The $B_s \to \mu^+\mu^-\gamma$ decay rate at large q^2 from lattice QCD

R. Frezzotti and N. Tantalo

Dipartimento di Fisica and INFN, Università di Roma "Tor Vergata",
Via della Ricerca Scientifica 1, I-00133 Roma, Italy

G. Gagliardi, F. Sanfilippo, and S. Simula

Istituto Nazionale di Fisica Nucleare, Sezione di Roma Tre,
Via della Vasca Navale 84, I-00146 Rome, Italy

V. Lubicz

Dipartimento di Fisica, Università Roma Tre and INFN, Sezione di Roma Tre,
Via della Vasca Navale 84, I-00146 Rome, Italy

G. Martinelli

Physics Department and INFN Sezione di Roma La Sapienza,
Piazzale Aldo Moro 5, 00185 Roma, Italy

C.T. Sachrajda

Department of Physics and Astronomy, University of Southampton,
Southampton SO17 1BJ, UK
(Dated: February 6, 2024)

We determine, by means of lattice QCD calculations, the local form factors describing the $B_s \to \mu^+\mu^-\gamma$ decay. For this analysis we make use of the gauge configurations produced by the ETM Collaboration with $N_f = 2 + 1 + 1$ flavour of Wilson-Clover twisted-mass fermions at maximal twist. To obtain the B_s meson form-factors, we perform simulations for several heavy-strange meson masses m_{H_s} in the range $m_{H_s} \in [m_{D_s}, 2m_{D_s}]$, and extrapolate to the physical B_s meson point $m_{B_s} \simeq 5.367$ GeV making use of the HQET scaling laws. We cover the region of large di-muon invariant masses $\sqrt{q^2} > 4.16$ GeV, and use our results to determine the branching fraction for $B_s \to \mu^+\mu^-\gamma$, which has been recently measured by LHCb in the region $\sqrt{q^2} > 4.9$ GeV. The largest contribution to the uncertainty in the partial branching fractions at values of $\sqrt{q^2} < 4.8$ GeV is now due to resonance and other long-distance effects, including those from "charming penguins", which we estimate by summing over the contributions from the $J_P = 1^-$ charmonium resonances.

I. INTRODUCTION

The flavour-changing-neutral current (FCNC) transition $B_s \to \mu^+\mu^-\gamma$, being strongly suppressed in the Standard Model (SM), represents an ideal channel to look for signals of New Physics (NP). Although there is an additional factor of $\alpha_{\rm em}$ in the amplitude for this process compared to that for the widely-studied $B_s \to \mu^+\mu^-$ decay, the presence of the final state energetic photon removes the helicity suppression making the rates for the two processes approximately comparable. The LHCb Collaboration has recently searched for signals of this process [1, 2] but found no significant events resulting in an upper limit for the branching ratio of $\mathcal{B}(B_s \to \mu^+\mu^-\gamma) < 2.0 \times 10^{-9}$ for photons γ emitted by the quarks[1] (the so-called initial-state radiation contribution, or ISR) and for di-muon inviariant masses $\sqrt{q^2} > 4.9$ GeV. Future measurements will be able to reduce the experimental uncertainties and cover a larger portion of the phase space reaching lower values of q^2. On the other hand, a first-principles theoretical prediction of the $B_s \to \mu^+\mu^-\gamma$ decay rate is currently missing. While the leading hadronic effects in the $B_s \to \ell^+\ell^-$ ($\ell = e, \mu, \tau$) decay amplitude depend only on the B_s-meson decay constant f_{B_s}, which is known to sub-percent precision from lattice computations, the determination of the amplitude for the $B_s \to \mu^+\mu^-\gamma$ decay is much more complex. In this case the non-perturbative hadronic effects depend not only on local form factors, but also on resonance contributions. Existing estimates of the rate are based on light-cone sum rules (LCSR) [3], on model/effective-theory calculations, such as the relativistic dispersion approach based on the constituent-quark picture [4] and, more recently, on the use of existing lattice QCD results for the radiative leptonic form factors of the D_s meson to estimate some of the $B_s \to \mu^+\mu^-\gamma$ transition form factors assuming vector-meson dominance (VMD) [5].

---

[1] The final-state radiation (FSR) contribution, in which the photon is emitted from a final-state muon, dominates at small photon energies and has been subtracted in Ref. [1]. The interference between FSR and ISR is instead found to be negligible.

The aim of this paper is to provide a first-principles determination, using lattice QCD, of the local form factors F_V, F_A, F_{TV}, F_{TA} and $\bar{F}_T$, which represent the only non-perturbative QCD input in the determination of the $\bar{B}_s \to \mu^+\mu^-\gamma$ transition matrix elements [2] $\langle\gamma(\varepsilon)\,\mu^+\mu^-|\mathcal{O}_{7,9,10}|\bar{B}_s\rangle$, where ε is the photon's polarization vector, and the $\mathcal{O}_i$ are the standard operators appearing in the effective weak Hamiltonian $H_{\text{eff}}^{b\to s}$ describing the FCNC $b \to s$ transition and are defined in Eq. (3) below. We explore the region of large di-muon invariant masses $\sqrt{q^2} > 4.16$ GeV. In this region, the impact of the contributions from the operators $\mathcal{O}_{1-6,8}$ (which are neglected at present) stemming from the four-quark operators and from the chromomagnetic penguin operator in $H_{\text{eff}}^{b\to s}$ is expected to be modest [6], and the rate can be reliably computed from the knowledge of the local form factors only. As an estimate of the systematic error induced by this approximation, we employ a phenomenological description of the charming-penguin contribution, illustrated in Figure 4 below, which is expected to be among the largest of the contributions we have neglected because of the presence of broad charmonium resonances which are near or within the region of q^2 we consider. Whilst we find that the differential branching fractions themselves are dominated by the form factors (in particular by F_V), the dominant uncertainty for $\sqrt{q^2} < 4.8$ GeV is that due to charming penguin contributions (see Figure 22) and therefore in order to improve the precision and to be able to reach lower values of q^2 the development of a rigorous treatment of the contributions from $\mathcal{O}_{1-6,8}$ will be necessary.

For this calculation we employ the same set of gauge configurations which have recently been used in our work on the radiative leptonic form factors of the D_s meson [7]. The configurations have been generated by the Extended Twisted Mass Collaboration (ETMC) with $N_f = 2 + 1 + 1$ flavors of Wilson-clover twisted-mass fermions at maximal twist, and sea-quark masses tuned very close to their physical values for all quark flavours. The ensembles correspond to four values of the lattice spacing a in the range $[0.056, 0.09]$ fm.

Our strategy for obtaining results for the physical $\bar{B}_s$ meson, is to perform simulations at a series of unphysical (lighter) heavy-strange pseudoscalar mesons $\bar{H}_s$, consisting of a heavy quark (h) and a strange anti-quark ($\bar{s}$), with $m_{H_s} \in [m_{D_s}, 2m_{D_s}]$. We then use heavy quark effective theory (HQET) relations to guide the extrapolation of the results to the physical $\bar{B}_s$ meson. For each heavy-quark mass we evaluate the form factors at four different values of the energy of the photon E_γ (as measured in the rest frame of the decaying meson), which we keep fixed in units of the heavy-strange meson mass m_{H_s}. The values are given by $x_\gamma \equiv 2E_\gamma/m_{H_s} = 0.1, 0.2, 0.3$ and 0.4, and for $m_{H_s} = m_{B_s}$ this corresponds to $q^2 > (4.16\,\text{GeV})^2$. Our main result is the calculation of $\mathcal{B}_{\text{SD}}(x_\gamma^{\text{cut}})$, which is the ISR contribution (to which we refer to in the paper as the structure-dependent contribution) to the branching fraction for $q^2 > m_{B_s}^2(1 - x_\gamma^{\text{cut}})$, and is given in Table VI. Our results at $x_\gamma^{\text{cut}} = 0.166$ (i.e. $q^2 > (4.9\,\text{GeV})^2$) is

$$\mathcal{B}_{\text{SD}}(x_\gamma^{\text{cut}} = 0.166) = 6.9(9) \times 10^{-11}\ , \tag{1}$$

which is well within the current upper-bound set by the LHCb, $\mathcal{B}(x_\gamma^{\text{cut}} = 0.166) < 2.0 \times 10^{-9}$. Anticipating that future experiments will be able to access values of q^2 below $(4.9\,\text{GeV})^2$, we present in Table VI the partial branching fractions corresponding to values of the lower cut-off $\sqrt{q_{\text{cut}}^2} = m_{B_s}\sqrt{1 - x_\gamma^{\text{cut}}}$ from 4.1 GeV to 5.2 GeV in steps of 0.1 GeV. We find that the partial branching fractions in the q^2-region we explored is dominated by the contribution of the vector form factor F_V; the combined contribution of all other local form factors $F_A, F_{TV}, F_{TA}, \bar{F}_T$ is of the order of $\mathcal{O}(10\%)$.

The plan for the remainder of this paper is as follows. In Section II we briefly recall the definition of the local form factors in terms of matrix elements of the operators in the effective weak Hamiltonian. In Section III we explain our strategy for the determination of the local form factors F_V, F_A, F_{TV} and F_{TA}, and present the results of the continuum extrapolation for each value of the simulated heavy-strange meson mass. We also discuss the heavy-quark scaling relations, which are then used to extrapolate the results to the mass of the physical B_s meson. In Section IV we present our strategy for evaluating the local form factor $\bar{F}_T$, whose lattice determination is complicated by the problem of the analytic continuation to Euclidean space-time of the relevant Minkowski correlation functions. We tackle this problem using the spectral reconstruction technique developed in Ref. [8]. In Section V we provide our determination of the differential cross-section for $B_s \to \mu^+\mu^-\gamma$ as well as the total differential rate for different q^2 intervals. We then compare our results with existing estimates as well as with the LHCb measurement [1, 2] corresponding to the interval $q^2 > (4.9\,\text{GeV})^2$. Finally, in Section VI we present our conclusions and outlook for future improvements.

---

[2] The text and diagrams here and below correspond to the decay of the $\bar{B}_s$ meson, which contains a valence b-quark.

II. THE EFFECTIVE WEAK HAMILTONIAN AND LOCAL FORM FACTORS

The low-energy effective weak Hamiltonian describing the $b \to s$ transition, neglecting doubly Cabibbo-suppressed contributions, is given by [9]

$$\mathcal{H}_{\text{eff}}^{b \to s} = 2\sqrt{2} G_F V_{tb} V_{ts}^* \left[\sum_{i=1,2} C_i(\mu) \mathcal{O}_i^c + \sum_{i=3}^{6} C_i(\mu) \mathcal{O}_i + \frac{\alpha_{\text{em}}}{4\pi} \sum_{i=7}^{10} C_i(\mu) \mathcal{O}_i \right] , \tag{2}$$

where G_F is the Fermi constant, C_i are the Wilson coefficients and $\mathcal{O}_i$ are local operators renormalized at the scale μ. The latter are given by $(P_{L(R)} = (1 \mp \gamma^5)/2)$

$$\mathcal{O}_1^c = (\bar{s}_i \gamma^\mu P_L c_j)(\bar{c}_j \gamma^\mu P_L b_i) , \qquad \mathcal{O}_2^c = (\bar{s}\gamma^\mu P_L c)(\bar{c}\gamma^\mu P_L b) , \tag{3}$$

$$\mathcal{O}_7 = -\frac{m_b}{e} \bar{s}\sigma^{\mu\nu} F_{\mu\nu} P_R b , \qquad \mathcal{O}_8 = -\frac{g_s m_b}{4\pi\alpha_{\text{em}}} \bar{s}\sigma^{\mu\nu} G_{\mu\nu} P_R b , \tag{4}$$

$$\mathcal{O}_9 = (\bar{s}\gamma^\mu P_L b)(\bar{\mu}\gamma_\mu \mu) , \qquad \mathcal{O}_{10} = (\bar{s}\gamma^\mu P_L b)(\bar{\mu}\gamma_\mu \gamma^5 \mu) \tag{5}$$

while the operators $\mathcal{O}_{3-6}$ are the QCD penguins. In the previous equations i, j are color indices, while $F_{\mu\nu}$ and $G_{\mu\nu}$ are the electromagnetic and gluonic field strength tensor, respectively. In the following, for the CKM matrix elements we use the PDG values $|V_{tb}| = 1.014(29)$ and $|V_{ts}| = 4.15(9) \times 10^{-2}$ [10]. Our conventions for the gamma matrices are

$$\gamma^5 = i\gamma^0 \gamma^1 \gamma^2 \gamma^3 , \qquad \sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu] , \tag{6}$$

while for the Levi-Civita tensor we adopt the convention $\varepsilon^{0123} = -1$. The transition amplitude for the decay of the $\bar{B}_s$ meson is given by

$$\mathcal{A}[\bar{B}_s \to \mu^+ \mu^- \gamma] = \langle \gamma(\boldsymbol{k}, \varepsilon) \mu^+(p_1) \mu^-(p_2)| - \mathcal{H}_{\text{eff}}^{b \to s} |\bar{B}_s(\boldsymbol{p})\rangle_{\text{QCD+QED}} , \tag{7}$$

where $k = (E_\gamma = |\boldsymbol{k}|, \boldsymbol{k})$ and $p = (E_{B_s}, \boldsymbol{p})$ are the momenta of the photon and $\bar{B}_s$ meson respectively, ε is the photon's polarization vector, and p_1 and p_2 the momenta of the μ^+ and μ^- respectively. The di-muon four-momentum is then $q = p_1 + p_2 = p - k$. The amplitude $\mathcal{A}$ is then expanded to leading non-vanishing order $(\mathcal{O}(\alpha_{\text{em}}^{3/2}))$ in the electromagnetic coupling α_{em}, and can be expressed as [9]

$$\mathcal{A}[\bar{B}_s \to \mu^+ \mu^- \gamma] = -e\frac{\alpha_{\text{em}}}{\sqrt{2}\pi} V_{tb} V_{ts}^* \varepsilon_\mu^* \left[\sum_{i=1}^{9} C_i H_i^{\mu\nu} L_{V\nu} + C_{10}\left(H_{10}^{\mu\nu} L_{A\nu} - \frac{i}{2} f_{B_s} L_A^{\mu\nu} p_\nu \right) \right] \tag{8}$$

where we have defined

$$L_V^\nu = \bar{u}(p_2)\gamma^\nu v(p_1) , \qquad L_A^\nu = \bar{u}(p_2)\gamma^\nu \gamma^5 v(p_1) . \tag{9}$$

The last term in Eq. (8), which only depends on the leptonic tensor $L_A^{\mu\nu}$ [9] and on the decay constant f_{B_s} of the B_s meson, corresponds to the final-state radiation (FSR) contribution (to which we refer in the following as to the point-like contribution). The non-perturbative contribution to the structure-dependent part of the amplitude is instead encoded in the hadronic tensors $H_i^{\mu\nu}$, which can be grouped into three different categories: the contribution from the semileptonic operators $\mathcal{O}_{9-10}$, the contribution from the photon penguin operator $\mathcal{O}_7$, and finally the contributions from the four-fermion operators $\mathcal{O}_{1-6}$ and from the chromomagnetic penguin operator $\mathcal{O}_8$.

The contributions from the semileptonic operators are depicted graphically in Figure 1. For these contributions the real photon γ is emitted directly from one of the two quarks. The corresponding tensors $H_{9-10}^{\mu\nu}$ are given by

$$H_9^{\mu\nu}(p, k) = H_{10}^{\mu\nu}(p, k) = i \int d^4y \, e^{iky} \, \hat{T}\langle 0|[\bar{s}\gamma^\nu P_L b](0) J_{\text{em}}^\mu(y)|\bar{B}_s(\boldsymbol{p})\rangle$$

$$= -i\left[g^{\mu\nu}(k \cdot q) - q^\mu k^\nu\right]\frac{F_A}{2m_{B_s}} + \varepsilon^{\mu\nu\rho\sigma} k_\rho q_\sigma \frac{F_V}{2m_{B_s}} , \tag{10}$$

where J_{em}^μ is the e.m. current, and $\hat{T}$ represents "time-ordered". The two tensors are parameterized by vector (F_V) and axial (F_A) form factors, which are scalar functions of the single invariant of the process, namely the di-muon

FIG. 1: *Graphical representation of the contribution to the $\bar{B}_s \to \mu^+\mu^-\gamma$ decay amplitude from the semileptonic operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$.*

FIG. 2: *Graphical representation of the contribution to the $\bar{B}_s \to \mu^+\mu^-\gamma$ decay amplitude from the photon penguin operator $\mathcal{O}_7$ in which the final-state photon is emitted directly from one of the valence quarks.*

inviarant mass $q^2 = (p-k)^2$. In the following, as in our previous papers, we present the form factors as functions of the dimensionless variable

$$x_\gamma = \frac{2\,p \cdot k}{m_{B_s}^2} = 1 - \frac{q^2}{m_{B_s}^2} \;, \qquad\qquad 0 \le x_\gamma \le 1 - \frac{4m_\mu^2}{m_{B_s}^2} \;, \tag{11}$$

which in the decaying meson rest frame reduces to $x_\gamma = 2E_\gamma/m_{B_s}$, where E_γ is the energy of the emitted photon.

The contribution from the photon penguin operator $\mathcal{O}_7$ is illustrated in Figures 2 and 3. In this case there are two types of contribution: those in which the final-state real photon is emitted by the valence quarks (Figure 2), and those in which the real photon is emitted by the penguin vertex (Figure 3). We indicate by $H_{7A}^{\mu\nu}$ and $H_{7B}^{\mu\nu}$ the hadronic tensor corresponding to the first and second contribution respectively, with $H_7^{\mu\nu} = H_{7A}^{\mu\nu} + H_{7B}^{\mu\nu}$. The hadronic tensor $H_{7A}^{\mu\nu}$ is given by

$$H_{7A}^{\mu\nu}(p,k) = i\frac{2m_b}{q^2} \int d^4y \; e^{iky} \; \hat{\mathrm{T}}\langle 0| \left[-i\bar{s}\sigma^{\nu\rho}q_\rho P_R b \right](0) J_{\mathrm{em}}^\mu(y)|\bar{B}_s(\boldsymbol{p})\rangle$$

$$= -i\left[g^{\mu\nu}(k \cdot q) - q^\mu k^\nu \right]\frac{F_{TA}m_b}{q^2} + \varepsilon^{\mu\nu\rho\sigma}k_\rho q_\sigma \frac{F_{TV}m_b}{q^2} \tag{12}$$

where the two tensor form factors F_{TV} and F_{TA} are again scalar functions of x_γ. Exploiting the relation $\gamma^5\sigma^{\mu\nu} = -i\varepsilon^{\mu\nu\rho\sigma}\sigma_{\rho\sigma}/2$ one can show that the two tensor form factors obey the kinematical constraint $F_{TV}(1) = F_{TA}(1)$ (see also Ref. [4]). The hadronic tensor $H_{7B}^{\mu\nu}$, corresponding to the emission of the real photon from the FCNC vertex is instead given by

$$H_{7B}^{\mu\nu}(p,k) = i\frac{2m_b}{q^2} \int d^4y \; e^{iqy} \; \hat{\mathrm{T}}\langle 0| \left[-i\bar{s}\sigma^{\mu\rho}k_\rho P_R b \right](0) J_{\mathrm{em}}^\nu(y)|\bar{B}_s(\boldsymbol{p})\rangle$$

$$= -i\left[g^{\mu\nu}(k \cdot q) - q^\mu k^\nu \right]\frac{\bar{F}_{TA}m_b}{q^2} + \varepsilon^{\mu\nu\rho\sigma}k_\rho q_\sigma \frac{\bar{F}_{TV}m_b}{q^2} \tag{13}$$

In this case, as discussed in Ref. [4], the two form factors obey $\bar{F}_{TV}(x_\gamma) = \bar{F}_{TA}(x_\gamma) \equiv \bar{F}_T(x_\gamma)$.[3] Moreover at $x_\gamma = 1$, i.e. at $q^2 = k^2 = 0$ one has

$$F_{TV}(1) = F_{TA}(1) = \bar{F}_T(1) \;. \tag{14}$$

----

[3] Again this can be shown making use of the relation $\gamma^5\sigma^{\mu\nu} = -i\varepsilon^{\mu\nu\rho\sigma}\sigma_{\rho\sigma}/2$.

FIG. 3: *Graphical representation of the contribution to the $\bar{B}_s \to \mu^+\mu^-\gamma$ amplitude from the photon penguin operator $\mathcal{O}_7$ in which the final-state photon is emitted from the penguin operator.*
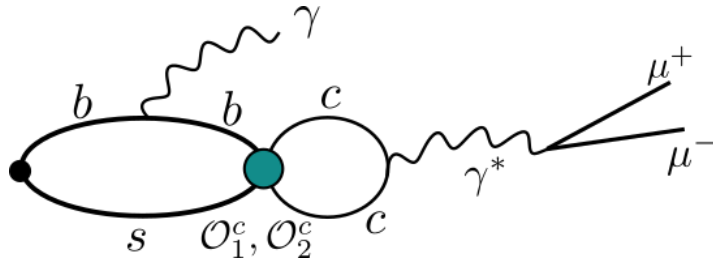

FIG. 4: *Graphical representation of the contribution to the $\bar{B}_s \to \mu^+\mu^-\gamma$ amplitude from the four-quarks operators $\mathcal{O}_1^c$ and $\mathcal{O}_2^c$ with the virtual photon γ^* emitted by the charm loop (the corresponding diagram with the real photon emitted from the strange valence quark is not shown).*

The form factor $\bar{F}_T(x_\gamma)$ is the most difficult to determine on the lattice. When the virtual photon γ^* is emitted by a valence strange quark, the presence of intermediate $J^P = 1^-$ $s\bar{s}$ resonance states forbids the analytic continuation to Euclidean spacetime of the relevant Minkowskian correlation functions needed to evaluate $\bar{F}_T(x_\gamma)$. In this case, in order to evaluate the form factor $\bar{F}_T$, we rely on the spectral density reconstruction technique developed in Ref. [8]. It is the form factors F_V, F_A, F_{TV}, F_{TA} and $\bar{F}_T$ which we evaluate from first principles via lattice QCD simulations. In the following we sometimes refer to them as local form factors.

The remaining contributions to the amplitude $\mathcal{A}[\bar{B}_s \to \mu^+\mu^-\gamma]$ are those corresponding to the four-quark operators and to the chromomagnetic penguin operator. The corresponding hadronic tensors $H_{i=1-6,8}^{\mu\nu}$ are given by

$$H_{i=1-6,8}^{\mu\nu}(p,k) = \frac{(4\pi)^2}{q^2} \int d^4y \, d^4x \, e^{iky}e^{iqx} \hat{\mathrm{T}}\langle 0|J_{\mathrm{em}}^\mu(y)J_{\mathrm{em}}^\nu(x)\mathcal{O}_i(0)|\bar{B}_s(\boldsymbol{p})\rangle \, . \tag{15}$$

In the high-q^2 region which we consider, as discussed in Ref. [6], the contribution of the presently neglected terms from $\mathcal{O}_{i=1-6,8}$ is expected to be small, since they are of higher-order in the $1/m_b$ expansion. Among them, one of the most important contribution is that of the charming-penguin diagram depicted in Figure 4, due to the presence of broad charmonium resonance contributions, which are near or within the region of q^2 we have explored.

To take into account this contribution, we follow Refs. [4–6] and include the charming-penguin diagram in Figure 4 as a q^2-dependent shift of the Wilson coefficient C_9, namely

$$C_9 \to C_9^{\mathrm{eff}}(q^2) = C_9 + \Delta C_9(q^2) \, , \tag{16}$$

where $\Delta C_9(q^2)$ can be phenomenologically modelled as a sum over the contributions from all the $J^P = 1^-$ charmonium resonances [4, 5, 11]

$$\Delta C_9(q^2) = -\frac{9\pi}{\alpha_{\mathrm{em}}^2} \left(C_1 + \frac{C_2}{3}\right) \sum_V |k_V| e^{i\delta_V} \frac{m_V \, B(V \to \mu^+\mu^-)\,\Gamma_V}{q^2 - m_V^2 + im_V\Gamma_V} \, , \tag{17}$$

where Γ_V is the total decay width of the resonance V, m_V its mass, and $B(V \to \mu^+\mu^-)$ the branching fraction for the decay into a di-muon. The coefficient k_V and the phase shift δ_V take into account deviations from the factorization approximation, which corresponds to $\delta_V = |k_V| - 1 = 0$. The values of (some of) the parameters entering Eq. (17), for the low-lying resonances, can be taken from experiments, but clearly this introduces a systematic error in our prediction.

