. Flynn, R. Hill, A. Juettner, A. Soni, J.T. Tsang and O. Witzel, *Form factors for semileptonic $B \rightarrow \pi$, $B_s \rightarrow K$ and $B_s \rightarrow D_s$ decays*, *PoS LATTICE2021* (2022) 306 [[2112.10580](#)].
- [36] P. Gambino, S. Hashimoto, S. Mächler, M. Panero, F. Sanfilippo, S. Simula et al., *Lattice QCD study of inclusive semileptonic decays of heavy mesons*, *JHEP* **07** (2022) 083 [[2203.11762](#)].
- [37] G. Backus and F. Gilbert, *The Resolving Power of Gross Earth Data*, *Geophysical Journal of the Royal Astronomical Society* **16** (1968) 169.
- [38] 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 [[2111.12774](#)].
- [39] EXTENDED TWISTED MASS COLLABORATION (ETMC) collaboration, *Probing the Energy-Smeared R Ratio Using Lattice QCD*, *Phys. Rev. Lett.* **130** (2023) 241901 [[2212.08467](#)].
- [40] Y. Shamir, *Chiral fermions from lattice boundaries*, *Nucl. Phys. B* **406** (1993) 90 [[hep-lat/9303005](#)].
- [41] V. Furman and Y. Shamir, *Axial symmetries in lattice QCD with Kaplan fermions*, *Nucl. Phys. B* **439** (1995) 54 [[hep-lat/9405004](#)].
- [42] Y. Iwasaki, *Renormalization Group Analysis of Lattice Theories and Improved Lattice Action. II. Four-dimensional non-Abelian $SU(N)$ gauge model*, [1111.7054](#).
- [43] C. Allton, D.J. Antonio, Y. Aoki, T. Blum, P.A. Boyle, N.H. Christ et al., *Physical results from 2+1 flavor domain wall QCD and $SU(2)$ chiral perturbation theory*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **78** (2008) [[0804.0473](#)].
- [44] P. Boyle, A. Yamaguchi, G. Cossu and A. Portelli, “Grid: Data parallel C++ mathematical object library.” <https://github.com/paboyle/Grid>.
- [45] P.A. Boyle, G. Cossu, A. Yamaguchi and A. Portelli, *Grid: A next generation data parallel C++ QCD library*, *PoS LATTICE2015* (2016) 023 [[1512.03487](#)].
- [46] A. Yamaguchi, P. Boyle, G. Cossu, G. Filaci, C. Lehner and A. Portelli, *Grid: OneCode and FourAPIs*, *PoS LATTICE2021* (2022) 035 [[2203.06777](#)].
- [47] A. Portelli, R. Abott, N. Asmussen, A. Barone, P.A. Boyle, F. Erben et al., *aportelli/hadrons: Hadrons v1.3*, Mar., 2022. [10.5281/zenodo.6382460](#).
- [48] P. Lepage and C. Gohlke, *gplepage/lqfit: lqfit version 12.0.3*, Dec., 2021. [10.5281/zenodo.5777652](#).
- [49] G.P. Lepage, B. Clark, C.T.H. Davies, K. Hornbostel, P.B. Mackenzie, C. Morningstar et al., *Constrained curve fitting*, *Nucl. Phys. B Proc. Suppl.* **106** (2002) 12 [[hep-lat/0110175](#)].
- [50] J.M. Flynn, R.C. Hill, A. Jüttner, A. Soni, J.T. Tsang and O. Witzel, *Semi-leptonic form factors for $B_s \rightarrow K\ell\nu$ and $B_s \rightarrow D_s\ell\nu$* , *PoS LATTICE2018* (2019) 290 [[1903.02100](#)].
- [51] J. Flynn, R. Hill, A. Jüttner, A. Soni, J.T. Tsang and O. Witzel, *Semileptonic $B \rightarrow \pi\ell\nu$* ,

- $B \rightarrow D\ell\nu$, $B_s \rightarrow K\ell\nu$, and $B_s \rightarrow D_s\ell\nu$ decays, *PoS LATTICE2019* (2019) 184 [[1912.09946](#)].
- [52] J.M. Flynn, R.C. Hill, A. Jüttner, A. Soni, J.T. Tsang and O. Witzel, *Exclusive semileptonic $B_s \rightarrow K\ell\nu$ decays on the lattice*, [2303.11280](#).
- [53] UKQCD collaboration, *Quark mass dependence of hadron masses from lattice QCD*, *Phys. Rev. D* **59** (1999) 074503 [[hep-lat/9810021](#)].
- [54] UKQCD collaboration, *Decay width of light quark hybrid meson from the lattice*, *Phys. Rev. D* **73** (2006) 074506 [[hep-lat/0603007](#)].
- [55] P.A. Boyle, A. Jüttner, C. Kelly and R.D. Kenway, *Use of stochastic sources for the lattice determination of light quark physics*, *JHEP* **08** (2008) 086 [[0804.1501](#)].
- [56] 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](#)].
- [57] C.T. Sachrajda and G. Villadoro, *Twisted boundary conditions in lattice simulations*, *Phys. Lett. B* **609** (2005) 73 [[hep-lat/0411033](#)].
- [58] M.G. Alford, T. Klassen and P. Lepage, *The $D234$ action for light quarks*, *Nucl. Phys. B Proc. Suppl.* **47** (1996) 370 [[hep-lat/9509087](#)].
- [59] A.C. Lichtl, *Quantum Operator Design for Lattice Baryon Spectroscopy*, other thesis, 9, 2006, [10.2172/917690](#), [[hep-lat/0609019](#)].
- [60] J.M. Flynn, T. Izubuchi, T. Kawanai, C. Lehner, A. Soni, R.S. Van de Water et al., *$B \rightarrow \pi\ell\nu$ and $B_s \rightarrow K\ell\nu$ form factors and $|V_{ub}|$ from 2+1-flavor lattice QCD with domain-wall light quarks and relativistic heavy quarks*, *Phys. Rev. D* **91** (2015) 074510 [[1501.05373](#)].
- [61] RBC AND UKQCD COLLABORATIONS collaboration, *Nonperturbative tuning of an improved relativistic heavy-quark action with application to bottom spectroscopy*, *Phys. Rev. D* **86** (2012) 116003.
- [62] A.X. El-Khadra, A.S. Kronfeld, P.B. Mackenzie, S.M. Ryan and J.N. Simone, *The Semileptonic decays $B \rightarrow \pi \ell \nu$ and $D \rightarrow \pi \ell \nu$ from lattice QCD*, *Phys. Rev. D* **64** (2001) 014502 [[hep-ph/0101023](#)].
- [63] A. Weiße, G. Wellein, A. Alvermann and H. Fehske, *The kernel polynomial method*, *Rev. Mod. Phys.* **78** (2006) 275.