In the evaluation of the rate, we will use Eq. (17), and estimate the associated systematic error in a conservative way by varying the input parameters over a sufficiently broad range. The conservative systematic we associate to the missing charming-penguin diagram in Figure 4 is expected to be sufficiently large to cover the uncertainty of all the other missing contributions from Eq. 15. In the future, in order to remove this source of systematic uncertainty and reduce the theory error, it will be extremely important to evaluate $H_{i=1-6,8}^{\mu\nu}$ on the lattice, in particular the charming-penguin contribution of Figure 4, which we plan to do. We now turn to the discussion of the calculation of the local form factors F_V, F_A, F_{TV}, F_{TA} and $\bar{F}_T$.

III. THE LOCAL FORM FACTORS F_V, F_A, F_{TV} AND F_{TA}

As illustrated in the previous section, the form factors $F_W(x_\gamma)$, $W = \{V, A, TV, TA\}$, can be computed from QCD matrix elements involving the e.m. and the following currents

$$J_A^\nu = Z_V \bar{s}\gamma^\nu\gamma_5 b, \qquad J_V^\nu = Z_A \bar{s}\gamma^\nu b, \qquad J_{TA}^\nu = -iZ_T(\mu)\bar{s}\sigma^{\nu\rho}\gamma_5 b \frac{q_\rho}{m_{B_s}}, \qquad J_{TV}^\nu = -iZ_T(\mu)\bar{s}\sigma^{\nu\rho}b \frac{q_\rho}{m_{B_s}}, \qquad (18)$$

where, as already stated, $q^\nu = p^\nu - k^\nu$ is the four-momentum of the charged muon pair. In the previous equation we have introduced the scheme- and scale-dependent renormalization constant (RC) $Z_T(\mu)$ of the tensor current, and the (finite) RCs of the axial and vector currents that in twisted-mass QCD are chirally rotated with respect to the ones of standard Wilson fermions. From now on, we work in the rest frame of the decaying meson and thus set $p = (m_{B_s}, \mathbf{0})$. In terms of the hadronic tensors

$$H_W^{\mu\nu}(p,k) \equiv i \int d^4y \; e^{iky} \; \hat{\mathrm{T}}\langle 0|J_W^\nu(0)J_{\mathrm{em}}^\mu(y)|\bar{B}_s(\mathbf{0})\rangle \;, \qquad W = \{V, A, TV, TA\} \;, \qquad (19)$$

and recalling the definitions given in Eqs. (10)-(12), one has that

$$H_A^{\mu\nu}(p,k) = i\left[(k\cdot q)g^{\mu\nu} - q^\mu k^\nu\right]\frac{F_A}{m_{B_s}} \;, \qquad H_V^{\mu\nu}(p,k) = \epsilon^{\mu\nu\rho\sigma}k_\rho p_\sigma \frac{F_V}{m_{B_s}}$$

$$H_{TA}^{\mu\nu}(p,k) = -i\left[(k\cdot q)g^{\mu\nu} - q^\mu k^\nu\right]\frac{F_{TA}}{m_{B_s}} \;, \qquad H_{TV}^{\mu\nu}(p,k) = \epsilon^{\mu\nu\rho\sigma}k_\rho p_\sigma \frac{F_{TV}}{m_{B_s}} \;. \qquad (20)$$

In Section III and Appendix B of Ref. [12] we show in detail that for the emission of a real photon, the hadronic tensor $H_W^{\mu\nu}$ can be extracted for all values of x_γ from the Euclidean three-point correlation function:

$$B_W^{\mu\nu}(t,\mathbf{k}) = a\sum_{t_y=0}^{T} a^3\sum_{\mathbf{y}} a^3\sum_{\mathbf{x}} \left(\theta(T/2 - t_y) + \theta(t_y - T/2)e^{-E_\gamma T}\right) e^{t_y E_\gamma - i\mathbf{k}\cdot\mathbf{y}} \; \langle 0|\hat{\mathrm{T}}\left[J_W^\nu(t,\mathbf{0})J_{\mathrm{em}}^\mu(t_y,\mathbf{y})\phi_{B_s}^\dagger(0,\mathbf{x})\right]|0\rangle \;, \quad (21)$$

where T is the temporal extent of the lattice[4], a is the lattice spacing, and $\phi_{B_s}^\dagger$ is an interpolating operator with the quantum numbers to create the $\bar{B}_s$ meson which, as in Ref. [7], we smear using Gaussian smearing. For the electromagnetic current J_{em}^μ we use the exactly-conserved point-split lattice operator

$$J_{\mathrm{em}}^\mu(x) = \sum_f J_f^\mu = -\sum_f q_f \left\{\bar{\psi}_f(x)\frac{ir_f\gamma_5 - \gamma^\mu}{2}U^\mu(x)\psi_f(x+\hat{\mu}) - \bar{\psi}_f(x+\hat{\mu})\frac{ir_f\gamma_5 + \gamma^\mu}{2}U^\mu(x)^\dagger\psi_f(x)\right\} \;. \qquad (22)$$

In the forward half of the lattice $0 \ll t \ll T/2$ one has

$$R_W^{\mu\nu}(t,\mathbf{k}) \equiv \frac{2m_{B_s}}{e^{-t(m_{B_s} - E_\gamma)}\langle\bar{B}_s(\mathbf{0})|\phi_{B_s}^\dagger(0)|0\rangle} B_W^{\mu\nu}(t,\mathbf{k}) = H_W^{\mu\nu}(p,k) + \cdots \;, \qquad (23)$$

where the ellipsis indicates terms that vanish exponentially in the large t limit. Eq. (21) is valid for $t < T/2$, however, as explained in Appendix B of Ref. [12], $H_W^{\mu\nu}(p,k)$ can also be obtained from the backward half of the lattice $T/2 \ll t \ll T$ exploiting time-reversal symmetry. The Wick contractions of the correlation function in Eq. (21) give rise to two distinct topologies of Feynman diagrams, namely to quark-line connected and quark-line disconnected diagrams; these are

---

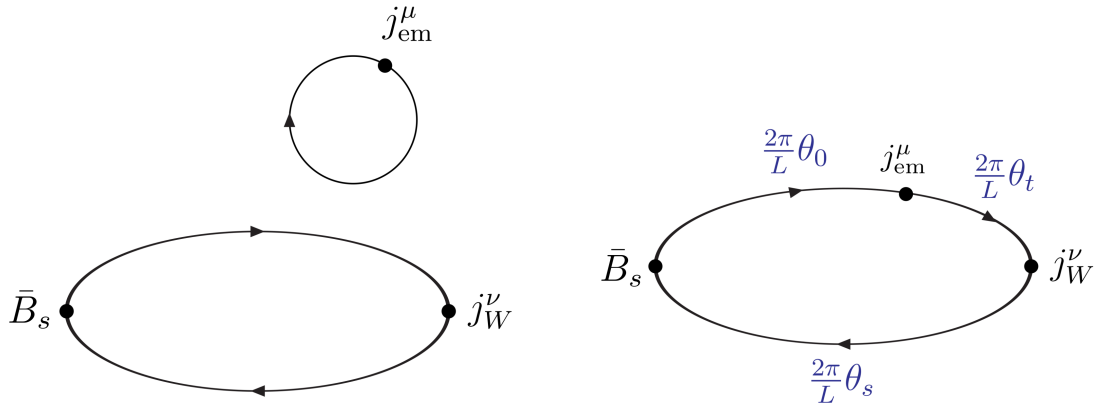[4] T is not to be confused with $\hat{\mathrm{T}}$ which represents "time-ordered".

FIG. 5: *The diagram on the left represents the quark-line disconnected contributions to the correlation function $B_W^{\mu\nu}$ in which the photon is emitted by a sea quark. In our numerical simulations we work in the electroquenched approximation and neglect such diagrams. The diagram on the right represents the quark-line connected contributions and illustrates our choice of the spatial boundary conditions, which allow us to set arbitrary values for the meson and photon spatial momenta. The spatial momenta of the valence quarks in terms of the twisting angles are as indicated. Each diagram implicitly includes all orders in QCD.*

illustrated in Figure 5. In the disconnected diagrams the photon is emitted from a sea quark. This contribution vanishes in the SU(3)-symmetric limit and when loop of charmed and heavier quarks are omitted, and is neglected in the present study; this is the so-called electroquenched approximation. We focus instead on the calculation of the dominant, quark-connected contributions for which only the strange- and bottom-quark components of the electromagnetic current $J_{\rm em}^\mu$ contribute.

As explained in Ref. [12], it is possible to use twisted boundary conditions to assign arbitrary values to momenta of the photon and $\bar{B}_s$-meson, $\boldsymbol{k}$ and $\boldsymbol{p}$ respectively, at the price of violations of unitarity which vanish exponentially with the lattice extent L [13–15]. This is achieved by treating the two quark propagators beginning or ending at y, i.e. the point at which the electromagnetic current is inserted in the right-hand diagram of Figure 5, as corresponding to two distinct quark fields ψ^0, ψ^t having the same mass and quantum number, but satisfying different spatial boundary conditions. Defining ψ^s to be the *spectator* quark-field in the right-hand diagram of Figure 5, we set the spatial boundary conditions of the three quark fields ψ^0, ψ^t, ψ^s as follows:

$$\psi^r(x + \boldsymbol{n}L) = \exp\{(2\pi i \boldsymbol{n}\cdot\boldsymbol{\theta}_r)\}\psi^r(x)\,, \qquad r = \{0, t, s\}\,, \tag{24}$$

where $\boldsymbol{\theta}_{\{0,t,s\}}$ are arbitrary spatial-vectors of angles, in terms of which the photon and meson lattice momenta can be written as

$$\boldsymbol{p} = \frac{2}{a}\sin\left(\frac{a\pi}{L}\,(\boldsymbol{\theta}_0 - \boldsymbol{\theta}_s)\right)\,, \qquad \boldsymbol{k} = \frac{2}{a}\sin\left(\frac{a\pi}{L}\,(\boldsymbol{\theta}_0 - \boldsymbol{\theta}_t)\right)\,, \tag{25}$$

We choose the photon momentum to be in the z-direction, $\boldsymbol{k} = (0, 0, k_z)$, and set

$$\boldsymbol{\theta}_0 = \boldsymbol{\theta}_s = \boldsymbol{0}\,, \qquad \boldsymbol{\theta_t} = (0, 0, \theta_t). \tag{26}$$

With such a choice of kinematics, the form factors can be obtained from the large time behaviour, $0 \ll t \ll T/2$, of the

ensemble	β	V/a^4	a (fm)	$a\mu_\ell$	m_π (MeV)	L (fm)	N_g
A48	1.726	$48^3 \cdot 128$	0.09075 (54)	0.00120	174.5 (1.1)	4.36	109
B64	1.778	$64^3 \cdot 128$	0.07957 (13)	0.00072	140.2 (0.2)	5.09	400
C80	1.836	$80^3 \cdot 160$	0.06821 (13)	0.00060	136.7 (0.2)	5.46	72
D96	1.900	$96^3 \cdot 192$	0.05692 (12)	0.00054	140.8 (0.2)	5.46	100

TABLE I: *Parameters of the ETMC ensembles used in this work. We present the light-quark bare mass, $a\mu_\ell = a\mu_u = a\mu_d$, the lattice spacing a, the pion mass m_π, the lattice size L and the number of gauge configurations N_g that have been used for each ensemble. The values of the lattice spacing are determined as explained in Appendix B of Ref. [22] using the value $f_\pi^{isoQCD} = 130.4(2)$ MeV for the pion decay constant.*

following estimators

$$R_V(t, \boldsymbol{k}) \equiv \frac{1}{2k_z} \left(R_V^{12}(t, \boldsymbol{k}) - R_V^{21}(t, \boldsymbol{k}) \right) \xrightarrow[0 \ll t \ll T/2]{} F_V(x_\gamma), \tag{27}$$

$$R_A(t, \boldsymbol{k}) \equiv \frac{i}{2E_\gamma} \left(R_A^{11}(t, \boldsymbol{k}) + R_A^{22}(t, \boldsymbol{k}) \right) \xrightarrow[0 \ll t \ll T/2]{} F_A(x_\gamma), \tag{28}$$

$$R_{TV}(t, \boldsymbol{k}) \equiv \frac{1}{2k_z} \left(R_{TV}^{12}(t, \boldsymbol{k}) - R_{TV}^{21}(t, \boldsymbol{k}) \right) \xrightarrow[0 \ll t \ll T/2]{} F_{TV}(x_\gamma), \tag{29}$$

$$R_{TA}(t, \boldsymbol{k}) \equiv -\frac{i}{2E_\gamma} \left(R_{TA}^{11}(t, \boldsymbol{k}) + R_{TA}^{22}(t, \boldsymbol{k}) \right) \xrightarrow[0 \ll t \ll T/2]{} F_{TA}(x_\gamma). \tag{30}$$

For each form factor it is useful to distinguish the two contributions due to the emission of the real photon from the bottom and strange quarks (left and right diagrams in Figures 1 and 2). We denote the two contributions by F_W^b and F_W^s for $W = \{V, A, TV, TA\}$. They are simply obtained by setting respectively the electric charges $q_s = 0$ and $q_b = 0$ in all the previous formulae. A minor complication arises in the axial channel $W = A$ due to the presence of a point-like contribution, proportional to $q_b f_{B_s}$ and $-q_s f_{B_s}$ respectively in F_A^b and F_A^s, which then cancels in the sum of the two contributions due to $q_b = q_s$. This point-like contribution, which is always present in the radiative leptonic decays of charged pseudoscalar mesons [12], can however, be easily removed by calculating the following zero-momentum-subtracted estimator

$$R_A^{(s,b)}(t, \boldsymbol{k}) \equiv \frac{i}{2E_\gamma} \left(R_A^{11,(s,b)}(t, \boldsymbol{k}) - R_A^{11,(s,b)}(t, \boldsymbol{0}) + R_A^{22,(s,b)}(t, \boldsymbol{k}) - R_A^{22,(s,b)}(t, \boldsymbol{0}) \right) \xrightarrow[0 \ll t \ll T/2]{} F_A^{(s,b)}(x_\gamma). \tag{31}$$

We refer the reader to Ref. [12] for more details on the removal of the point-like contribution.

A.   Numerical results for F_V, F_A, F_{TV}, F_{TA}

We now turn to the discussion of our numerical results for F_V, F_A, F_{TV} and F_{TA}. They have been obtained using the gauge field configurations generated by the Extended Twisted Mass Collaboration (ETMC) employing the Iwasaki gluon action [16] and $N_f = 2 + 1 + 1$ flavours of Wilson-Clover twisted-mass fermions at maximal twist [17]. This framework guarantees the automatic $\mathcal{O}(a)$ improvement of parity-even observables [18, 19]. A detailed description of the ETMC ensembles can be found in Refs. [20–23], and we also refer to Ref. [7] for additional information on the tuning of the sea and valence quark masses. In Table I we present the parameters of the ETMC ensembles that have been used in the present computation, while in Table II we collect the relevant RCs used to renormalize the vector, axial, and tensor currents.

The presently available lattice spacings are not small enough to perform simulations at the physical bottom quark mass. For this reason our strategy to reach the physical B_s meson mass, is to perform simulations for a series of heavy-strange quark masses, and then extrapolate to the physical point using heavy-quark effective theory (HQET) scaling relations, to be discussed in the next sections. For each of the ensembles of Table I, we have performed simulations at five different values of m_{H_s}, the mass of the lightest pseudoscalar meson composed of a valence heavy quark of mass m_h

ensemble	Z_V	Z_A	$Z_T(\overline{\text{MS}}, 5\,\text{GeV})$
A48	0.68700 (15)	0.7284 (18)	0.7541 (98)
B64	0.706379 (24)	0.74294 (24)	0.7735 (93)
C80	0.725404 (19)	0.75830 (16)	0.7928 (85)
D96	0.744108 (12)	0.77395 (12)	0.8141 (74)

TABLE II: *The values of the vector (Z_V), axial (Z_A), and tensor (Z_T) renormalization constants, for the ETMC ensembles of Table I. Z_T values were kindly provided to us by the ETMC, and are from a preliminary analysis [24]. For the present work, we increased their uncertainties by a factor of 3. The scale-independent renormalization constants Z_V and Z_A have been determined in Ref. [22] using Ward-identity methods.*

and a strange antiquark with mass m_s. The five values correspond to the following five m_h/m_c ratios (m_c is the mass of the charm quark determined by the condition $m_{\eta_c} = 2.984(4)$ GeV, see Refs. [8, 22]):

$$\frac{m_h}{m_c} \simeq 1,\ 1.5,\ 2,\ 2.5,\ 3\ . \tag{32}$$

Such values of the heavy quark masses m_h give rise to heavy-strange meson masses m_{H_s} in the range $m_{H_s}/m_{D_s} \in [1, 2]$. For each ensemble and heavy quark mass m_h, we evaluate the Euclidean three point function $B_W^{\mu\nu}(t; k, p)$ at four evenly-spaced values of the dimensionless variable x_γ:

$$x_\gamma = \frac{2E_\gamma}{m_{H_s}} = 0.1, 0.2, 0.3, 0.4\ . \tag{33}$$

For an illustration of the quality of the plateaus, we present in Figure 6 the estimators $R_W^{(s,b)}(t, x_\gamma) \equiv R_W^{(s,b)}(t, (0, 0, k_z(x_\gamma)))$, $W = \{V, A, TV, TA\}$, obtained at $x_\gamma = 0.2$ on the finest lattice spacing ensemble (D96) for $m_h/m_c \simeq 2$. In each figure the blue band shows our estimate of the corresponding form factor, obtained from a constant fit in the region where the estimators $R_W^{(s,b)}(t, x_\gamma)$ display a plateau. The band already includes the systematic error due to the choice of the fit-interval, which is estimated by performing a second fit shifting the fit-interval forward in time by an amount $\Delta T = 0.4, 0.35, 0.30, 0.27, 0.25\,\text{fm}$, respectively for $m_h/m_c \simeq 1, 1.5, 2, 2.5, 3$, and then adding the difference between the central values obtained in the two different fits as a systematic error[5]. For the tensor form factors the results are obtained using the preliminary values of $Z_T(\mu)$ in the $\overline{\text{MS}}$ scheme at $\mu = 5$ GeV, provided to us by the ETMC [24].

The ensembles of Table I all correspond to lattices with a spatial extent in the range $L \simeq 4.4$ - $5.4\,\text{fm}$. These volumes are expected to be large enough for the finite size effects (FSEs) on the form factors to be small. For the smallest heavy quark mass considered, $m_h = m_c$, and for the form factors F_A and F_V, this has been explicitly checked in Ref. [7] using an additional ensemble, the B96, which has a large spatial extent L of more than 7.5 fm. Here, using the B96 ensemble, we have checked that FSEs are very small (at the level of our statistical uncertainty or smaller) also for the tensorial form factors F_{TV} and F_{TA}, and we therefore consider our results on the ensembles listed in Table I as infinite-volume quantities.

Next we consider the cut-off effects. For each value of x_γ and m_{H_s}, the extrapolation to the continuum limit is performed using the following Ansatz

$$F_W^{(b,s)}(x_\gamma, m_{H_s}, a) = F_W^{(b,s)}(x_\gamma, m_{H_s})\left(1 + D_W^{(b,s)}(x_\gamma, m_{H_s})\, a^2\right)\ , \qquad W = \{V, A, TV, TA\}\ , \tag{34}$$

where $F_W^{(b,s)}(x_\gamma, m_{H_s})$ and $D_W^{(b,s)}(x_\gamma, m_{H_s})$ are fit parameters which depend on x_γ and m_{H_s}, and are different for the four channels $W = \{V, A, TV, TA\}$ and for the two contributions F_W^s and F_W^b. We estimate the systematic uncertainty due to the continuum-limit extrapolation by performing two different linear extrapolations: in the first one we include the full dataset, and in the second one we remove the measurements on the ensemble with the largest lattice spacing (A48). The two results are combined as follows: let f_A and f_B represent generically the continuum values of $F_W(x_\gamma)$, for a given $W = \{V, A, TV, TA\}$ and $x_\gamma \in \{0.1, \ldots, 0.4\}$, obtained respectively from the linear fit by including or omitting the result at the coarsest lattice spacing. We determine the final central value $\bar{f}$ through a weighted average of the form

$$\bar{f} = w_A\, f_A + w_B\, f_B, \qquad w_A + w_B = 1\ . \tag{35}$$

---

[5] The choice of ΔT, for each value of m_h, has been adjusted so that this is small enough to avoid the region of large times where the signal-to-noise ratio of the estimator $R_W^{s,b}$ is very small, and at the same time large enough to provide a reasonable estimate of the systematics due to the choice of the fit interval.
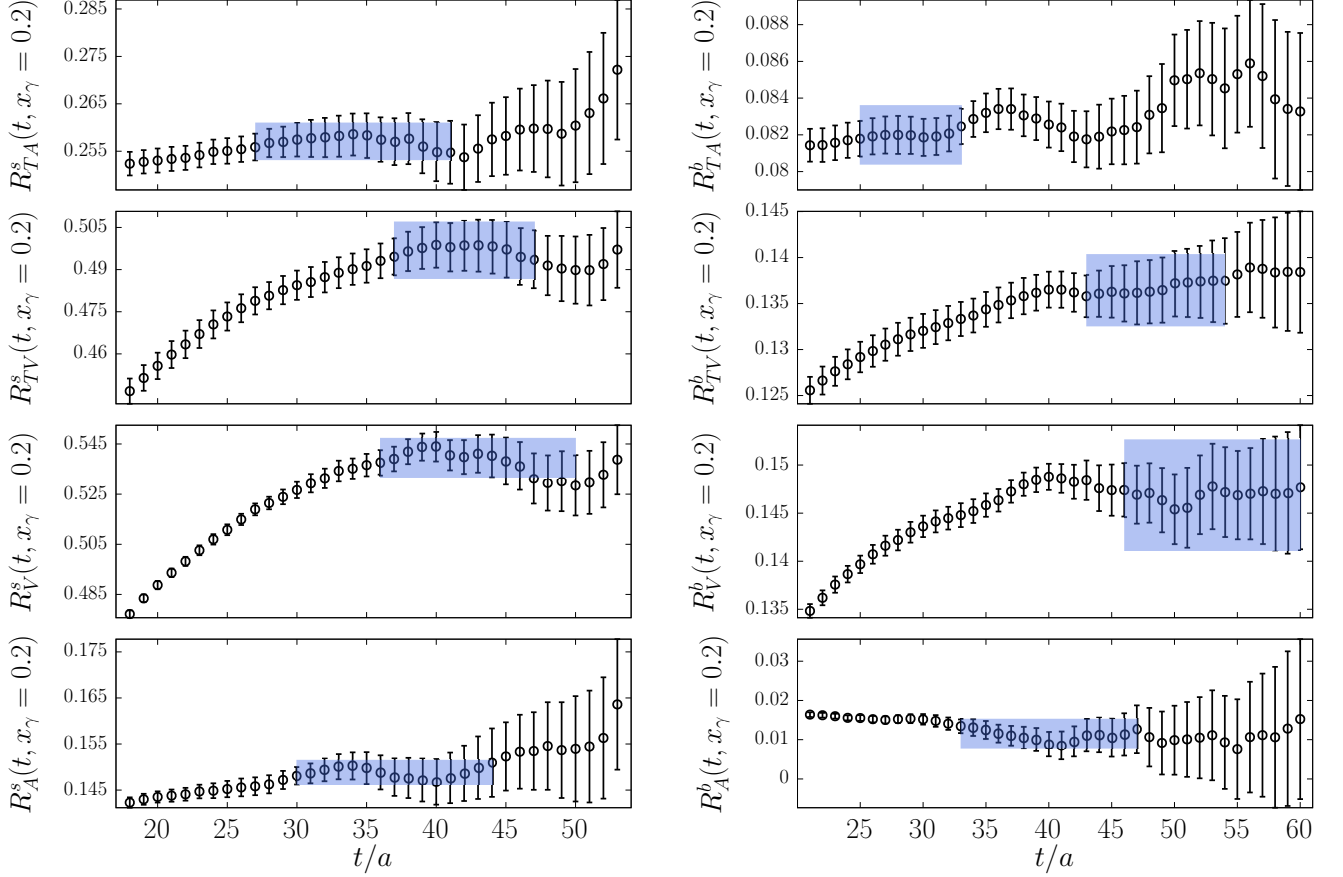
FIG. 6: *The estimators $R_W^s(t,x_\gamma)$ (left) and $R_W^b(t,x_\gamma)$ (right) for $W = \{V, A, TV, TA\}$. The data corresponds to the D96 ensemble at $x_\gamma = 0.2$ and for $m_h \simeq 2m_c$. The blue bands show our estimates of the form factors from a constant fit in the region where the estimators $R_W^{(s,b)}(t,x_\gamma)$ display a plateau.*

Our estimate of the systematic error, which is added (linearly to be conservative) to the statistical uncertainty, is then obtained using

$$\sigma_{\rm syst}^2 = \sum_{i=A,B} w_i \ (f_i - \bar{f})^2 \ . \tag{36}$$

The weights w_i, with $i = \{A, B\}$, are chosen according to the Akaike Information Criterion [25] (AIC), namely

$$w_i \propto e^{-\left(\chi_i^2 + 2N_{\rm pars}^{(i)} - N_{\rm data}^{(i)}\right)/2} \ , \tag{37}$$

where χ_i^2 is the total χ^2 obtained in the i-th fit, and $N_{\rm pars}^{(i)}$ and $N_{\rm meas}^{(i)}$ are the corresponding number of fit parameters and measurements.

In Figures 7 and 8 we show the results of our continuum fits, for the smallest ($x_\gamma = 0.1$) and largest ($x_\gamma = 0.4$) simulated values of x_γ. The fits shown in the figures are those for which the full dataset has been used. Clearly for large quark masses m_h, as a consequence of the Parisi-Lepage theorem [26, 27], the statistical noise of the data rapidly increases. The quality of the fits is very good, and in Figure 9 we show the histogram of the reduced χ^2 distribution corresponding to the 160 continuum extrapolations we have performed.

B. Extrapolating the results for the local form factors F_V, F_A, F_{TV}, F_{TA} to the physical B_s meson

In this section we discuss the asymptotic formulae used to extrapolate the form factors, computed for $m_{H_s} \in [m_{D_s}, 2m_{D_s}]$, to the physical point $m_{H_s} = m_{B_s} \simeq 5.367 \, {\rm GeV}$. For heavy quark masses m_h and energetic photons, there
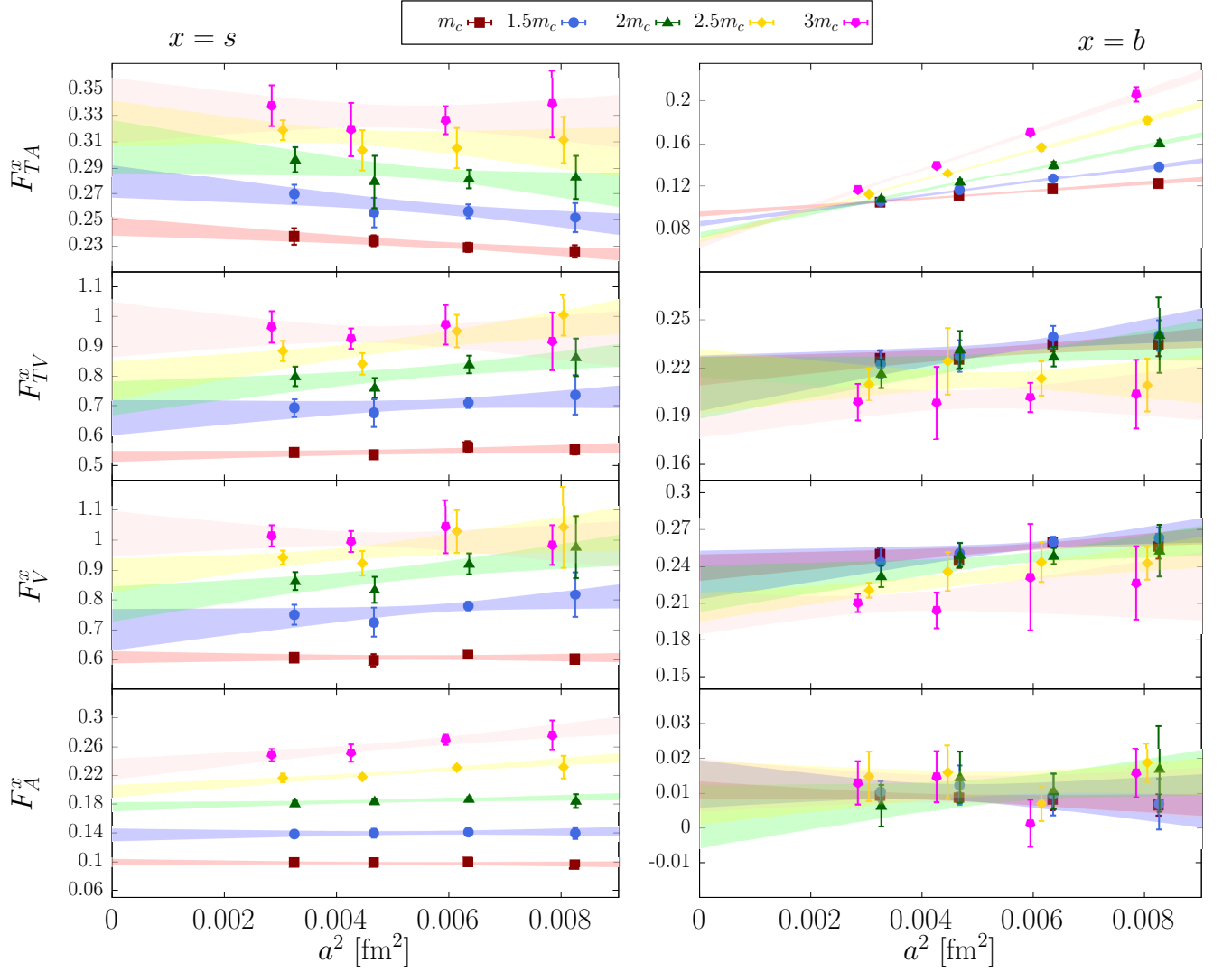
FIG. 7: *Continuum limit extrapolation of the lattice data for F_W^s (left) and F_W^b (right) for $x_\gamma = 0.1$. The transparent bands correspond to the best-fit function obtained in the linear a^2 fit employing the full dataset. In the panels, the different colors correspond to different values of the heavy quark mass m_h.*

are elegant and simple relations relating the four form factors. In Ref. [28] (see also [29–32]), the authors studied in detail the behaviour of the axial and vector form factors contributing to the radiative $B \to \gamma \ell \nu$ decay amplitude in the framework of the HQET and large-photon-energy expansions. The relations derived in Ref. [28] imply that up to (and including) order $\mathcal{O}(1/m_{H_s}, 1/E_\gamma)$ terms in the heavy-quark and large-photon-energy expansion, the axial and vector form factors F_A and F_V are given by

$$\frac{F_V(x_\gamma, m_{H_s})}{f_{H_s}} = \frac{|q_s|}{x_\gamma} \left(\frac{R(E_\gamma, \mu)}{\lambda_B(\mu)} + \xi(x_\gamma, m_{H_s}) + \frac{1}{m_{H_s} x_\gamma} + \frac{|q_b|}{|q_s|} \frac{1}{m_h} \right), \tag{38}$$

$$\frac{F_A(x_\gamma, m_{H_s})}{f_{H_s}} = \frac{|q_s|}{x_\gamma} \left(\frac{R(E_\gamma, \mu)}{\lambda_B(\mu)} + \xi(x_\gamma, m_{H_s}) - \frac{1}{m_{H_s} x_\gamma} - \frac{|q_b|}{|q_s|} \frac{1}{m_h} \right), \tag{39}$$

where $f_{\bar{H}_s}$ is the decay constant of the $\bar{H}_s = \bar{s}h$ pseudoscalar meson of mass $m_{\bar{H}_s}$, $\lambda_B(\mu)$ is the first inverse moment of the B_s-meson light-cone distribution amplitude (LCDA), and $R(E_\gamma, \mu) = 1 + \mathcal{O}(\alpha_s)$ is a radiative correction factor that is the same for F_V and F_A. Finally, $\xi(x_\gamma, m_{H_s})$ is a power-suppressed term, common to both form factors, that
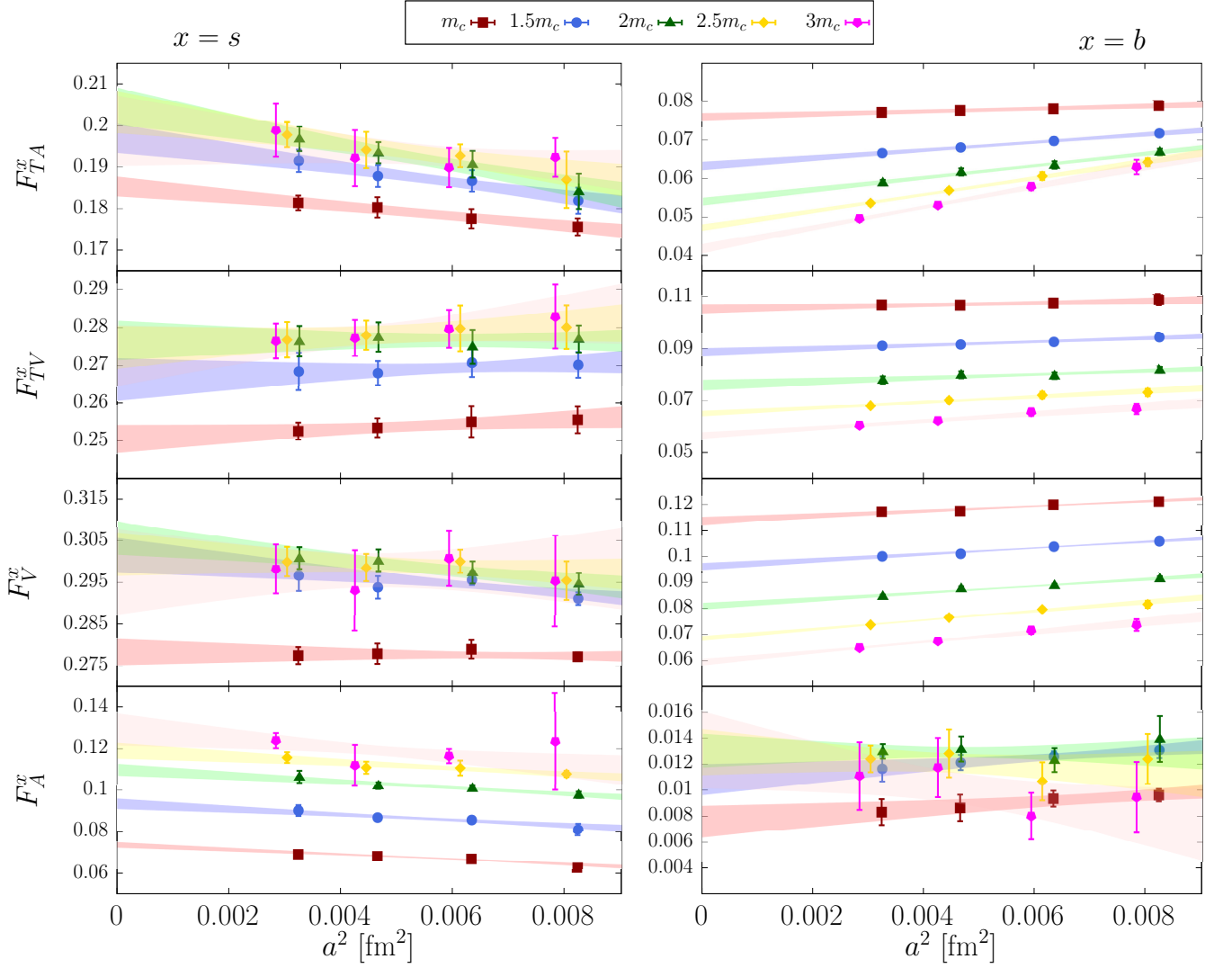
FIG. 8: *Same as in Figure 7 for $x_\gamma = 0.4$.*

can be written as [33]

$$\xi(x_\gamma, m_{H_s}) = \frac{A}{m_{H_s}} + \frac{B}{m_{H_s} x_\gamma} . \tag{40}$$

In Eqs. (38)-(39), perturbative radiative corrections to the subleading terms and $\mathcal{O}(m_s/(m_{H_s}^2))$ terms have been neglected. The leading contribution to the form factors comes from the emission of the photon from the strange quark. Radiation from the heavy quark is suppressed by a factor proportional to $1/m_{H_s}$ and the corresponding subleading terms are proportional to $|q_b|$ in Eqs. (38)-(39).

The large mass/photon-energy behaviour of the tensor form factors F_{TA} and F_{TV} including order $\mathcal{O}(1/m_{H_s}, 1/E_\gamma)$ corrections has been investigated in Ref. [9] and is given by

$$\frac{F_{TV}(x_\gamma, m_{H_s}, \mu)}{f_{H_s}} = \frac{|q_s|}{x_\gamma} \left(\frac{R_T(E_\gamma, \mu)}{\lambda_B(\mu)} + \xi(x_\gamma, m_{H_s}) + \frac{1 - x_\gamma}{m_{H_s} x_\gamma} + \frac{|q_b|}{|q_s|} \frac{1}{m_{H_s}} \right) , \tag{41}$$

$$\frac{F_{TA}(x_\gamma, m_{H_s}, \mu)}{f_{H_s}} = \frac{|q_s|}{x_\gamma} \left(\frac{R_T(E_\gamma, \mu)}{\lambda_B(\mu)} + \xi(x_\gamma, m_{H_s}) - \frac{1 - x_\gamma}{m_{H_s} x_\gamma} + \frac{|q_b|}{|q_s|} \frac{1}{m_{H_s}} \right) , \tag{42}$$

where $R_T(E_\gamma, \mu) = 1 + \mathcal{O}(\alpha_s)$ is the radiative correction. Again, perturbative radiative corrections to the subleading terms and $\mathcal{O}(m_s/(m_{H_s}^2))$ terms have been neglected. In Eqs. (41), (42) we have explicitly inserted in the l.h.s.
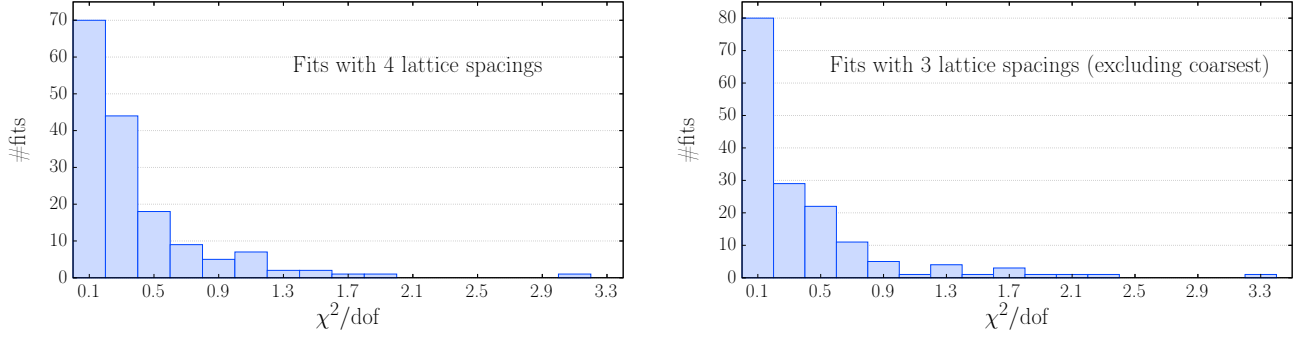
FIG. 9: *Histograms of the $\chi^2/$dof distribution corresponding to the 160 continuum limit extrapolations we have performed using the full dataset (left) and excluding the coarsest lattice spacing ensemble, the A48 (right).*

the dependence on the renormalization scale μ, which is instead absent in F_V and F_A which are scale-independent quantities. The previous relations imply that, neglecting power suppressed contributions and radiative corrections, one has $F_{TV}(x_\gamma) = F_{TA}(x_\gamma) = F_V(x_\gamma) = F_A(x_\gamma)$.

We now explain that the above asymptotic relations for the form factors, being valid in the limit of large E_γ, are not sufficient to describe their behaviour in the range of the simulated values of m_{H_s} and x_γ because of the presence of sizeable non-asymptotic contributions from resonances. To highlight this point, we start from the canonical decomposition of the form factors in terms of intermediate-state contributions. The hadronic tensor $H_W^{\mu\nu}(p,k)$ in Eq. (19)

$$H_W^{\mu\nu}(p,k,m_{H_s}) \equiv i \int d^4y \; e^{iky} \; \hat{\mathrm{T}} \langle 0| J_W^\nu(0) J_{\mathrm{em}}^\mu(y) |\bar{H}_s(\mathbf{0})\rangle \; , \tag{43}$$

can be decomposed as

$$H_W^{\mu\nu}(p,k,m_{H_s}) = i \int_{-\infty}^0 dt \, e^{iE_\gamma t} \langle 0| J_W^\nu(0) J_{\mathrm{em}}^\mu(t,\mathbf{k}) |\bar{H}_s(\mathbf{0})\rangle + i \int_0^\infty dt \, e^{iE_\gamma t} \, \langle 0| J_{\mathrm{em}}^\mu(t,\mathbf{k}) J_W^\nu(0) |\bar{H}_s(\mathbf{0})\rangle$$

$$\equiv H_{W,1}^{\mu\nu}(p,k,m_{H_s}) + H_{W,2}^{\mu\nu}(p,k,m_{H_s}) \; , \tag{44}$$

where

$$J_{\mathrm{em}}^\mu(t,\mathbf{k}) \equiv \int d^3x \, e^{-i\mathbf{k}\mathbf{x}} \, J_{\mathrm{em}}^\mu(t,\mathbf{x}) \; . \tag{45}$$

We now focus on the contribution from the first time-ordering, $H_{W,1}^{\mu\nu}$, which can be written as

$$H_{W,1}^{\mu\nu}(p,k,m_{H_s}) = i \int_{-\infty}^0 dt \, e^{iE_\gamma t} \langle 0| J_W^\nu(0) e^{i(\hat{H}-m_{H_s}-i\varepsilon)t} J_{\mathrm{em}}^\mu(0,\mathbf{k}) |\bar{H}_s(\mathbf{0})\rangle$$

$$= \langle 0| J_W^\nu(0) \frac{1}{\hat{H} + E_\gamma - m_{H_s} - i\varepsilon} J_{\mathrm{em}}^\mu(0,\mathbf{k}) |\bar{H}_s(\mathbf{0})\rangle = \sum_n \frac{\langle 0| J_W^\nu(0)|n\rangle \langle n| J_{\mathrm{em}}^\mu(0,\mathbf{k})|\bar{H}_s(\mathbf{0})\rangle}{2E_n(-\mathbf{k})(E_n(-\mathbf{k}) + E_\gamma - m_{H_s})} \; , \tag{46}$$

where $\hat{H}$ is the QCD Hamiltonian. The contributing intermediate states $|n\rangle$ are $B = -1$, $S = 1$ states with $J^P = 1^-$ for $W = \{V, TV\}$ and $J^P = 1^+$ for $W = \{A, TA\}$. Their energies are given by $E_n(-\mathbf{k}) = \sqrt{m_n^2 + E_\gamma^2}$. In the following, in order to model the x_γ and mass behaviour of the form factors, we only consider the contributions coming from the resonances that we treat as stable particles. Using the following relations (η_n is the polarization of the vector meson $|n\rangle$, $k_n = (E_n, -\mathbf{k})$, $p_\gamma = k_n - p$):

$$\langle 0| J_V^\nu|n\rangle = \eta_n^\nu \, m_n \, f_n^V \; , \qquad \langle 0| J_{TV}^\nu|n\rangle = f_n^{TV} \left[\frac{\eta_n^\nu}{m_{H_s}}(k_n \cdot q) - \frac{k_n^\nu}{m_{H_s}}(\eta_n \cdot q) \right]$$

$$\langle 0| J_A^\nu|n\rangle = i\eta_n^\nu \, m_n \, f_n^A \; , \qquad \langle 0| J_{TA}^\nu|n\rangle = -i f_n^{TA} \left[\frac{\eta_n^\nu}{m_{H_s}}(k_n \cdot q) - \frac{k_n^\nu}{m_{H_s}}(\eta_n \cdot q) \right]$$

$$\langle n| J_{\mathrm{em}}^\mu|\bar{H}_s(\mathbf{0})\rangle = g_n \, \varepsilon^{\mu\nu\gamma\beta}(\eta_n^*)_\nu \, k_{n,\gamma} \, p_\beta \qquad\qquad W = \{V, TV\} \; ,$$

$$\langle n| J_{\mathrm{em}}^\mu|\bar{H}_s(\mathbf{0})\rangle = g_n' \left[(\eta_n^*)^\mu(p \cdot p_\gamma) - (\eta_n^* \cdot p_\gamma)p^\mu \right] + \mathcal{O}(p_\gamma^2) \; , \qquad W = \{A, TA\} \; , \tag{47}$$

and given that at leading-order in m_{H_s} one has $f_n^W \propto f_{H_s}$, and assuming that at leading-order the form factors g_n and g_n' are constant, one obtains that each of the intermediate states in Eq. (46) gives a contribution $F_{W,n}(x_\gamma)$ to the form factor $F_W(x_\gamma)$ which scales as ($r_n \equiv m_n/m_{H_s}$)

$$F_{W,n}(x_\gamma) \propto \frac{f_{H_s}}{\sqrt{r_n^2 + \frac{x_\gamma^2}{4} + \frac{x_\gamma}{2} - 1}} \ . \tag{48}$$

In the static limit, since r_n approaches one, the scaling relations in Eqs. (38) - (42) are recovered. However, it is important to notice that for $x_\gamma = 0$ the denominator in Eq. (48) develops a pole for $r_n \to 1$, signalling the fact that the scaling laws are different at $x_\gamma = 0$ and $x_\gamma \neq 0$ (in this last case the denominator approaches a non-zero value $x_\gamma/2 + \mathcal{O}(x_\gamma^2)$ in the $m_h \to \infty$ limit). For small enough values of x_γ, the presence of a quasi-pole may generate large corrections to the scaling relations in Eqs. (38) - (42), which we now discuss. We start by recalling the following HQET relations for the masses $m_{\bar{H}_s^*}$ and $m_{H_{s1}}$ of the lowest-lying vector mesons $\bar{H}_s^*$ and $\bar{H}_{s1}$ in the $J^P = 1^-$ and $J^P = 1^+$ channel [34]

$$m_{\bar{H}_s^*}^2 - m_{\bar{H}_s}^2 = 2\lambda_2 + \mathcal{O}\left(\frac{1}{m_h}\right) , \qquad \lambda_2 \simeq 0.24 \text{ GeV}^2 , \tag{49}$$

$$m_{\bar{H}_{s1}} - m_{\bar{H}_s} = \Lambda_1 + \mathcal{O}\left(\frac{1}{m_h}\right) , \qquad \Lambda_1 \simeq 0.5 \text{ GeV} . \tag{50}$$

As is well known, the first relation comes from the fact that the ground-state pseudoscalar ($J^P = 0^-$) and vector ($J^P = 1^-$) mesons, are members of the same HQET spin-doublet, and so they become degenerate in the infinite heavy-quark mass limit. The mass-splitting between the ground-state pseudoscalar and axial-vector meson is instead of order $\mathcal{O}(\Lambda_{\text{QCD}})$. This implies that for the lowest-lying intermediate state ($n = 1$) contributing to F_W, for $W = \{V, TV\}$, one has

$$r_1 = \frac{m_{H_s^*}}{m_{H_s}} \simeq 1 + \frac{\lambda_2}{m_{H_s}^2} \qquad \Longrightarrow \qquad \sqrt{r_1^2 + \frac{x_\gamma^2}{4} + \frac{x_\gamma}{2} - 1} \simeq \frac{\lambda_2}{m_{H_s}^2} + \frac{x_\gamma}{2} + \dots , \tag{51}$$

where the ellipses indicate subleading corrections at large m_h and small x_γ. For $W = \{A, TA\}$, one has instead

$$r_1 = \frac{m_{H_{s1}}}{m_{H_s}} \simeq 1 + \frac{\Lambda_1}{m_{H_s}} \qquad \Longrightarrow \qquad \sqrt{r_1^2 + \frac{x_\gamma^2}{4} + \frac{x_\gamma}{2} - 1} \simeq \frac{\Lambda_1}{m_{H_s}} + \frac{x_\gamma}{2} + \dots . \tag{52}$$

The previous equations show that for small values of x_γ the quasi-pole produces an *enhancement* of the form factor of order $\mathcal{O}(m_h^2)$ and $\mathcal{O}(m_h)$, respectively in the vector-like and axial-like channels. In the following section we will combine the leading-order relations Eqs. (38)- (42), with the quasi-pole behaviour described by Eqs. (51), (52) in order to extrapolate the form factors to the physical B_s meson mass.

C. Numerical results at the physical B_s mass

Guided by the analysis in the previous section, we introduce some model-dependent interpolating formulae for the form factors which describe their q^2 dependence in the resonance region and have the correct asymptotic behaviour in the limit of large m_h. We have extrapolated our results for F_W, $W = \{V, A, TV, TA\}$, obtained at the five different simulated values of the heavy quark mass m_h in Eq. (32), employing the following fit Ansatz ($z = 1/m_{H_s}$)

$$\frac{F_V(x_\gamma, z)}{f_{H_s}} = \frac{|q_s|}{x_\gamma} \frac{1}{1 + C_V \frac{2z^2}{x_\gamma}} \left(K + (1 + \delta_z)\frac{z}{x_\gamma} + \frac{1}{z^{-1} - \Lambda_H} + A_m z + A_{x_\gamma}\frac{z}{x_\gamma} + B_m^V z^2 + B_{x_\gamma}^V \frac{z^2}{x_\gamma} \right) , \tag{53}$$

$$\frac{F_A(x_\gamma, z)}{f_{H_s}} = \frac{|q_s|}{x_\gamma} \frac{1}{1 + C_A \frac{2z}{x_\gamma}} \left(K - (1 + \delta_z)\frac{z}{x_\gamma} - \frac{1}{z^{-1} - \Lambda_H} + A_m z + (A_{x_\gamma} + 2KC_A)\frac{z}{x_\gamma} + B_m^A z^2 + B_{x_\gamma}^A \frac{z^2}{x_\gamma} \right) , \tag{54}$$

$$\frac{F_{TV}(x_\gamma, z)}{f_{H_s}} = \frac{|q_s|}{x_\gamma} \frac{1 + 2C_V z^2}{1 + C_V \frac{2z^2}{x_\gamma}} \left(K_T + (A_m^T + 1)z + A_{x_\gamma}^T \frac{z}{x_\gamma} + (1 + \delta_z')z\frac{1 - x_\gamma}{x_\gamma} + B_m^T z^2 + B_{x_\gamma}^{TV}(1 - x_\gamma)\frac{z^2}{x_\gamma} \right) , \tag{55}$$

$$\frac{F_{TA}(x_\gamma, z)}{f_{H_s}} = \frac{|q_s|}{x_\gamma} \frac{1 + 2C_A^T z}{1 + C_A^T \frac{2z}{x_\gamma}} \left(K_T + (A_m^T + 1)z + A_{x_\gamma}^T \frac{z}{x_\gamma} - (1 + \delta_z' - 2K_T C_A^T)z\frac{1 - x_\gamma}{x_\gamma} + B_m^T z^2 + B_{x_\gamma}^{TA}(1 - x_\gamma)\frac{z^2}{x_\gamma} \right) , \tag{56}$$

where $C_A^T = C_A + \delta C_A^T$, $K_T = K + \delta K_T$, and $K, \delta K_T, \delta_z, \Lambda_H, C_V, C_A, \delta C_A^T, A_{m/x_\gamma}, A_m^T, A_{x_\gamma}^T, \delta_z'$ and B_{m/x_γ}^W ($W = \{V, A, TV, TA\}$) are free fit parameters. Our strategy to extrapolate the form factors to the physical B_s-meson consists in a simultaneous global fit of the mass and x_γ dependence of all four form factors. The phenomenological fit Ansatz described by Eqs. (53)-(56) takes into account the constraints discussed in the previous section, and contains the quasi-pole corrections to the asymptotic scaling described by Eqs. (38), (39), (41) and (42). We however relaxed the constraint $A_m^T = A_m$ and $A_{x_\gamma} = A_{x_\gamma}^T$ due to the presence of the same function $\xi(x_\gamma, m_{H_s})$ in the expression for the tensor-like and vector-like form factors in Eqs (38)-(42). The position of the pole is taken to be the same in the vector and tensor-vector channel, while we allow for the possibility of having a different pole in F_A and F_{TA} (i.e. $\delta C_A^T \neq 0$). This is due to the fact that while in the vector channel vector-meson-dominance is expected to work well since the vector (H_s^*) and pseudoscalar (H_s) ground-state mesons become degenerate in the static limit, this is not the case in the axial channel where many resonances with masses of order $m_{H_s} + \mathcal{O}(\Lambda_{\mathrm{QCD}})$ are present. The axial pole should be considered to be an *effective pole*, and its position can therefore be slightly different in the axial and tensor-axial channel due to the different couplings to the excited states. Moreover, in order to account for the fact that the tree-level equality between tensor-like and vector-like form factors is spoiled by the radiative corrections, we also allow for the possibility that $\delta K_T \neq 0$. Notice that in the tensor form factors, the numerator in the pole term is inserted to ensure the validity of the kinematical constraint $F_{TA}(1) = F_{TV}(1)$. Finally, we have included a parameter Λ_H to account for the fact that the hadron mass m_{H_s} differs from the heavy quark mass m_h by an amount of order $\mathcal{O}(\Lambda_{\mathrm{QCD}})$ (see the last terms in Eqs. (38) and (39)), and two parameter δ_z, δ_z' to account for violations of the relations in Eqs. (38)-(42) which are only exactly valid in the limit of a massless strange quark and neglecting radiative corrections to the power suppressed terms. Our determination of the decay constant f_{H_s}, on the same configurations used for the computation of the form factors, is discussed in detail in Appendix A. In the same appendix we also discuss our determination of f_{B_s}, for which we get the value

$$f_{B_s} = 224.5(5.0)\,\mathrm{MeV}\,. \tag{57}$$

Our determination of f_{B_s} agrees with the $N_f = 2 + 1 + 1$ FLAG average $f_{B_s}^{\mathrm{FLAG}} = 230.3(1.3)\,\mathrm{MeV}$ at the level of 1.1σ, although our uncertainty is larger. Using the Ansätze in Eqs. (53)-(56) we have performed a total of $N \simeq \mathcal{O}(500)$ fits which differ on whether the fit parameters δK_T, δC_A^T, δ_z, δ_z' are set to zero or not, and on whether we include or not the fit parameters describing the $\mathcal{O}(1/m_{H_s}^2)$ corrections. The total number of measurements is 80 and the maximum number of fit parameters used is 14. To stabilize the fits, large Gaussian priors are imposed on the fit parameters $\delta K_T, \delta C_A^T, \delta_z, \delta_z', \Lambda_H$ and B_{m/x_γ}^W. These are

$$\delta K_T = 0 \pm 0.4\,\mathrm{GeV}^{-1}\,, \quad \delta_z^{(')} = 0 \pm 1\,, \quad \delta C_A^T = 0 \pm 0.4\,\mathrm{GeV}\,, \quad \Lambda_H = 0.75 \pm 0.5\,\mathrm{GeV}\,, \quad B_{m/x_\gamma}^W = 0 \pm 2.5\,\mathrm{GeV}\,. \tag{58}$$

We minimize a correlated χ^2 function which takes fully into account the correlations between the values of a given form factor at the different simulated values of x_γ and m_{H_s}. However, in order to avoid having an ill-conditioned covariance matrix, we assume, in the construction of the χ^2, that the different form factors are instead uncorrelated. The error on the fit parameters are always properly estimated, since they are obtained from the dispersion of the results obtained repeating the fits for each jackknife sample.

Many of the fit parameters entering Eqs. (53)-(55) are not needed in order to obtain a good χ^2/dof, and a good description of the data is already obtained by setting $\delta K_T = \delta C_A^T = \delta_z = \delta_z' = 0$ and neglecting the $\mathcal{O}(1/m_{H_s}^2)$ corrections. However, in order to estimate correctly the systematic errors due to the mass extrapolation, it is important to span over a sufficiently large number of fit Ansätze.

We combine the results of the N different fits using two different criteria. The first one is based on the AIC discussed in Section III A, i.e. we assign to each of the N fits a weight w_i given by

$$w_i \propto \exp\!\left(-\left(\chi_{(i)}^2 + 2N_{\mathrm{pars}}^{(i)} - N_{\mathrm{data}}^{(i)}\right)/2\right)\,, \qquad \sum_{i=1}^N w_i = 1\,, \tag{59}$$

where $\chi_{(i)}^2$ is the total χ^2 of the i-th fit, and $N_{\mathrm{pars}}^{(i)}$ and $N_{\mathrm{meas}}^{(i)}$ are the corresponding number of fit parameters and measurements. The second criterion consists in selecting only those fits leading to a good χ^2/dof and assigning them an uniform weight, i.e. using

$$w_i \propto \theta\!\left(c - \chi_{(i)}^2/N_{\mathrm{dof}}^{(i)}\right)\,, \qquad \sum_{i=1}^N w_i = 1\,, \tag{60}$$

and we set $c = 1.4$ which corresponds approximately to $1 + 2\sqrt{2/N_{\mathrm{dof}}^{(i)}}$, where $\sqrt{2/N_{\mathrm{dof}}^{(i)}}$ is the standard deviation of the reduced χ^2 distribution with $N_{\mathrm{dof}}^{(i)}$ degrees of freedom. Then, with a given choice for the weights w_i, the final central
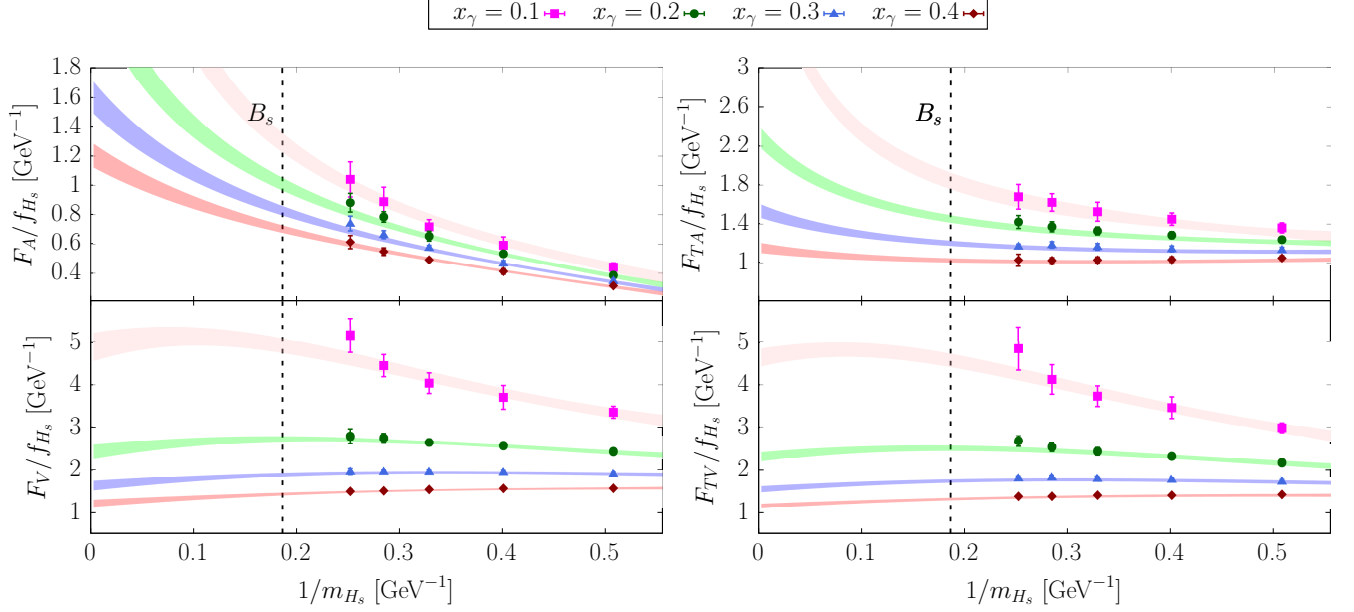
FIG. 10: *Extrapolation to the physical B_s meson of the four form factors F_A (top left) , F_{TA} (top right), F_V (bottom left) and F_{TV} (bottom right). The form factors are divided by the decay constant f_{H_s} of the heavy-strange pseudoscalar meson. The different colors correspond to the different simulated values of the Lorentz invariant x_γ. Finally, the continuum bands correspond to the best-fit function obtained after applying the model-averaged procedure described by Eqs. (61)-(62) using the weights in Eq. (60).*

value $\bar{x}$ is obtained from a weighted average:

$$\bar{x} = \sum_{i=1}^{N} w_i x_i \ , \tag{61}$$

where x_i is the result obtained from the i-th fit. The sum in Eq. (61) is evaluated in a correlated way, so that the statistical errors of the x_i are correctly propagated to $\bar{x}$. The systematic error, which is added in quadrature to the statistical error of $\bar{x}$, is then given by

$$\sigma_{x,\text{syst}}^2 = \sum_{i=1}^{N} w_i \left(x_i - \bar{x}\right)^2 \ . \tag{62}$$

We have found that the results obtained using the weights in Eq. (59) and Eq. (60) are consistent well within the uncertainties. However, at small x_γ the errors obtained using the AIC are typically smaller than those obtained using Eq. (60). In order to be conservative, we take the results obtained using the weights in Eq. (60) to obtain our final results for the form factors.

The results of the extrapolation are collected in the plots of Figure 10. The continuum bands in the figure correspond to the best-fit function obtained after applying the above procedure with the weights in Eq. (60). We obtain for the pole coefficients C_V, C_A and $C_A + \delta C_A^T$ the values

$$C_V = (0.57(3) \text{ GeV})^2 \ , \qquad C_A = 0.70(7) \text{ GeV} \ , \qquad C_A + \delta C_A^T = 0.77(4) \text{ GeV} \ . \tag{63}$$

The result for C_V can be compared with the value expected from the HQET relation in Eq. (49), namely $C_V \simeq \lambda_2 \simeq (0.5 \text{ GeV})^2$. Although slightly larger (recall that the Ansatz we use is a phenomenological description of the full form factors where excited-states contributions are always present), our determination is in line with expectations, and provides nice evidence that the reason behind the steep rise of the vector form factors at small x_γ is due to the presence of the quasi-pole. Concerning the position of the axial pole, the value we obtained for C_A and δC_A^T is also qualitatively in line with the expectations $C_A, C_A + \delta C_A^T \simeq \mathcal{O}(\Lambda_1)$. We did not find clear evidences of non-zero values of δ_z, δ_z' and δK_T. We obtain

$$1 + \delta_z = 1.02(9) \ , \qquad 1 + \delta_z' = 1.06(8) \ , \qquad K = 1.46(10) \text{ GeV}^{-1} \ , \qquad K + \delta K_T = 1.39(6) \text{ GeV}^{-1} \ , \tag{64}$$

x_γ	F_V	F_A	F_{TV}	F_{TA}
0.1	1.103(38)	0.290(13)	1.026(35)	0.413(17)
0.2	0.610(13)	0.226(8)	0.564(15)	0.326(8)
0.3	0.422(8)	0.186(6)	0.389(10)	0.270(6)
0.4	0.322(6)	0.157(5)	0.297(8)	0.230(5)

TABLE III: *Our results for the form factors F_V, F_A, F_{TV} and F_{TA} extrapolated to the physical mass m_{B_s}, for the four simulated values of $x_\gamma = 0.1, 0.2, 0.3, 0.4$.*
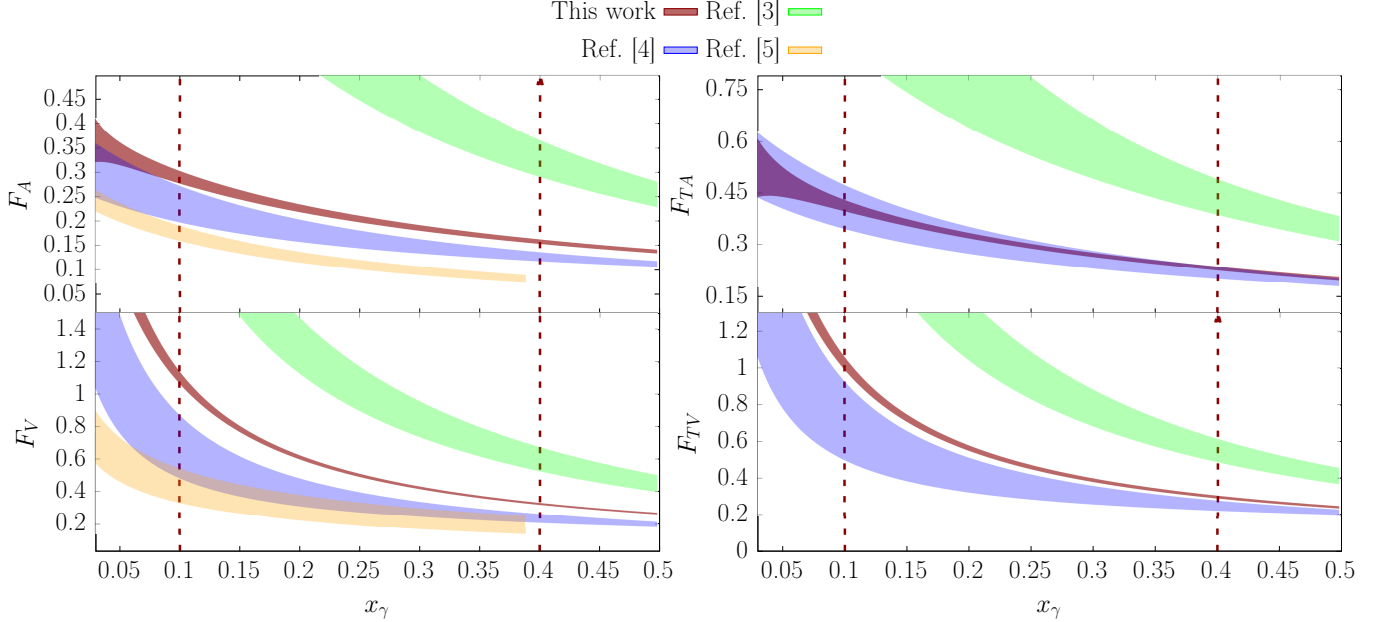

FIG. 11: *Comparison between our results for the form factors F_W with $W = \{V, A, TV, TA\}$ (shown in the figure by the red bands), and existing model-dependent results [3–5]. The region between the vertical red dashed lines corresponds to the region of simulated x_γ, and therefore within this region our results are obtained through an interpolation of our lattice data.*

and for the fit parameter Λ_H we obtain the value $\Lambda_H = 0.70(17)$ GeV. Finally, for the parameters A_m, A_m^T, A_{x_γ} and $A_{x_\gamma}^T$ we obtain

$$A_m = 0.8(5), \qquad A_m^T = 1.4(3), \qquad A_{x_\gamma} = -1.0(1), \qquad A_{x_\gamma}^T = -1.0(1). \tag{65}$$

The relation $A_{x_\gamma} = A_{x_\gamma}^T$ which holds in the HQET and large-photon-energy expansion neglecting perturbative radiative corrections and non-zero strange-quark mass effects, appears to be well reproduced by our data. As for the relation $A_m = A_m^T$, we find that the fitted values of A_m and A_m^T are slightly different, which can be attributed to radiative corrections and/or $\mathcal{O}(m_s/m_{H_s})$ effects as well as to statistical fluctuations. In Table III we provide our results for the four form factors, extrapolated to the physical mass m_{B_s} and for the four simulated values of x_γ. The fit parameters, including their correlations, are available upon request from the authors.

Our results for the form factors can be compared with available phenomenological and model estimates. The form factors F_W, $W = \{V, A, TV, TA\}$, have been previously obtained using relativistic dispersion relations [4], light-cone sum rules [3] and recently a hybrid approach [5] in which the existing lattice results for the form factors F_V and F_A in $D_s \to \ell \nu \gamma$ decays are used to obtain the form factors F_V and F_A entering $B_s \to \mu^+ \mu^- \gamma$ decays using a VMD-inspired ansatz. The comparison between our determination of F_W, $W = \{V, A, TV, TA\}$ and the existing model-dependent results is shown in Figure 11. Our results are given by the red curves, and outside the region of measured x_γ are obtained by using the best-fit function obtained in the global fits discussed above. Our results for F_{TV} and F_{TA} turn out to be in rather good agreement with the estimate of Ref. [4], taking into account that the results of the relativistic dispersion approach contain a systematic uncertainty which is difficult to quantify. However we find significant

differences with respect to the results of Ref. [3] for F_{TV} and F_{TA}. For the axial form factor F_A, the differences between our results and those of Ref. [4] are of similar size as the one present for F_{TV}, while more significant deviations are observed for the vector form factor F_V. Moreover, we disagree with both the estimates given in Ref. [3] and Ref. [5] for F_V and F_A. The disagreement with the light-cone sum rule calculation was somehow expected, given that large differences with respect to lattice QCD calculations have been already observed in the radiative leptonic decays of the D_s meson [7]. The smaller value of F_V obtained in Ref. [5] could be, at least partially, traced back to the fact that their estimate of the strange-quark contribution to the form factor $g_{D_s^* D_s \gamma}$, an essential input parameter of their VMD-inspired approach, turns out to be substantially smaller than the one obtained by the HPQCD Collaboration in Ref. [35] and in our recent paper [7] (which are instead in very nice agreement with each other). Before discussing the implications of our results for the branching fraction $\mathcal{B}(B_s \to \mu^+ \mu^- \gamma)$, we present now our results for the form factor $\bar{F}_T$.

IV. THE LOCAL FORM FACTOR $\bar{\mathbf{F}}_{\mathbf{T}}$

The form factor $\bar{F}_T$ can be computed from the knowledge of the hadronic tensor

$$H_{\bar{T}}^{\mu\nu}(p,k) = i \int d^4x \; e^{i(p-k)x} \; \hat{\mathrm{T}} \langle 0 | J_{\bar{T}}^{\nu}(0) J_{\mathrm{em}}^{\mu}(x) | \bar{B}_s(\mathbf{0}) \rangle = -\varepsilon^{\mu\nu\rho\sigma} k_\rho p_\sigma \frac{\bar{F}_T}{m_{B_s}} \;, \tag{66}$$

where

$$J_{\bar{T}}^{\nu} = -i Z_T(\mu) \bar{s} \, \sigma^{\nu\rho} \, b \frac{k_\rho}{m_{B_s}} \;. \tag{67}$$

As in the case of the currents J_{TV}^{ν} and J_{TA}^{ν}, we renormalize the tensor current $J_{\bar{T}}$ using the non-perturbative determination of $Z_T(\mu)$ in the $\overline{\mathrm{MS}}$ scheme at the scale $\mu = 5$ GeV given in Table II. Note that $H_{\bar{T}}^{\mu\nu}(p,k) = H_{TV}^{\mu\nu}(p, p-k)$, and recall that $\bar{F}_T(1) = F_{TV}(1) = F_{TA}(1)$ (see Eq. (14)). A significant complication is that the hadronic tensor $H_{\bar{T}}^{\mu\nu}(p,k)$ suffers from problems of analytic continuation to Euclidean spacetime. To demonstrate this, we start by writing explicitly the contributions to the hadronic tensor $H_{\bar{T}}^{\mu\nu}$ from the two time orderings, namely

$$H_{\bar{T}}^{\mu\nu}(p,k) = i \int_{-\infty}^{0} dt \, e^{i(m_{B_s} - E_\gamma)t} \, \langle 0 | J_{\bar{T}}^{\nu}(0) \, J_{\mathrm{em}}^{\mu}(t, -\mathbf{k}) | \bar{B}_s(\mathbf{0}) \rangle + i \int_{0}^{\infty} dt \, e^{i(m_{B_s} - E_\gamma)t} \, \langle 0 | J_{\mathrm{em}}^{\mu}(t, -\mathbf{k}) J_{\bar{T}}^{\nu}(0) | \bar{B}_s(\mathbf{0}) \rangle \;. \tag{68}$$

Making use of

$$J_{\mathrm{em}}^{\mu}(t, \mathbf{k}) = e^{i(\hat{H} - i\varepsilon)t} \, J_{\mathrm{em}}^{\mu}(0, \mathbf{k}) \, e^{-i(\hat{H} - i\varepsilon)t} \;, \tag{69}$$

where $\hat{H}$ is the QCD Hamiltonian, one has $H_{\bar{T}}^{\mu\nu}(p,k) = H_{\bar{T},1}^{\mu\nu}(p,k) + H_{\bar{T},2}^{\mu\nu}(p,k)$ with

$$H_{\bar{T},1}^{\mu\nu}(p,k) = \langle 0 | J_{\bar{T}}^{\nu}(0) \frac{1}{\hat{H} - E_\gamma - i\varepsilon} J_{\mathrm{em}}^{\mu}(0, -\mathbf{k}) | \bar{B}_s(0) \rangle \;,$$

$$H_{\bar{T},2}^{\mu\nu}(p,k) = \langle 0 | J_{\mathrm{em}}^{\mu}(0, -\mathbf{k}) \frac{1}{\hat{H} + E_\gamma - m_{B_s} - i\varepsilon} J_{\bar{T}}^{\nu}(0) | \bar{B}_s(0) \rangle \;. \tag{70}$$

The two integrals in Eq. (68) can only be Wick-rotated from Minkowskian time t to Euclidean time $\tau = it$ if the following positivity conditions are met

$$\langle n | \hat{H} - E_\gamma | n \rangle > 0, \qquad \langle m | \hat{H} + E_\gamma - m_{B_s} | m \rangle > 0 \;, \tag{71}$$

where $|n\rangle$ and $|m\rangle$ are the intermediate states contributing respectively to the first and second time ordering in Eq. (70). In the rest frame of the $\bar{B}_s$ meson in which we work, all intermediate states contributing to the hadronic tensor have three momentum $|\mathbf{k}| = E_\gamma$, therefore the condition $\langle n | \hat{H} - E_\gamma | n \rangle > 0$ is always satisfied and one can safely set $\varepsilon = 0$ in the first contribution on the r.h.s. of Eq. (70). This is not the case for the second condition in Eq. (71) due to the presence of light unflavoured $J^P = 1^-$ intermediate states. Indeed, defining m_{V_0} to be the mass of the lightest hadronic state contributing to the second time ordering, the analytic continuation is obstructed if the photon energy E_γ satisfies

$$\sqrt{m_{V_0}^2 + E_\gamma^2} + E_\gamma < m_{B_s} \implies x_\gamma < 1 - \left(\frac{m_{V_0}}{m_{B_s}} \right)^2 \;. \tag{72}$$

As in the case of the local form factors F_W, $W = \{V, A, TV, TA\}$, we can distinguish the two contributions $\bar{F}_T^b$ and $\bar{F}_T^s$ to the form factor $\bar{F}_T$, corresponding respectively to the emission of the virtual photon γ^* from the bottom (Figure 3 left) and strange (Figure 3 right) quark line [6]. The lightest hadronic intermediate state in the second time ordering are given, respectively for the bottom- and strange-quark contributions, by the $\Upsilon(1S)$ resonance and by $K^+ K^-$ states in a P-wave [7]. Given that $m_\Upsilon \simeq 9460$ MeV $> m_{B_s}$, the bottom quark contribution is not affected by the problem of analytic continuation, which is only present in the strange quark contribution. Indeed, with $2m_K \simeq 1$ GeV, one finds that analytic continuation is obstructed for

$$x_\gamma < x_\gamma^{\text{th}} \equiv 1 - \left(\frac{2m_K}{m_{B_s}}\right)^2 \simeq 0.96 \;, \tag{73}$$

i.e. for all the values of x_γ that we are considering.

Recently, some of us have proposed a novel strategy [8] to circumvent the problem of analytic continuation of electroweak amplitudes of the type present in Eq. (68), i.e. involving an hadron-to-vacuum QCD matrix element of the product of two currents. In order to briefly summarise the strategy, we focus on the strange-quark contribution to $H_{T,2}^{\mu\nu}(p,k)$. To keep the notation simple, we set $E \equiv m_{B_s} - E_\gamma$ and define

$$C_{q,2}^{\mu\nu}(t,\boldsymbol{k}) \equiv \langle 0 | J_q^\mu(t, -\boldsymbol{k}) J_{\bar{T}}^\nu(0) | \bar{B}_s(\boldsymbol{0}) \rangle \;, \qquad H_{\bar{T}_q,2}^{\mu\nu}(E,\boldsymbol{k}) \equiv i \int_0^\infty dt \, e^{iEt} \, C_{q,2}^{\mu\nu}(t,\boldsymbol{k}) \;, \qquad q = s, b \;, \tag{74}$$

so that $H_{\bar{T},2}^{\mu\nu}(p,k) = H_{\bar{T}_b,2}^{\mu\nu}(E,\boldsymbol{k}) + H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k})$. The main idea for circumventing the problem of the analytic continuation of $H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k})$, is to consider the spectral-density representation of the time-dependent correlation function $C_{s,2}^{\mu\nu}(t,\boldsymbol{k})$,

$$C_{s,2}^{\mu\nu}(t,\boldsymbol{k}) = \int_{E^*}^\infty \frac{dE'}{2\pi} e^{-iE't} \rho^{\mu\nu}(E',\boldsymbol{k}) \;, \tag{75}$$

where $[E^*, \infty)$ is the support of the spectral density $\rho^{\mu\nu}(E', \boldsymbol{k})$, and in our case $E^* = \sqrt{m_{V_0}^2 + E_\gamma^2}$ with $m_{V_0} = 2m_K$. Combining Eqs. (74) and (75) it follows that (see Ref. [8] for details)

$$H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k}) = \lim_{\varepsilon \to 0^+} \int_{E^*}^\infty \frac{dE'}{2\pi} \frac{\rho^{\mu\nu}(E',\boldsymbol{k})}{E' - E - i\varepsilon} = \text{PV} \int_{E^*}^\infty \frac{dE'}{2\pi} \frac{\rho^{\mu\nu}(E',\boldsymbol{k})}{E' - E} + \frac{i}{2}\rho^{\mu\nu}(E,\boldsymbol{k}) \;, \tag{76}$$

where PV denotes the principal value of the integral. The Minkowski correlator $C_{q,2}^{\mu\nu}(t,\boldsymbol{k})$ can always be analytically continued to Euclidean spacetime. The Euclidean correlator $\bar{C}_{s,2}^{\mu\nu}(t,\boldsymbol{k}) \equiv C_{s,2}^{\mu\nu}(-it,\boldsymbol{k})$ is then related to the spectral density $\rho^{\mu\nu}(E', \boldsymbol{k})$ via

$$\bar{C}_{s,2}^{\mu\nu}(t,\boldsymbol{k}) = \int_{E^*}^\infty \frac{dE'}{2\pi} e^{-E't} \rho^{\mu\nu}(E',\boldsymbol{k}) \;. \tag{77}$$

Since $\bar{C}_{s,2}^{\mu\nu}(t,\boldsymbol{k})$ can be computed using Monte Carlo simulations, we have formally solved the problem of analytic continuation; by inverting the relation in Eq. (77) to determine $\rho^{\mu\nu}(E')$ we can then obtain $H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k})$ using Eq. (76). However, in order to determine $\rho^{\mu\nu}(E', \boldsymbol{k})$ using Eq. (77), an inverse Laplace transform of the Euclidean correlator $\bar{C}_{s,2}^{\mu\nu}(t,\boldsymbol{k})$ is required. This is a well-known ill-posed numerical problem when $\bar{C}_{s,2}^{\mu\nu}(t,\boldsymbol{k})$ is only known on a finite set of points in time and is affected by uncertainties, which is the typical situation encountered in a lattice calculation. In Ref. [8] it has been proposed to use the $-i\varepsilon$ term appearing in the denominator of Eq. (76) as a *regulator* of the problem by introducing the smeared amplitude $H_{s,2}^{\mu\nu}(E, \boldsymbol{k}; \varepsilon)$

$$H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k};\varepsilon) = \int_{E^*}^\infty \frac{dE'}{2\pi} \frac{\rho^{\mu\nu}(E',\boldsymbol{k})}{E' - E - i\varepsilon} = \int_{E^*}^\infty \frac{dE'}{2\pi} K(E' - E; \varepsilon) \, \rho^{\mu\nu}(E',\boldsymbol{k}) \;, \tag{78}$$

where

$$K(x;\varepsilon) \equiv \frac{1}{x - i\varepsilon} = \frac{x}{x^2 + \varepsilon^2} + i\frac{\varepsilon}{x^2 + \varepsilon^2} \;. \tag{79}$$

----

[6] The two contributions are obtained by the replacements $J_{\text{em}}^\mu \to J_b^\mu$ and $J_{\text{em}}^\mu \to J_s^\mu$ in Eqs. (68)-(70) (see Eq. (22)).
[7] In the electroquenched approximation in which we work the Zweig-suppressed contributions from $u\bar{u}$, $d\bar{d}$ and $c\bar{c}$ resonances are absent.

The key point is that for non-zero values of ε, the kernel function $K(x;\varepsilon)$ is smooth, and its convolution integral with the spectral density $\rho^{\mu\nu}(E',\boldsymbol{k})$ can be evaluated, from the knowledge of $\bar{C}_{s,2}^{\mu\nu}(t,\boldsymbol{k})$ only, using the Hansen-Lupo-Tantalo (HLT) method introduced in Ref. [36] (see also Refs. [37–39] for recent applications of the method). The idea is to numerically evaluate the smeared amplitude $H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k};\varepsilon)$ for finite values of the smearing parameter ε using the HLT method (to be discussed in the next section), and then to extrapolate to $\varepsilon = 0$, exploiting the fact that (see Ref. [8] for a proof)

$$H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k};\varepsilon) = H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k}) + A(E,\boldsymbol{k})\varepsilon + \mathcal{O}(\varepsilon^2)\,. \tag{80}$$

We stress that the problem of evaluating $H_{\bar{T}}^{\mu\nu}(E,\boldsymbol{k})$ is ill-posed only for $E > E^*$, i.e. if the inequality in Eq. (72) is satisfied. Instead, for $E < E^*$, one can directly set $\varepsilon = 0$ in Eq. (76) (in this case the integrand is non-singular), and by using

$$\frac{1}{E' - E} = \int_0^\infty dt\, e^{-(E'-E)t} \qquad (\text{valid for } E < E')\,, \tag{81}$$

one arrives at (see Ref. [8])

$$H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k}) = \int_0^\infty dt\, e^{Et} \bar{C}_{s,2}^{\mu\nu}(t,\boldsymbol{k}) < \infty\,, \tag{82}$$

which is the standard formula used to evaluate the form factors in absence of problems of analytic continuation (see e.g. Eq. (21)) and the one we apply here to determine $H_{\bar{T}_b,2}^{\mu\nu}(E,\boldsymbol{k})$.

To summarize, we evaluate the hadronic tensor $H_{\bar{T}}^{\mu\nu}(p,k)$ in Eq. (66) as the sum of the following terms

$$H_{\bar{T}}^{\mu\nu}(p,k) \equiv H_{\bar{T}_b}^{\mu\nu}(E,\boldsymbol{k}) + H_{\bar{T}_s}^{\mu\nu}(E,\boldsymbol{k})\,, \tag{83}$$

where

$$H_{\bar{T}_b}^{\mu\nu}(E,\boldsymbol{k}) \equiv H_{\bar{T}_b,1}^{\mu\nu}(E,\boldsymbol{k}) + H_{\bar{T}_b,2}^{\mu\nu}(E,\boldsymbol{k})\,, \tag{84}$$

$$H_{\bar{T}_s}^{\mu\nu}(E,\boldsymbol{k}) \equiv \lim_{\varepsilon \to 0^+} H_{\bar{T}_s,1}^{\mu\nu}(E,\boldsymbol{k}) + H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k};\varepsilon)\,, \tag{85}$$

and we have defined the first time ordering contribution as $(q = b, s)$

$$H_{\bar{T}_q,1}^{\mu\nu}(E,\boldsymbol{k}) \equiv \int_{-\infty}^0 dt\, e^{Et} \bar{C}_{q,1}^{\mu\nu}(t,\boldsymbol{k})\,, \qquad \bar{C}_{q,1}^{\mu\nu}(t,\boldsymbol{k}) \equiv C_{q,1}^{\mu\nu}(-it,\boldsymbol{k})\,, \qquad C_{q,1}^{\mu\nu}(t,\boldsymbol{k}) \equiv \langle 0|J_{\bar{T}}^\nu(0)J_q^\mu(t,-\boldsymbol{k})|\bar{B}_s(\boldsymbol{0})\rangle\,, \tag{86}$$

where the Euclidean correlator $\bar{C}_{q,1}^{\mu\nu}(t,\boldsymbol{k})$ is the lattice input. On the lattice, because of the discretization of spacetime, the relations above get slightly modified, as we will discuss in the next section.

A. Numerical results for $\bar{F}_T$

In order to evaluate the form factor $\bar{F}_T$ we have performed simulations on a subset of the ensembles in Table I. These are the B64 and D96 ensembles. The computations have been performed at all four values of x_γ in Eq. (33) but only at the following three values of the heavy quark mass,

$$\frac{m_h}{m_c} \simeq 1,\ 1.5,\ 2.5\,. \tag{87}$$

The reasons for reducing the number of ensembles and values of m_h which we use are two-fold. Firstly, the use of the spectral representation technique to overcome the difficulty in the continuation to Euclidean space is computationally expensive and secondly the contribution to the differential rates from $\bar{F}_T$ is small and so this form factor is not required with the same precision as those studied in Section III.

Our strategy to compute $\bar{F}_T$ consists in evaluating on the lattice the following three-point Euclidean correlation function

$$M_{\bar{T}_q}^{\mu\nu}(t,t_{\text{sep}},\boldsymbol{k}) \equiv a^3 \sum_{\boldsymbol{x}} \langle 0|J_q^\mu(t+t_{\text{sep}},-\boldsymbol{k})\, J_{\bar{T}}^\nu(t_{\text{sep}})\, \phi_{B_s}^\dagger(0,\boldsymbol{x})|0\rangle \qquad q = s, b\,, \tag{88}$$

where $\phi^\dagger_{B_s}(0,\boldsymbol{x})$ is the same interpolating operator as was used in Eq. (21), while $t_{\rm sep}$ is the fixed time where the tensor FCNC $J^\nu_{\bar T}$ is inserted, which must be chosen large enough to ensure the dominance of the ground state. In the limit of large $t_{\rm sep}$ one has

$$M^{\mu\nu}_{\bar T_q}(t, t_{\rm sep}, \boldsymbol{k}) = \frac{\langle \bar B_s(0)|\phi^\dagger_{B_s}(0)|0\rangle}{2m_{B_s}} e^{-m_{B_s} t_{\rm sep}} \left(\bar C^{\mu\nu}_q(t) + \dots\right) , \qquad q = s, b , \tag{89}$$

where the dots represent terms that are exponentially suppressed at large $t_{\rm sep}$, and

$$\bar C^{\mu\nu}_q(t, \boldsymbol{k}) = \theta(-t)\, \bar C^{\mu\nu}_{q,1}(t, \boldsymbol{k}) + \theta(t)\, \bar C^{\mu\nu}_{q,2}(t, \boldsymbol{k}) , \qquad q = s, b, \tag{90}$$

where the correlators $\bar C^{\mu\nu}_{q,1}(t, \boldsymbol{k})$ and $\bar C^{\mu\nu}_{q,2}(t, \boldsymbol{k})$ were introduced in the previous section.

Notice that the time t in the previous equations corresponds to the time separation between the electromagnetic and tensor currents and is different from the time t introduced in Eq. (21). The choice of $t_{\rm sep}$ has been adapted depending on the contribution being considered. For $q = b$ (and both $t > 0$ and $t < 0$) and for $q = s$, $t < 0$ we have chosen a large $t_{\rm sep} \simeq 2\,{\rm fm}$, while for $q = s$ and $t > 0$, which is the only contribution requiring the spectral density reconstruction method of Eq. (78) and for which statistical accuracy is of the upmost importance [8], we have chosen $t_{\rm sep} \simeq 1\,{\rm fm}$, after checking ground-state-dominance using the larger value $t_{\rm sep} \simeq 1.7\,{\rm fm}$. For the same reason, the inversions of the Dirac operator for $q = s$ and $t > 0$ have been performed using a number of stochastic sources which is eight times larger than that used for $q = b$ and $q = s$, $t < 0$.

We now discuss our determination of $\bar F_T$ starting from the b-quark contribution $\bar F^b_T$. In this case, since there is no problem of analytic continuation, we proceed as in Eq. (74), and evaluate the hadronic tensor $H^{\mu\nu}_{\bar T_b}$ using [9]

$$H^{\mu\nu}_{\bar T_b}(E, \boldsymbol{k}) = a \sum_{t=-t_{\rm sep}}^{T/2-t_{\rm sep}} e^{Et}\, \bar C^{\mu\nu}_b(t, \boldsymbol{k}) , \qquad E = m_{B_s} - E_\gamma . \tag{91}$$

For any simulated heavy-strange meson mass m_{H_s} the corresponding energy E in the previous equation is understood to be

$$E = m_{H_s} - E_\gamma = m_{H_s}\left(1 - \frac{x_\gamma}{2}\right) . \tag{92}$$

From the knowledge of $H^{\mu\nu}_{\bar T_b}(E)$ we use Eq. (66) to determine the b-quark contribution $\bar F^b_T$ to the form factor $\bar F_T$. In the rest frame of the decaying meson, and with our choice of the photon momentum ($\boldsymbol{k} = (0, 0, k_z)$), the form factor can be obtained using

$$\bar F^b_T(x_\gamma) = -\frac{1}{2k_z}\left(H^{12}_{\bar T_b}(E, \boldsymbol{k}) - H^{21}_{\bar T_b}(E, \boldsymbol{k})\right) , \qquad |k_z| = E_\gamma = m_{H_s}\frac{x_\gamma}{2} . \tag{93}$$

Our determination of $\bar F^b_T$ for the four different simulated values of x_γ and for the three different heavy-strange meson masses m_{H_s}, is shown in the left panel of Figure 12. The blue and red colors in the left panel correspond to our results on the B64 and D96 ensemble, respectively. As the figure shows, we find that cut-off effects are very small, the x_γ behaviour is almost linear and the form factor decreases as the heavy-quark mass m_{H_s} increases. However, we postpone the discussion of the extrapolation to the physical mass m_{B_s} to Section IV B, and concentrate here only on the issue of the continuum extrapolation. Having only two lattice spacings available, and given the smallness of the observed UV cut-off effects, we opt for extrapolating to the continuum limit at fixed m_{H_s} and x_γ, employing either a constant or linear Ansatz in a^2. We then combine the results of the two extrapolations using the following criterion: if the constant fit gives a χ^2 smaller than two, we combine the results of the linear and constant fit using the weighted average already illustrated in Eqs. (35)-(36) but using same weights for the linear and constant a^2 extrapolation, otherwise the final result is given by the result obtained using the linear a^2 Ansatz. The result of the continuum-limit extrapolation is illustrated in the right panel of Figure 12.

We now turn into the discussion of the more involved strange-quark contribution. In this case, as already discussed, the form factor cannot be obtained as in Eq. (91) since

$$\lim_{T\to\infty} a \sum_{t=-t_{\rm sep}}^{T/2-t_{\rm sep}} e^{Et}\, \bar C^{\mu\nu}_s(t, \boldsymbol{k}) = \infty , \tag{94}$$

---

[8] The statistical accuracy of the computed $M^{\mu\nu}_{\bar T,q}(t, t_{\rm sep}, \mathbf{k})$ decreases as $t_{\rm sep}$ increases.

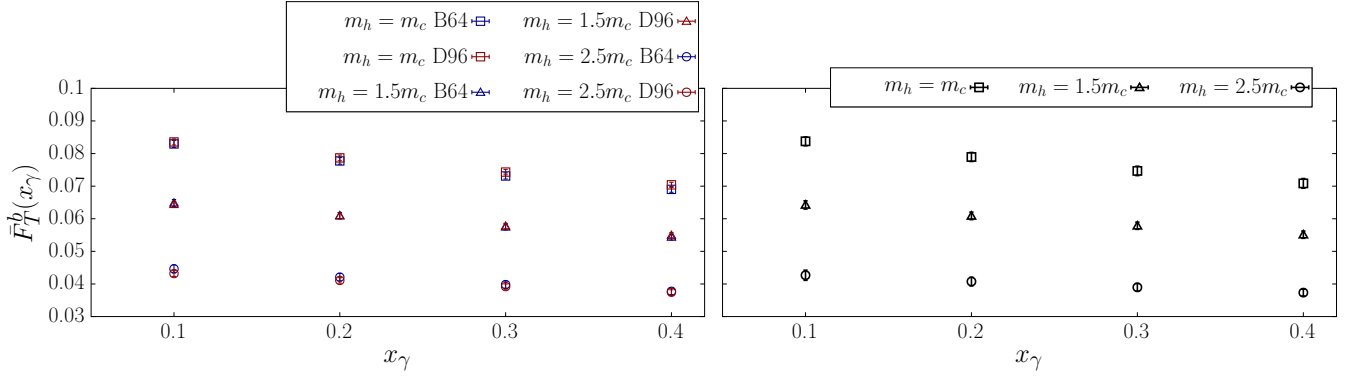[9] With respect to Eq. (21), we have dropped the e^{-ET} term, which is numerically negligible.

FIG. 12: *Left: values of $\bar{F}_T^b$ obtained on the B64 (blue data points) and D96 (red data points) ensembles as a function of x_γ. Right: results for $\bar{F}_T^b$ after extrapolation to the continuum as a function of x_γ, obtained following the procedure described in the text. The different symbols correspond to the different values of the simulated heavy quark mass m_h.*

due to the fact that for large and positive times t the correlation function behaves approximately as

$$\bar{C}_s^{\mu\nu}(t,\boldsymbol{k}) \simeq e^{-E^* t}, \qquad E^* < E . \tag{95}$$

Our strategy to evaluate the contribution from the second time-ordering, which is the only one affected by the problem of the analytic continuation, is to consider the smeared (or regularized) hadronic amplitude introduced in Eq. (78), namely

$$H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k};\varepsilon) = \int_{E^*}^\infty \frac{dE'}{2\pi} K(E'-E;\varepsilon)\,\rho^{\mu\nu}(E',\boldsymbol{k}) , \tag{96}$$

or equivalently, separating the real and imaginary part,

$$\mathrm{Re}\big[H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k};\varepsilon)\big] = \int_{E^*}^\infty \frac{dE'}{2\pi} K_{\mathrm{Re}}(E'-E;\varepsilon)\,\rho^{\mu\nu}(E',\boldsymbol{k}) , \qquad K_{\mathrm{Re}}(x;\varepsilon) = \mathrm{Re}\big[K(x;\varepsilon)\big] , \tag{97}$$

$$\mathrm{Im}\big[H_{\bar{T}_s,2}^{\mu\nu}(E,\boldsymbol{k};\varepsilon)\big] = \int_{E^*}^\infty \frac{dE'}{2\pi} K_{\mathrm{Im}}(E'-E;\varepsilon)\,\rho^{\mu\nu}(E',\boldsymbol{k}) , \qquad K_{\mathrm{Im}}(x;\varepsilon) = \mathrm{Im}\big[K(x;\varepsilon)\big] , \tag{98}$$

and then to perform the extrapolation to $\varepsilon = 0$. Eqs. (97) and (98) can be evaluated, from the knowledge of $C_{s,2}^{\mu\nu}(t,\boldsymbol{k})$ only, using the HLT method and we now briefly summarize the main ingredients of the procedure. To simplify the notation, we concentrate directly on the Lorentz indices that are relevant for the determination of the form factor with our choice of kinematics (decaying meson at rest, and $k = (0,0,k_z)$), and define

$$\bar{C}(t,\boldsymbol{k}) \equiv \frac{1}{2}\big[\bar{C}_s^{12}(t,\boldsymbol{k}) - \bar{C}_s^{21}(t,\boldsymbol{k})\big] , \tag{99}$$

$$H_2(E,\boldsymbol{k},\varepsilon) \equiv \frac{1}{2}\Big[H_{\bar{T}_s,2}^{12}(E,\boldsymbol{k},\varepsilon) - H_{\bar{T}_s,2}^{21}(E,\boldsymbol{k},\varepsilon)\Big] , \tag{100}$$

$$\rho(E',\boldsymbol{k}) \equiv \frac{1}{2}\big[\rho^{12}(E',\boldsymbol{k}) - \rho^{21}(E',\boldsymbol{k})\big] . \tag{101}$$

The final goal is to find, for fixed ε, the best approximation of the kernel functions $K_{\mathrm{Re/Im}}(E'-E;\varepsilon)$, in terms of the basis function $\{e^{-aE'n}\}_{n=1,\ldots,n_{max}}$, namely

$$K_{\mathrm{I}}\left(E'-E;\varepsilon\right) \simeq \sum_{n=1}^{n_{max}} g_{\mathrm{I}}(n,E,\varepsilon)\,e^{-aE'n} \equiv \widetilde{K}_{\mathrm{I}}\left(E',E;\varepsilon\right) , \tag{102}$$

where I = {Re, Im}. In this way, once the coefficients g_{I} are known, the smeared hadronic amplitude can be reconstructed,

from the knowledge of $\bar{C}$, using

$$
\begin{aligned}
H_2(E, \boldsymbol{k}; \varepsilon) &\equiv \int_{E^*}^{\infty} \frac{dE'}{2\pi} \rho(E', \boldsymbol{k}) \, K(E' - E; \varepsilon) \\
&\simeq \sum_{n=1}^{n_{\max}} \left(g_{\mathrm{Re}}(n, E, \varepsilon) + i g_{\mathrm{Im}}(n, E, \varepsilon) \right) \int_{E^*}^{\infty} \frac{dE'}{2\pi} e^{-aE'n} \rho(E', \boldsymbol{k}) \\
&= \sum_{n=1}^{n_{\max}} \left(g_{\mathrm{Re}}(n, E, \varepsilon) + i g_{\mathrm{Im}}(n, E, \varepsilon) \right) \bar{C}(na, \boldsymbol{k}) \; .
\end{aligned}
\tag{103}
$$

The problem of finding the coefficients g_{I} presents a certain number of technical difficulties. Any determination of the real and imaginary part of the smeared hadronic amplitude based on Eqs. (102) and (103) will inevitably be affected by both systematic errors (due to the inexact reconstruction of the kernels) and statistical uncertainties (due to the fluctuations of the correlator $\bar{C}(t, \boldsymbol{k})$), which need to be simultaneously kept under control. The HLT method finds an optimal balance between the size of the statistical and systematic errors. This is achieved by minimizing a linear combination

$$
W_{\mathrm{I}}[\boldsymbol{g}] \equiv \frac{A_{\mathrm{I}}[\boldsymbol{g}]}{A_{\mathrm{I}}[\boldsymbol{0}]} + \lambda B[\boldsymbol{g}] \; ,
\tag{104}
$$

of the norm-functional

$$
A_{\mathrm{I}}[\boldsymbol{g}] = \int_{E_{\min}}^{\infty} \mathrm{d}E' \left| \sum_{n=1}^{n_{\max}} g(n) e^{-aE'n} - K_{\mathrm{I}}\left(E' - E; \varepsilon\right) \right|^2 \; ,
\tag{105}
$$

which quantifies the difference between the approximated and the target kernel, and of the error-functional

$$
B[\boldsymbol{g}] = B_{\mathrm{norm}} \sum_{n_1, n_2=1}^{n_{\max}} g(n_1) \, g(n_2) \, \mathrm{Cov}(an_1, an_2) \; ,
\tag{106}
$$

where $\mathrm{Cov}(an_1, an_2)$ is the covariance matrix of the correlator $\bar{C}(an)$, and B_{norm} is a normalization factor introduced to render the error-functional dimensionless. The algorithmic parameter $E_{\min}$ should only satisfy the constraint $E_{\min} < E^*$, and we choose $E_{\min} = 0.9 \sqrt{m_\phi^2 + E_\gamma^2}$. For each simulated value of x_γ and m_{H_s}, we choose $n_{\max}$ by requiring that the statistical error on the correlation function $\bar{C}(t, \boldsymbol{k})$ for all times $t \leq an_{\max}$ must be smaller than 30%. The parameter λ in Eq. (104) is the so-called *trade-off* parameter, and for a given value of λ, the minimization of the functional $W_{\mathrm{I}}^\alpha[\boldsymbol{g}]$ gives the coefficients $\boldsymbol{g}_{\mathrm{I}}^{\boldsymbol{\lambda}}$. In the presence of statistical errors, the second term in Eq. (104) disfavours coefficients $\boldsymbol{g}$ leading to too large statistical uncertainties in the reconstructed value of the smeared hadronic amplitude. The optimal balance between having small statistical errors (small $B[\boldsymbol{g}]$) and small systematic errors in the kernel reconstruction (small $A_{\mathrm{I}}[\boldsymbol{g}]$) can be achieved by tuning λ appropriately. This is done performing the so-called *stability-analysis*, which is discussed in detail in Refs. [8, 37]. In brief, using the stability-analysis one monitors the evolution of the reconstructed values of the real and imaginary part of $H_2(E, \boldsymbol{k}; \varepsilon)$ as a function of λ. The optimal value, $\lambda^\star$, (which is generally different for the real and imaginary parts) is chosen to be in the so-called statistically-dominated regime, where λ is sufficiently small that the systematic error due to the kernel reconstruction is smaller than the statistical one (in this region the results are therefore stable under variations of λ), but large enough to still have reasonable statistical uncertainties. Finally, having determined the optimal value $\lambda^\star$, we repeat the calculation using a second (smaller) value of $\lambda = \lambda^{\star\star}$, which is determined by imposing the validity of the following condition

$$
\frac{B[\boldsymbol{g}_{\mathrm{I}}^{\boldsymbol{\lambda}^{\star\star}}]}{A_{\mathrm{I}}[\boldsymbol{g}_{\mathrm{I}}^{\boldsymbol{\lambda}^{\star\star}}]} = \kappa \, \frac{B[\boldsymbol{g}_{\mathrm{I}}^{\boldsymbol{\lambda}^\star}]}{A_{\mathrm{I}}[\boldsymbol{g}_{\mathrm{I}}^{\boldsymbol{\lambda}^\star}]} \; ,
\tag{107}
$$

with $\kappa = 10$. Any statistically-significant difference between the values of the real and imaginary part of $H_2(E, \boldsymbol{k})$ corresponding to the two choices $\lambda = \lambda^\star$ and $\lambda = \lambda^{\star\star}$ is added as a systematic uncertainty in our final error. We refer the reader to Ref. [37] for further details on this point.

At a finite lattice spacing, similarly to what had been done in Ref. [8], we adopt the kernel function

$$
K(x; \varepsilon) = \frac{a}{\sinh\left[a(x - i\varepsilon)\right]} = \frac{1}{x - i\varepsilon} + \mathcal{O}(a^2) \; ,
\tag{108}
$$

which differs from the one in Eq. (79) only by $\mathcal{O}(a^2)$ cut-off effects. A major difference in the analysis of $H_2(E, \boldsymbol{k})$ compared to the strategy followed for $H_{\bar{T}_b}^{\mu\nu}(E, \boldsymbol{k})$ and $H_{\bar{T}_s,1}^{\mu\nu}(E, \boldsymbol{k})$, concerns the scaling of the energy E with the heavy-strange meson mass, m_{H_s}. While the energy-scaling given by Eq. (92) leads to a smooth mass dependence for the latter two contributions, this is not the case for $H_2(E, \boldsymbol{k})$: the main contributions to the spectral density $\rho(E', \boldsymbol{k})$ are expected to depend on the position of the ϕ, $\phi(1680)$, $\phi(2170)$ (and possibly heavier) resonances. By scaling the energy E according to Eq. (92), to our lightest simulated mass $m_{H_s} = m_{D_s}$ would correspond an energy E smaller or very close to that of one of the main $s\bar{s}$ peaks. On the other hand, the energy $E = m_{B_s}(1 - x_\gamma/2)$, corresponding to the physical mass of the B_s meson, is much larger than the energy of such resonances. Since the behaviour of $H_2(E, \boldsymbol{k}; \varepsilon)$ below (or close to) the main $s\bar{s}$ resonances is expected to be very different from the one at much larger energies of order $\mathcal{O}(m_{B_s})$, the mass scaling of $H_2(E, \boldsymbol{k}; \varepsilon)$ that would result from the use of Eq. (92) is very complicated and difficult to handle. At the same time setting $E = m_{B_s}(1 - x_\gamma/2)$ for all simulated m_{H_s} is problematic, as it leads to large cut-off effects. For $H_2(E, \boldsymbol{k}; \varepsilon)$ we thus chose to scale the energy E with the heavy-strange meson mass m_{H_s} according to

$$E(r) = (r\, m_{H_s} + (1-r)\, m_{B_s}) \left(1 - \frac{x_\gamma}{2}\right) . \tag{109}$$

Note that any fixed r is allowed since

$$\lim_{m_{H_s} \to m_{B_s}} E(r) = m_{B_s} \left(1 - \frac{x_\gamma}{2}\right) , \tag{110}$$

and for $0 \le r \le 1$ one interpolates between the scaling in Eq. (92) ($r = 1$) and the case of a fixed energy $E = m_{B_s}(1 - x_\gamma/2)$ ($r = 0$). For each x_γ we tune the value of r in such a way that for $m_{H_s} = m_{D_s}$, since $m_{H_s}(1 - x_\gamma/2)$ is the closest to the resonance region, the corresponding energy $E(r)$ is above that of the main $s\bar{s}$ peaks, and at the same time small enough to avoid large cut-off effects. We choose $r = 0.65, 0.60, 0.57, 0.55$, respectively for $x_\gamma = 0.1, 0.2, 0.3$ and 0.4 for all three values of m_{H_s}. Finally, we define the *smeared* form factor as

$$\bar{F}_T^s(x_\gamma; \varepsilon) \equiv -\frac{H(E, E(r), \boldsymbol{k}; \varepsilon)}{k_z} , \qquad \boldsymbol{k} = (0, 0, k_z) , \tag{111}$$

$$H(E, E(r), \boldsymbol{k}; \varepsilon) = H(0, \boldsymbol{k}) + H_1^{\text{sub}}(E, \boldsymbol{k}) + H_2^{\text{sub}}(E(r), \boldsymbol{k}; \varepsilon) , \tag{112}$$

where

$$H(0, \boldsymbol{k}) = a \sum_{t=-t_{\text{sep}}}^{T/2-t_{\text{sep}}} \bar{C}(t, \boldsymbol{k}) , \qquad H_1^{\text{sub}}(E, \boldsymbol{k}) = a \sum_{t=-t_{\text{sep}}}^{0} \left(e^{Et} - 1\right) \bar{C}(t, \boldsymbol{k}) \tag{113}$$

$$H_2^{\text{sub}}(E(r), \boldsymbol{k}; \varepsilon) = H_2(E(r), \boldsymbol{k}; \varepsilon) - H_2(0, \boldsymbol{k}; 0) . \tag{114}$$

In the combined $m_{H_s} \to m_{B_s}$ and $\varepsilon \to 0$ limits, the smeared form factor tends to $\bar{F}_T^s(x_\gamma)$. The zero-energy subtraction allows us to define the contributions from the two time-orderings in such a way that cut-off effects start at order $\mathcal{O}(a^2)$ for both time-orderings. This is because they both are now free of the contact term $\bar{C}(0, \boldsymbol{k})$. This contact term does not belong to either the first or second time ordering, and cannot be simply removed as this generates $\mathcal{O}(a)$ cut-off effects. Since $H_2(E(r), \boldsymbol{k}; \varepsilon)$ is evaluated via the HLT reconstruction method using the kernel function in Eq. (108), to avoid the presence of $\mathcal{O}(a)$ cut-off effects also $H_2(0, \boldsymbol{k}, 0)$ is evaluated via the HLT method using the same type of kernel function. Being able to define the two time-orderings separately turns out to be useful if a model for the spectral density $\rho(E', \boldsymbol{k})$ is used to perform the $\varepsilon \to 0$ extrapolation, as will be discussed below.

In the plot of Figure 13, we give an example of the stability analysis in the case of the lowest simulated quark mass, and for $x_\gamma = 0.1$ and $\varepsilon \simeq 1.4$ GeV. In the figure we show the real and imaginary part of the smeared form factor $\bar{F}_T^s(x_\gamma; \varepsilon)$, obtained employing different values of the trade-off parameter λ. The results are shown as a function of $A_{\text{I}}[\boldsymbol{g}^\lambda]/A_{\text{I}}[\boldsymbol{0}]$ which is a measure of the goodness of the reconstruction. When the systematic error due to the inexact kernel reconstruction becomes smaller than the statistical uncertainty, the reconstructed smeared form factor is stable under variation of λ. In this region we determine $\lambda^\star$ and $\lambda^{\star\star}$ which are given respectively by the red and blue data points in the figure. The reconstructed kernel functions corresponding to our choice of $\lambda^\star$ are then shown in Figure 14.

We have repeated the analysis for different values of ε and for all simulated x_γ and m_{H_s}. The smearing parameter ε cannot be however reduced arbitrarily since the uncertainties on $F_T^s(x_\gamma; \varepsilon)$ generally increase as ε decreases, and at the same time the reconstruction of the kernel function becomes poorer. The smallest value of ε for which the errors are still under control is determined by both the statistical uncertainties on $\bar{C}(t, \boldsymbol{k})$ and by the size n_{max} of the exponential basis.
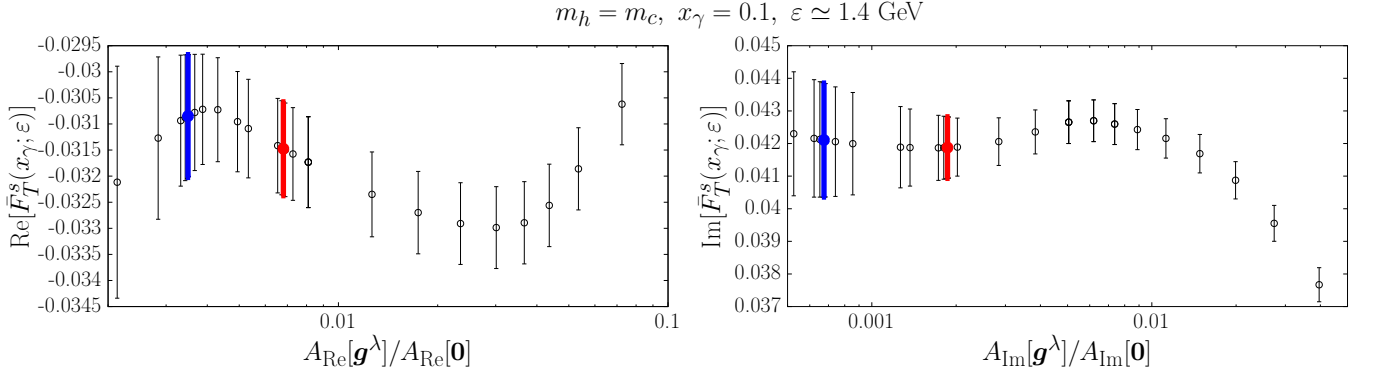
FIG. 13: *The real (left panel) and imaginary (right panel) part of the form factor $\bar{F}_T^s(x_\gamma;\varepsilon)$ on the B64 ensemble, for the lowest simulated value of $m_h = m_c$, and for $\varepsilon \simeq 1.4$ GeV and $x_\gamma \simeq 0.1$, as a function of the ratio $A_{\rm I}[\boldsymbol{g}^\lambda]/A_{\rm I}[\boldsymbol{0}]$ indicating the quality of the kernel reconstruction obtained employing different values of λ. The plot shows an example of our stability analysis. The red and blue data points in both panels correspond to the reconstructions obtained for $\lambda = \lambda^\star$ and $\lambda = \lambda^{\star\star}$, respectively.*
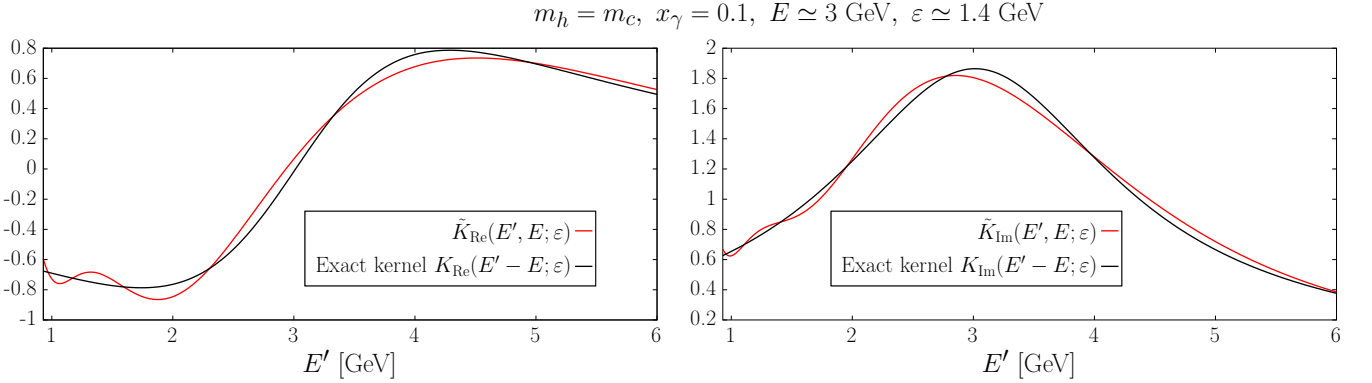
FIG. 14: *The reconstructed smearing kernels $\widetilde{K}_{\rm I}(E',E;\varepsilon)$ obtained using the coefficients $\boldsymbol{g}_{\rm I}^{\lambda^\star}$ employed in the reconstruction of $\bar{F}_T^s(x_\gamma,\varepsilon)$ in Figure 13. The black lines correspond to the real (right panel) and imaginary (left panel) part of the exact kernel function $K(E'-E;\varepsilon)$ of Eq. (108). Th kernels are given in lattice units.*

In the plots of Figures 15 and 16 we show the ε-behaviour of the real and imaginary parts of the smeared form factors for the different simulated heavy-strange meson masses m_{H_s}, and for the smallest (0.1) and largest (0.4) simulated values of x_γ. In the figure we show the results obtained on both the B64 and D96 ensembles. Few comments are in order. First of all the observed cut-off effects are smaller or of the same size of the statistical error for all contributions, with the exception of $\bar{F}_T^s(x_\gamma;\varepsilon)$ for $m_{H_s} \simeq 1.78\,m_{D_s}$. Such behaviour can be expected since larger masses correspond to higher energies $E(r)$. In addition, we observe that both the real and imaginary part of $\bar{F}_T^b(x_\gamma)$ decrease in magnitude as m_{H_s} increases.

We extrapolate the smeared form factor $\bar{F}_T^s(x_\gamma;\varepsilon)$ to the continuum limit at fixed ε, x_γ and m_h, following the same procedure used for $\bar{F}_T^b(x_\gamma)$. Next we perform the $\varepsilon \to 0$ extrapolation at fixed x_γ and m_h, which is the most delicate step of the analysis. As already stated, using the kernel function $K(x;\varepsilon) = (x - i\varepsilon)^{-1}$, the leading corrections to the $\varepsilon = 0$ limit are expected to be of the form

$$\bar{F}_T^s(\varepsilon) = \bar{F}_T^s + A_1\,\varepsilon + A_2\,\varepsilon^2 + \mathcal{O}(\varepsilon^3)\ , \tag{115}$$

and in the following we indicate by asymptotic regime, the regime in which the corrections to the vanishing-ε limit can be described by a low-degree polynomial in ε. The onset of the asymptotic regime for $H_2(E,\boldsymbol{k};\varepsilon)$ at a given energy E, as discussed in detail in Ref. [8], depends on the (unknown) typical size, $\Delta(E)$, of the interval around E in which $H_2(E,\boldsymbol{k})$ is significantly varying. Parametrically one must then have $\varepsilon \ll \Delta(E)$ and at the same time $\varepsilon \gg 1/L$ to avoid large FSEs. Assuming that $H_2(E,\boldsymbol{k})$ is dominated by the contribution from a single resonance, which at fixed
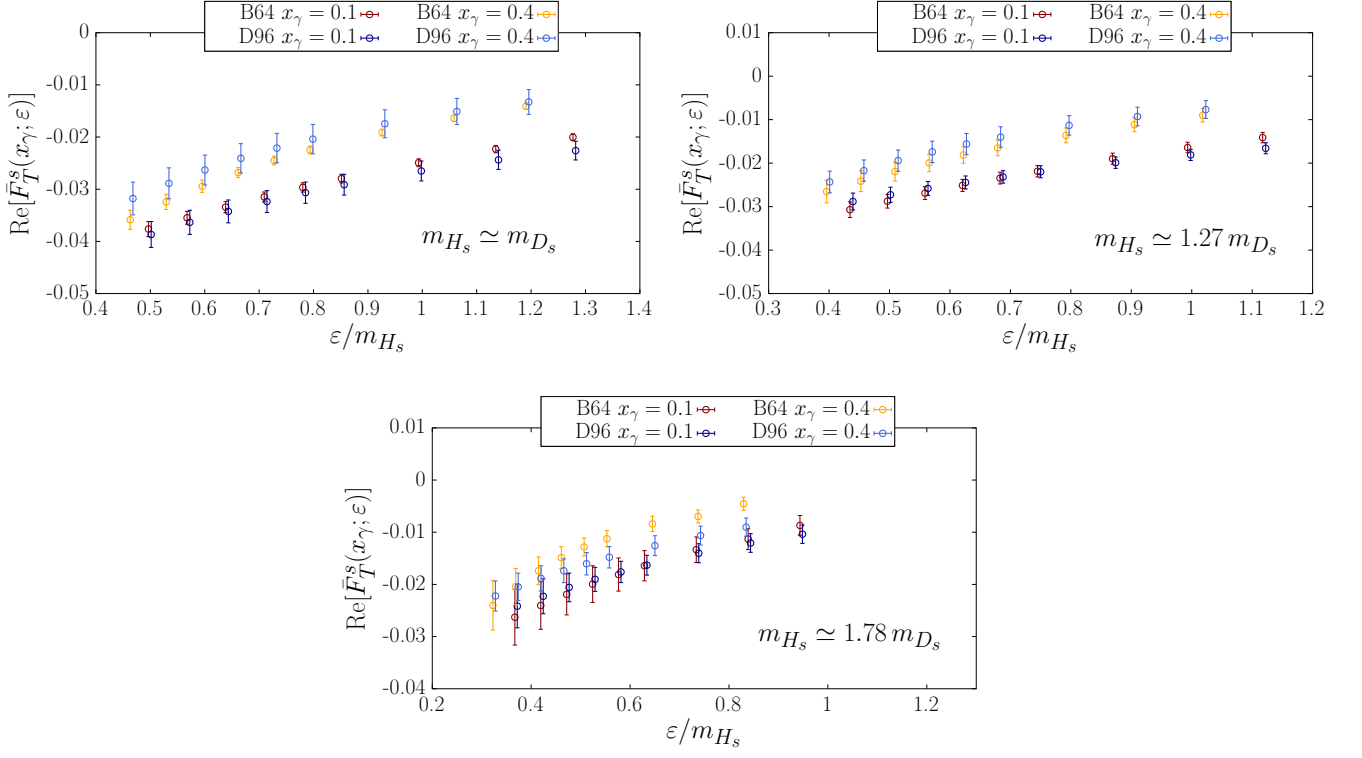
FIG. 15: ε-dependence of the real part of the smeared form factor $\bar{F}_T^s(x_\gamma; \varepsilon)$ for the three different masses $m_{H_s} \simeq m_{D_s}$ (top-left panel), $m_{H_s} \simeq 1.27\, m_{D_s}$ (top-right panel), and $m_{H_s} \simeq 1.78\, m_{D_s}$ (bottom panel). For each figure we show the results obtained at $x_\gamma = 0.1$ and $x_\gamma = 0.4$ on the B64 and D96 ensembles.
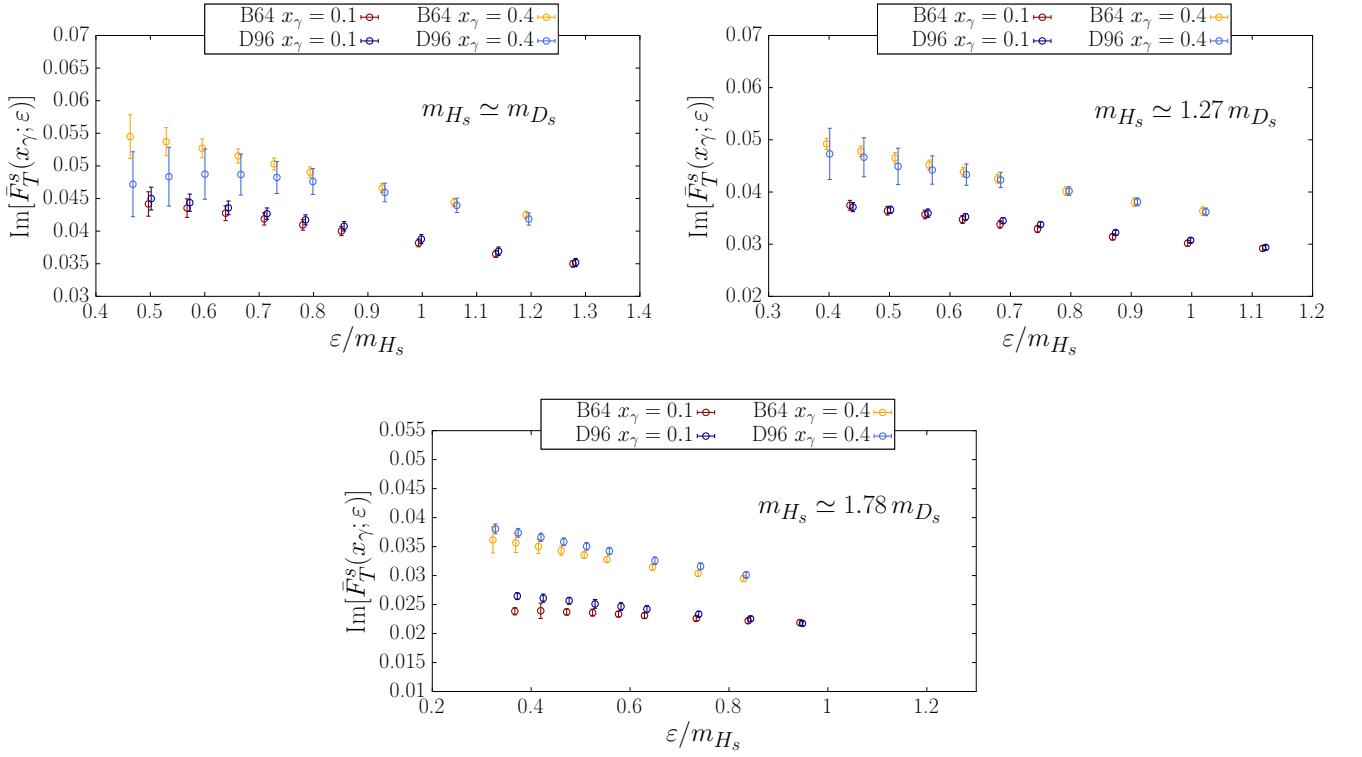

FIG. 16: Same as in Figure 15 for the imaginary part of the smeared form factor $\bar{F}_T^s(x_\gamma; \varepsilon)$.

x_γ and m_{H_s} we approximate with a Breit-Wigner distribution centered at M and of width Γ, i.e.

$$H_2(E, \boldsymbol{k}) \simeq H_2^{\mathrm{BW}}(E) = \frac{R}{M - E - i\frac{\Gamma}{2}} \quad \implies \quad H_2(E, \boldsymbol{k}; \varepsilon) \simeq \frac{R}{M - E - i(\frac{\Gamma}{2} + \varepsilon)} \ , \tag{116}$$

then within this approximation we have $\Delta(E) = \sqrt{(E - M)^2 + \Gamma^2/4}$. In our case the energy E is given for each x_γ and m_{H_s} by $E(r)$ in Eq. (109), and it ranges from $E(r) \simeq 3\,\mathrm{GeV}$ at the lowest mass $m_{H_s} = m_{D_s}$ to $E(r) \simeq 4\,\mathrm{GeV}$ at $m_{H_s} \simeq 1.78\,m_{D_s}$. The peaks of the main $s\bar{s}$ resonances, the ϕ, $\phi(1680)$ and $\phi(2170)$, are at $M \simeq 1, 1.7, 2.2$ GeV respectively, with a mild dependence on the value of $|\boldsymbol{k}|$. In our computations we have $\varepsilon \simeq \mathcal{O}(1\,\mathrm{GeV})$ or higher and it is not clear whether such values of ε are in the asymptotic regime, despite an approximate linear scaling in ε being observed in Figure 15 and 16. To account for this source of systematic error we proceed as follows: we first carry out the extrapolation to $\varepsilon = 0$ assuming that the observed behaviour is the asymptotic scaling, and perform a polynomial extrapolation in ε (in practice, as explained below, we perform a quadratic extrapolation in ε, unless there is no signal of a ε^2 term, in which case we perform a linear extrapolation). In addition to the polynomial extrapolation, we follow a second approach, performing the vanishing-ε extrapolation assuming the following model for the spectral density

$$\rho^{\mathrm{mod}}(E', \boldsymbol{k}) = \sum_{n=1,2,3} \theta(E_\theta - E') \frac{R_n(\boldsymbol{k})\Gamma_n}{\left(E_n(\boldsymbol{k}) - E'\right)^2 + \left(\frac{\Gamma_n}{2}\right)^2} + \theta(E' - E_\theta(\boldsymbol{k})) \frac{D(\boldsymbol{k})}{(E')^z}, \qquad E_n(\boldsymbol{k}) = \sqrt{M_n^2 + |\boldsymbol{k}|^2} \ . \tag{117}$$

with $z = 1/2, 1$, and where M_n and Γ_n are the mass and the decay width of the ϕ, $\phi(1680)$ and $\phi(2170)$ resonances respectively for $n = 1, 2, 3$, which we take from the PDG[10]. The last term in Eq. (117) mimicks the continuum behaviour at large energies starting at a threshold $E_\theta(\boldsymbol{k}) > E_3(\boldsymbol{k})$, and is compatible with the physical constraint $\lim_{E' \to \infty} \rho(E', \boldsymbol{k}) = 0$. [10] For any fixed values of x_γ and m_{H_s}, our model for the spectral density contains four free real parameters: the three amplitudes $R_1(\boldsymbol{k})$, $R_2(\boldsymbol{k})$, $R_3(\boldsymbol{k})$, and the threshold energy $E_\theta(\boldsymbol{k})$. The parameter $D(\boldsymbol{k})$ is instead determined by the requirement that the spectral density is continuous at $E' = E_\theta(\boldsymbol{k})$. The smeared hadronic amplitude $H_2^{\mathrm{mod}}(E, \boldsymbol{k}; \varepsilon)$ associated with $\rho^{\mathrm{mod}}(E', \boldsymbol{k})$ is then given by the convolution of $\rho^{\mathrm{mod}}(E', \boldsymbol{k})$ with the kernel function $K(E - E'; \varepsilon)$. Finally, we can use $H_2^{\mathrm{mod}}(E', \boldsymbol{k}; \varepsilon)$ to obtain the corresponding model smeared form factor,

$$\bar{F}_T^{s,\mathrm{mod}}(x_\gamma; \varepsilon) \equiv -\frac{H^{\mathrm{mod}}(E, E(r), \boldsymbol{k}; \varepsilon)}{k_z} \ , \tag{118}$$

$$H^{\mathrm{mod}}(E, E(r), \boldsymbol{k}; \varepsilon) \equiv H_2^{\mathrm{mod}}(E(r), \boldsymbol{k}; \varepsilon) - H_2^{\mathrm{mod}}(0, \boldsymbol{k}; 0) + H_1^{\mathrm{sub}}(E, \boldsymbol{k}) + H(0, \boldsymbol{k}) \ . \tag{119}$$

In Figure 17 we show the results of the polynomial extrapolation to vanishing ε, which we perform separately for each x_γ and m_{H_s}. In the figure we show, as an illustration, the results obtained for $x_\gamma = 0.1$ and 0.4. The extrapolation has been carried out using the following Ansatz for the smeared form factor

$$\bar{F}_T^s(\varepsilon) = A + A_1\,\varepsilon + A_2\,\varepsilon^2 \ , \tag{120}$$

where A, A_1 and A_2 are complex-valued free fit parameters, which are different for each x_γ and m_{H_s}. In order to avoid overfitting, for those cases when there is no signal of ε^2-dependence visible in the data, we have set $A_2 = 0$. We have minimized a χ^2-function constructed without taking into account the correlation between the values of the smeared form factors corresponding to different ε, since they are too correlated, and the resulting correlation matrix is ill-conditioned. In this way, the reduced χ^2 resulting from the minimization, which is always well below one, cannot be taken as a quantitative measure of the quality of the fit. To estimate the systematic error of the polynomial extrapolation, we have also performed for all the cases a second fit, linear in ε, using only the five smallest simulated values of ε. Any statistically-significant deviation from the results obtained in the fit with all simulated values of ε included (i.e. those whose resulting best-fit functions are given by the coloured bands of Figure 17) is then added as a systematic error. In Figure 17 the data points at $\varepsilon = 0$ correspond to our final results from the polynomial extrapolation, after including the systematic error determined following the procedure described above. As is clear from the figure the real and imaginary part of the form factor $\bar{F}_T^s(x_\gamma)$ decrease in magnitude as the mass m_{H_s} increases, and already for $m_{H_s} \simeq 1.78\,m_{D_s}$ they are both one order of magnitude smaller than the tensor form factors F_{TV} and F_{TA} determined in the previous section.

As discussed above, since the simulated values of ε may not be in the asymptotic regime, we have also performed non-polynomial extrapolations in ε using the model in Eq. (117) with $z = 1/2$ and $z = 1$. The fits have been performed

---

[10] The spectral density must vanish in the infinite-energy limit in order to have a finite $\mathrm{Re}[H_2(E, \boldsymbol{k}; \varepsilon)]$.
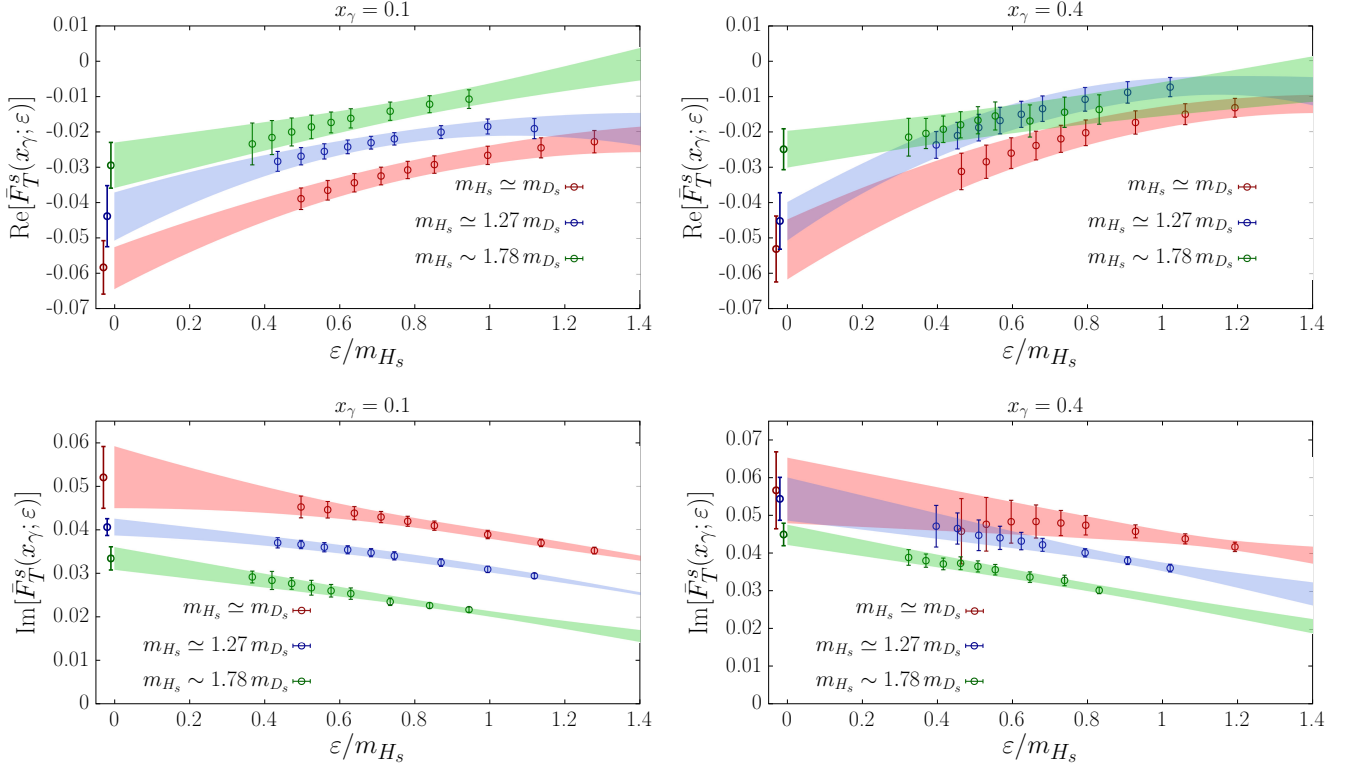
FIG. 17: *The extrapolation in ε of the real (top panels) and imaginary (bottom panels) components of the smeared form factors, using the Ansatz in Eq. (120). For those cases when no ε^2-dependence is visible in the data, we have set $A_2 = 0$. The different colours correspond to the three different values of m_{H_s} and the corresponding bands are the best-fit functions obtained from the fits with all values of ε included. The left and right panels correspond to $x_\gamma = 0.1$ and $x_\gamma = 0.4$ respectively. The data points at $\varepsilon = 0$ (which in the figures are slightly shifted horizontally for better visualization) correspond to our final results after including the systematic error, determined as discussed in the text.*

imposing Gaussian priors on all four fit parameters. The prior corresponding to the amplitude $R_1(\boldsymbol{k})$ of the ϕ resonance is

$$R_1(\boldsymbol{k}) = \langle R_1(\boldsymbol{k}) \rangle \left(1 \pm 0.1 \right) , \tag{121}$$

where $\langle R_1(\boldsymbol{k}) \rangle$ has been estimated from an effective residue analysis of the correlation function $\bar{C}(t, \boldsymbol{k})$ at large times. We use $\langle R_1(\boldsymbol{k}) \rangle / |\boldsymbol{k}| = -0.045, -0.042, -0.040$ GeV, respectively for $m_{H_s}/m_{D_s} \simeq 1, 1.27, 1.78$. The priors corresponding to the two amplitudes $R_2(\boldsymbol{k})$ and $R_3(\boldsymbol{k})$ are instead

$$R_2(\boldsymbol{k}) = \frac{\langle R_1(\boldsymbol{k}) \rangle}{2} \left(1 \pm 1 \right) , \qquad R_3(\boldsymbol{k}) = \frac{\langle R_1(\boldsymbol{k}) \rangle}{2} \left(1 \pm 1 \right) , \tag{122}$$

i.e. we assume that, within one standard deviation, they are at most of the same size as the contribution from the ϕ resonance. Finally the prior on the threshold parameter $E_\theta(\boldsymbol{k})$ is

$$E_\theta(\boldsymbol{k}) = E_3(\boldsymbol{k}) + (0.5 \pm 0.5) \text{ GeV} , \tag{123}$$

i.e. we assume that the onset of the perturbative regime occurs at an energy which is $\mathcal{O}(\Lambda_{\text{QCD}})$ larger than that of the heaviest known $s\bar{s}$ resonance. We have found that both values of z describe the data well at all masses and x_γ.[11]

For the real part of $\bar{F}_T^s(x_\gamma)$, the results of the extrapolation to $\varepsilon = 0$, obtained using the model with either $z = 1/2$ and 1 are in good agreement with those of the polynomial extrapolation. For the imaginary part, instead, we find that the

---

[11] Additionally, we have tried to fit our data using the model in Eq. (117) with $z = 2$, but found that it does not provide an equally good description of the smeared form factor.
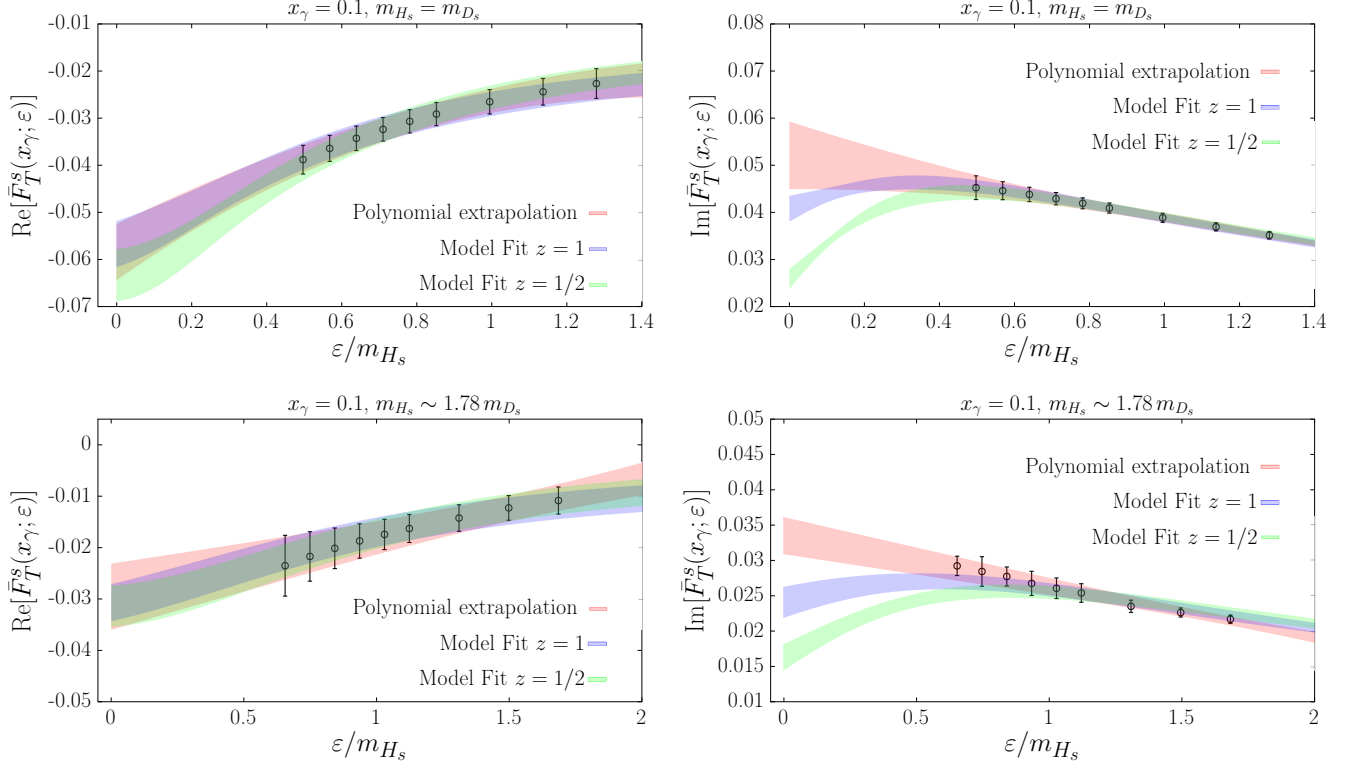
FIG. 18: *Comparison between the results from the polynomial extrapolation (given by the red band), and those obtained fitting the smeared form factor using the model in Eq. (117) with $z = 1/2$ (given by the green band) and $z = 1$ (given by the blue band). The top panels correspond to $x_\gamma = 0.1$, $m_{H_s} = m_{D_s}$, while the bottom panels to $x_\gamma = 0.1$, $m_{H_s} \sim 1.78 m_{D_s}$.*

model results (in particular for $z = 1/2$) are significantly smaller than those obtained from the polynomial extrapolation. The comparison is shown in Figure 18, for the case $x_\gamma = 0.1$ and for both $m_{H_s} = m_{D_s}$ and $m_{H_s} \sim 1.78 m_{D_s}$. All the other cases are very similar. The lower value obtained for $\text{Im}[\bar{F}_T^s(x_\gamma)]$ assuming the model $\rho^{\text{mod}}(E', \boldsymbol{k})$ for the spectral density, could be due to the fact that at the simulated values of ε, the imaginary part of the kernel function still has a sizeable overlap with the peaks of the nearby resonances (e.g. the $\phi(2170)$ resonance). In this case, the imaginary part of the smeared form factor, $\text{Im}[\bar{F}_T^s(x_\gamma; \varepsilon)]$, is expected to decrease in value for smaller, presently unreachable, values of ε. This behaviour cannot be captured by the polynomial extrapolation, but is in-built in our model for the spectral density. To have a realistic estimate of the systematic uncertainty for $\bar{F}_T^s(x_\gamma; \varepsilon)$, we average the results of the polynomial and model-dependent extrapolation with $z = 1/2$, and include a systematic error equal to half the difference between the two results.

B. Extrapolating $\bar{F}_T$ to the mass of the physical B_s-meson

We now discuss the extrapolation of the form factor $\bar{F}_T$ to the mass of the physical B_s-meson. We start from the b-quark contribution $\bar{F}_T^b(x_\gamma)$, which we determined for three values of the heavy-strange meson mass $m_{H_s}/m_{D_s} \simeq 1, 1.27$ and 1.78 as shown in the right panel of Figure 12. To perform the mass extrapolation we make use of a phenomenological VMD-inspired Ansatz to describe the combined x_γ and m_{H_s} dependence of the form factor. At the physical B_s mass point, the form factor $\bar{F}_T^b(x_\gamma)$ is expected to be dominated by the contributions of neutral, $J^P = 1^-$, $b\bar{b}$ resonance states (e.g. $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$, ...). The contribution to the form factor $\bar{F}_T^b(x_\gamma)$ of a given vector resonance state containing an heavy quark (h) and an heavy anti-quark ($\bar{h}$), for a given value of m_{H_s}, and approximating the resonance as a stable state, is of the form

$$\bar{F}_{T,n}^b(x_\gamma) = \frac{q_b f_n m_n g_n^+(0)}{E_n(E_n + E_\gamma - m_{H_s})} + \text{regular terms},\qquad(124)$$
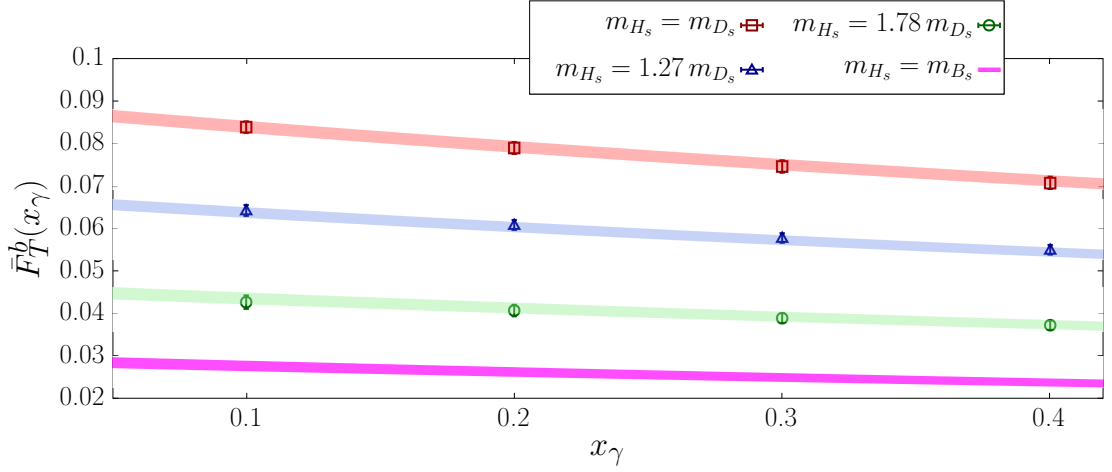
FIG. 19: *Extrapolation to m_{B_s} of the form factor $\bar{F}_T^b(x_\gamma)$ using the Ansatz in Eq. (129). The red, blue and green bands correspond to the best fit-function obtained for $m_{H_s}/m_{D_s} \simeq 1, 1.27, 1.78$, respectively. The magenta band correspond to our final result for $\bar{F}_T^b$ at $m_{H_s} = m_{B_s}$.*

where $E_n = \sqrt{m_n^2 + E_\gamma^2}$, and m_n and f_n are respectively the mass and the electromagnetic decay constant of the vector resonance. The latter is defined through

$$\langle 0|\bar{h}\gamma^\mu h|n(-\boldsymbol{k},\eta)\rangle = \eta^\mu f_n m_n \ , \tag{125}$$

where $|n(-\boldsymbol{k},\eta)\rangle$ is the vector resonance state with given polarization η. The coupling $g_n^+(0)$ is defined by [4]

$$\langle n(-\boldsymbol{k},\eta)|\,\bar{s}\sigma^{\mu\nu}h\,|\bar{H}_s(\boldsymbol{0})\rangle = i\eta_\beta^*\epsilon^{\mu\nu\beta\gamma}\left[g_n^+(p_\gamma^2)(p+q_n)_\gamma + g_n^-(p_\gamma^2)(p-q_n)_\gamma\right] + ig_n^0(p_\gamma^2)(\eta^*\cdot p)\epsilon^{\mu\nu\beta\gamma}p_\beta q_{n,\gamma} \ , \tag{126}$$

with $q_n = (E_n, -\boldsymbol{k})$, $p_\gamma = p - q_n$. In the heavy-quark limit, $m_h \to \infty$, the following scaling laws hold

$$f_n \propto \frac{1}{\sqrt{m_h}} + \ldots \propto \frac{1}{\sqrt{m_{H_s}}} + \ldots \ , \qquad \frac{m_n}{m_{H_s}} = 2 + \frac{\Lambda_T^n}{m_{H_s}} + \ldots \ , \tag{127}$$

where $\Lambda_T^n \simeq \mathcal{O}(\Lambda_{\mathrm{QCD}})$, and the ellipses represent higher-order terms in the heavy-quark expansion. In light of the previous relations $\bar{F}_{T,n}^b(x_\gamma)$ can be further approximated with

$$\bar{F}_{T,n}^b(x_\gamma) = \frac{q_b}{m_{H_s}}\frac{f_n\,g_n^+(0)}{1 + \frac{x_\gamma}{2} + \frac{\Lambda_T^n}{m_{H_s}}}\left(1 + \mathcal{O}\left(x_\gamma, \frac{\Lambda_{\mathrm{QCD}}}{m_{H_s}}\right)\right) \ . \tag{128}$$

Our strategy to extrapolate $\bar{F}_T^b(x_\gamma)$ to the physical mass m_{B_s}, consists in approximating the tower of contributions of type 124, with a single effective pole. This is achieved through the use of the following fit Ansatz for the combined mass and x_γ dependence of the form factor

$$\bar{F}_T^b(x_\gamma, m_{H_s}) = \frac{1}{m_{H_s}}\frac{A + B\,x_\gamma}{1 + \frac{x_\gamma}{2} + \frac{\Lambda_T}{m_{H_s}}} \ , \tag{129}$$

where A, B and Λ_T are free fit parameters, and the effective-pole mass is $m_{\mathrm{eff}} = 2m_{H_s} + \Lambda_T$. We assume that A and B are mass-independent, which is consistent with our data, as illustrated below[12].

Using the Ansatz in Eq. (129) we have performed a combined fit of the x_γ and m_{H_s} dependence of our data. The total number of measurement entering the χ^2 minimization is 12, and the number of fit parameters is 3. The χ^2/dof resulting from the minimization is very good and well below unity, although in this case we have employed an uncorrelated χ^2 function, since we find that the covariance matrix is ill-conditioned. To illustrate the quality of the fit,

---

[12] The Ansatz in Eq. (129) assumes that $g_n^+ \propto \sqrt{m_{H_s}}$ for which however we are not aware of any formal proof in the HQET.

	x_γ			
	0.1	0.2	0.3	0.4
$\bar{F}_T^b$	0.028(1)	0.026(1)	0.025(1)	0.0239(9)

TABLE IV: *Our results for the form factor $\bar{F}_T^b$ extrapolated to the physical mass m_{B_s}, for the four simulated values of $x_\gamma = 0.1, 0.2, 0.3, 0.4$.*
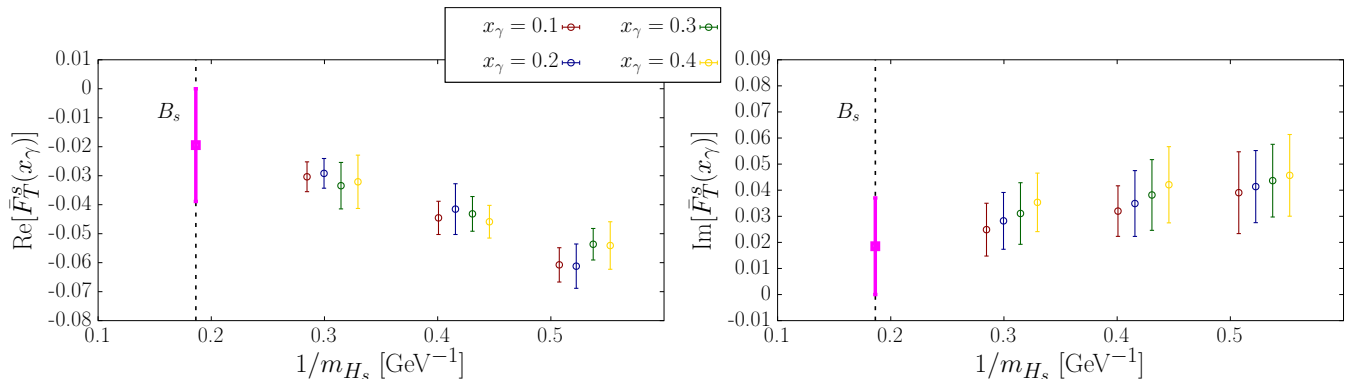

FIG. 20: *Mass dependence of the real (left panel) and imaginary (right panel) part of $\bar{F}_T^s(x_\gamma)$ for the four simulated values of x_γ. The data points corresponding to the four simulated values of x_γ have been slightly shifted horizontally for better visualization. The data point in magenta corresponds to our final (conservative) estimate for $\bar{F}_T^s(x_\gamma)$ at the physical mass $m_{H_s} = m_{B_s}$.*

we show in Figure 19 the best-fit functions obtained from the global fit. As is clear from the figure, the VMD-inspired Ansatz perfectly captures both the mass and x_γ behaviour of our data. The resulting value of the parameter Λ_T is

$$\Lambda_T = -0.32\,(11)\ \text{GeV}\ , \tag{130}$$

which implies that at the physical B_s mass the effective pole is located at $m_{\text{eff}} = 10.4\,(1)\,\text{GeV}$, i.e. around the mass of the $\Upsilon(2S)$ resonance. To check for possible systematic errors due to the mass extrapolation we have repeated the fit setting to zero the parameter B in Eq. (129). However, we did not find significant differences within uncertainties. The magenta band in Figure 19 correspond to our final result for $\bar{F}_T^b(x_\gamma)$ at $m_{H_s} = m_{B_s}$. This contribution turns out to be small compared to the tensor form factors $F_{TV}(x_\gamma)$ and $F_{TA}(x_\gamma)$ described in Section III B, which are more than one order of magnitude larger. In Table IV we give our results for $\bar{F}_T^b$ extrapolated at the physical mass m_{B_s}, for the four simulated values of x_γ.

We now turn into the discussion of the mass extrapolation of $\bar{F}_T^s(x_\gamma)$. In this case the uncertainties are significantly larger than those affecting $\bar{F}_T^b(x_\gamma)$. Moreover, after including the systematic errors due to the $\varepsilon \to 0$ extrapolation, only a very smooth x_γ-dependence is visible in the data within uncertainties. This is shown in Figure 20, where we plot the real and imaginary part of $\bar{F}_T^s(x_\gamma)$ as a function of $1/m_{H_s}$ for all simulated values of x_γ. As is clear from the figure, both the real and imaginary part of $\bar{F}_T^s(x_\gamma)$ decrease as m_{H_s} increases. This is expected since the form factor vanishes in the static limit. In this case, to have a conservative error estimate, we take the results at the largest simulated mass $m_{H_s} \simeq 1.78\,m_{D_s}$ as a bound for the value of the form factor at the physical point, $m_{H_s} = m_{B_s}$. Since no clear x_γ-dependence is visible in the data, we associate the same central value and errors to all x_γ. Our final determination is

$$\text{Re}[\bar{F}_T^s(x_\gamma)] = -0.019(19)\ , \qquad \text{Im}[\bar{F}_T^s(x_\gamma)] = 0.018(18)\,, \tag{131}$$

which correspond to the data points in magenta in the panels of Figure 20.

V.   THE $B_s \to \mu^+\mu^-\gamma$ DECAY RATE

The doubly-differential cross section for the $B_s \to \mu^+\mu^-\gamma$ decay can be written as

$$\frac{\mathrm{d}^2\Gamma}{\mathrm{d}x_\gamma\,\mathrm{d}(\cos\theta)} = \frac{\mathrm{d}^2\Gamma^{(\text{PT})}}{\mathrm{d}x_\gamma\,\mathrm{d}(\cos\theta)} + \frac{\mathrm{d}^2\Gamma^{(\text{INT})}}{\mathrm{d}x_\gamma\,\mathrm{d}(\cos\theta)} + \frac{\mathrm{d}^2\Gamma^{(\text{SD})}}{\mathrm{d}x_\gamma\,\mathrm{d}(\cos\theta)}\,, \tag{132}$$

where the superscript (PT) refers to the point-like contribution (which becomes negligible for large x_γ), (SD) labels the structure-dependent contribution and (INT) labels the contribution from the interference between the point-like and structure-dependent terms in the amplitude. In Eq. (132) θ is the angle between the three-momenta of the μ^+ and the photon in the rest frame of the $\mu^+\mu^-$ pair. Recalling that $x_\gamma = 2p \cdot k/m_{B_s}$, $\cos\theta$ is written in terms of Lorentz invariant quantities in Eq. (135) below. We now present the expressions for the three terms on the right-hand side of Eq. (132), neglecting the contributions from $\mathcal{O}_{1-6,8}$ except for the charming penguin diagram in Figure 4 which is included in the effective Wilson coefficient C_9^{eff}.

The structure-dependent contribution, which depends quadratically on the form factors, can be written as [4] [13]

$$\frac{\mathrm{d}^2\Gamma^{(\text{SD})}}{\mathrm{d}x_\gamma\,\mathrm{d}(\cos\theta)} = \frac{G_F^2\,\alpha_{\text{em}}^3\,m_{B_s}^5}{2^{10}\,\pi^4}\,|V_{tb}V_{ts}^*|^2\,J(x_\gamma)\left[x_\gamma^2\,B_0\left(x_\gamma\right) + x_\gamma\,\xi\left(x_\gamma,\hat{t}\right)\tilde{B}_1\left(x_\gamma\right) + \xi^2\left(x_\gamma,\hat{t}\right)\tilde{B}_2\left(x_\gamma\right)\right], \tag{133}$$

where $\hat{t} = (p - p_1)^2/m_{B_s}^2$ and p_1 is the momentum of the μ^+ lepton. The function $\xi(x_\gamma,\hat{t})$ is defined as

$$\xi(x_\gamma,\hat{t}) = x_\gamma + 2\hat{m}_\mu^2 - 2\hat{t}\,, \tag{134}$$

where $\hat{m}_\mu = m_\mu/m_{B_s}$ [14], in terms of which the angle θ is given by

$$\cos\theta = \frac{\xi(x_\gamma,\hat{t})}{x_\gamma\sqrt{1 - 4\hat{m}_\mu^2/(1 - x_\gamma)}}\,. \tag{135}$$

The Jacobian $J(x_\gamma)$ is

$$J(x_\gamma) = \frac{x_\gamma}{2}\sqrt{1 - \frac{4\hat{m}_\mu^2}{1 - x_\gamma}}\,. \tag{136}$$

Finally the functions $B_0(x_\gamma)$, $\tilde{B}_1(x_\gamma)$ and $\tilde{B}_2(x_\gamma)$ are as follows:

$$B_0\left(x_\gamma\right) = \left(1 - x_\gamma + 4\hat{m}_\mu^2\right)\left(F_1\left(x_\gamma\right) + F_2\left(x_\gamma\right)\right) - 8\hat{m}_\mu^2\,|C_{10}(\mu)|^2\left(F_V^2\left(x_\gamma\right) + F_A^2\left(x_\gamma\right)\right)\,, \tag{137}$$

$$\tilde{B}_1\left(x_\gamma\right) = 8\left[\left(1 - x_\gamma\right)F_V(x_\gamma)\,F_A(x_\gamma)\,\mathrm{Re}\left(C_9^{\text{eff}\,*}(\mu,x_\gamma)\,C_{10}(\mu)\right)\right.$$

$$\left. + \hat{m}_b\,F_V(x_\gamma)\,\mathrm{Re}\left(C_7^*(\mu)\,F_{TA}^{\text{eff}\,*}(x_\gamma)\,C_{10}(\mu)\right) + \hat{m}_b\,F_A(x_\gamma)\,\mathrm{Re}\left(C_7^*(\mu)\,F_{TV}^{\text{eff}\,*}(x_\gamma)\,C_{10}(\mu)\right)\right]\,, \tag{138}$$

$$\tilde{B}_2\left(x_\gamma\right) = \left(1 - x_\gamma\right)\left(F_1\left(x_\gamma\right) + F_2\left(x_\gamma\right)\right)\,, \tag{139}$$

where $\hat{m}_b = m_b/m_{B_s}$,

$$F_1\left(x_\gamma\right) = \left(\left|C_9^{\text{eff}}(\mu,x_\gamma)\right|^2 + |C_{10}(\mu)|^2\right)F_V^2(x_\gamma) + \left(\frac{2\hat{m}_b}{1 - x_\gamma}\right)^2\left|C_7(\mu)\,F_{TV}^{\text{eff}}(x_\gamma)\right|^2 \tag{140}$$

$$+ \frac{4\hat{m}_b}{1 - x_\gamma}\,F_V(x_\gamma)\,\mathrm{Re}\left(C_7(\mu)\,F_{TV}^{\text{eff}}(x_\gamma)\,C_9^{\text{eff}\,*}(\mu,x_\gamma)\right),$$

$$F_2\left(x_\gamma\right) = \left(\left|C_9^{\text{eff}}(x_\gamma,\mu)\right|^2 + |C_{10}(\mu)|^2\right)F_A^2(x_\gamma) + \left(\frac{2\hat{m}_b}{1 - x_\gamma}\right)^2\left|C_7(\mu)\,F_{TA}^{\text{eff}}(x_\gamma)\right|^2 \tag{141}$$

$$+ \frac{4\hat{m}_b}{1 - x_\gamma}\,F_A(x_\gamma)\,\mathrm{Re}\left(C_7(\mu)\,F_{TA}^{\text{eff}}(x_\gamma)\,C_9^{\text{eff}\,*}(\mu,x_\gamma)\right) \tag{142}$$

---

[13] In Ref. [4] the authors chose $\hat{s} = (p - k)^2/m_{B_s}^2 = 1 - x_\gamma$ and $\hat{t} = (p - p_1)^2/m_{B_s}^2$, where p_1 is the four-momentum of the μ^+, as the independent variables. We choose x_γ and $\cos\theta$ and $J(x_\gamma)$ is the Jacobian relating the two sets of variables.

[14] $\xi(x_\gamma,\hat{t})$ can also be written as $\hat{u} - \hat{t}$, where $\hat{u} = (p - p_2)^2/m_{B_s}^2$ and p_2 is the momentum of the μ^-.

$V_{c\bar{c}}$	$M_{V_{c\bar{c}}}$ [GeV]	Γ [MeV]	$\mathcal{B}(V_{c\bar{c}} \to \mu^+\mu^-)$
J/ψ	3.096900(6)	0.0926(17)	0.05961(33)
$\Psi(2S)$	3.68610(6)	0.294(8)	$8.0(6) \cdot 10^{-3}$
$\Psi(3770)$	3.7737(4)	27.2(1.0)	$^*9.6(7) \cdot 10^{-6}$
$\Psi(4040)$	4.039(1)	80(10)	$^*1.07(16) \cdot 10^{-5}$
$\Psi(4160)$	4.191(5)	70(10)	$^*6.9(3.3) \cdot 10^{-6}$
$\Psi(4230)$	4.2225(24)	48(8)	$3.2(2.9) \cdot 10^{-5}$
$\Psi(4415)$	4.421(4)	62(20)	$2(1) \cdot 10^{-5}$
$\Psi(4660)$	4.630(6)	72^{+14}_{-12}	not seen

TABLE V: Masses, decay widths and branching fractions into a $\mu^+\mu^-$ pair for the lowest-lying charmonium resonances $V_{c\bar{c}}$ [10]. For some of the charmonium resonances, in absence of information on the branching into $\mu^+\mu^-$, we provide the branching into e^+e^-, which is expected to provide a good approximation of $\mathcal{B}(V_{c\bar{c}} \to \mu^+\mu^-)$ given that $M_{V_{c\bar{c}}} \gg m_\mu, m_e$. In those cases where the branching into e^+e^- is given, the numerical value of the branching is preceded by an asterisk. For the last resonance in the table, the $\Psi(4660)$ resonance, neither the branching into e^+e^- or $\mu^+\mu^-$ has been measured, and we input the fiducial value $\mathcal{B}(\Psi(4660) \to \mu^+\mu^-) = 1(1) \cdot 10^{-5}$.

and

$$F_{TV}^{\text{eff}}(x_\gamma) = F_{TV}(x_\gamma) + \bar{F}_T(x_\gamma) \ , \tag{143}$$

$$F_{TA}^{\text{eff}}(x_\gamma) = F_{TA}(x_\gamma) + \bar{F}_T(x_\gamma) \ . \tag{144}$$

The interference and point-like contributions are given by

$$\frac{\mathrm{d}^2\Gamma^{(\text{INT})}}{\mathrm{d}x_\gamma\,\mathrm{d}(\cos\theta)} = -\frac{G_F^2\,\alpha_{\text{em}}^3\,m_{B_s}^5}{2^{10}\,\pi^4}\,|V_{tb}V_{ts}^*|^2\,\frac{16\,f_{B_s}}{m_{B_s}}\,\hat{m}_\mu^2 J(x_\gamma)\,\frac{x_\gamma^2}{(x_\gamma + \hat{m}_\mu^2 - \hat{t})(\hat{t} - \hat{m}_\mu^2)} \times$$

$$\left[\frac{2x_\gamma\hat{m}_b}{1 - x_\gamma}\,\text{Re}\left(C_{10}^*(\mu)C_7(\mu)F_{TV}^{\text{eff}}(x_\gamma)\right) + x_\gamma\,F_V(x_\gamma)\,\text{Re}\left(C_{10}^*(\mu)C_9^{\text{eff}}(\mu,x_\gamma)\right) + \xi(x_\gamma,\hat{t})\,F_A(x_\gamma)\,|C_{10}(\mu)|^2\right] \tag{145}$$

and

$$\frac{\mathrm{d}^2\Gamma^{(\text{PT})}}{\mathrm{d}x_\gamma\,\mathrm{d}(\cos\theta)} = \frac{G_F^2\,\alpha_{\text{em}}^3\,m_{B_s}^5}{2^{10}\,\pi^4}\,|V_{tb}V_{ts}^*|^2\,\left(\frac{8\,f_{B_s}}{m_{B_s}}\right)^2\,\hat{m}_\mu^2\,|C_{10}(\mu)|^2\,J(x_\gamma)\left[\frac{1 - x_\gamma + x_\gamma^2/2}{(x_\gamma + \hat{m}_\mu^2 - \hat{t})(\hat{t} - \hat{m}_\mu^2)} - \right.$$

$$\left.\left(\frac{x_\gamma\,\hat{m}_\mu}{(x_\gamma + \hat{m}_\mu^2 - \hat{t})\,(\hat{t} - \hat{m}_\mu^2)}\right)^2\right]\ . \tag{146}$$

In the following we will use the Wilson coefficients evaluated in the $\overline{\text{MS}}$ scheme at the scale $\mu = 5\,\text{GeV}$, which corresponds to the same scheme and scale at which we calculated the tensor form factors. In the calculation of the rate we input the value $m_b = m_b(5\,\text{GeV}) = 4.073(11)\,\text{GeV}$ obtained from $m_b(m_b) = 4.203(11)\,\text{GeV}$ [40] using the four-loop quark-mass anomalous dimension [41].

We now discuss our strategy for estimating in a conservative way the systematic error due to the charming-penguin diagram in Figure 4, corresponding to the emission of the $\mu^+\mu^-$ pair from the $c\bar{c}$ loop. As already discussed in Section II, this contribution can be written as a process and $q^2 = m_{B_s}^2(1 - x_\gamma)$ dependent shift of the Wilson coefficient $C_9 \to C_9^{\text{eff}}(q^2) = C_9 + \Delta C_9(q^2)$, and we rely on the phenomenological parameterization in Eq.(17) which we rewrite here for convenience,

$$\Delta C_9(q^2) = -\frac{9\pi}{\alpha_{\text{em}}^2}\left(C_1 + \frac{C_2}{3}\right)\sum_V |k_V|e^{i\delta_V}\frac{m_V\,B(V \to \mu^+\mu^-)\,\Gamma_V}{q^2 - m_V^2 + im_V\Gamma_V}\ . \tag{147}$$

The values of the parameters $m_V, \Gamma_V, B(V \to \mu^+\mu^-)$ are known experimentally for the lowest-lying resonances, and are collected in Table V. Instead, the value of the coefficients $|k_V|$ and the phases δ_V are largely unknown: $\delta_V = |k_V| - 1 = 0$ only holds in the factorization approximation. In order to estimate the systematic error induced in the parameterization of Eq. (147) by the poor knowledge of some of the parameters, we follow a (conservative) procedure similar to the
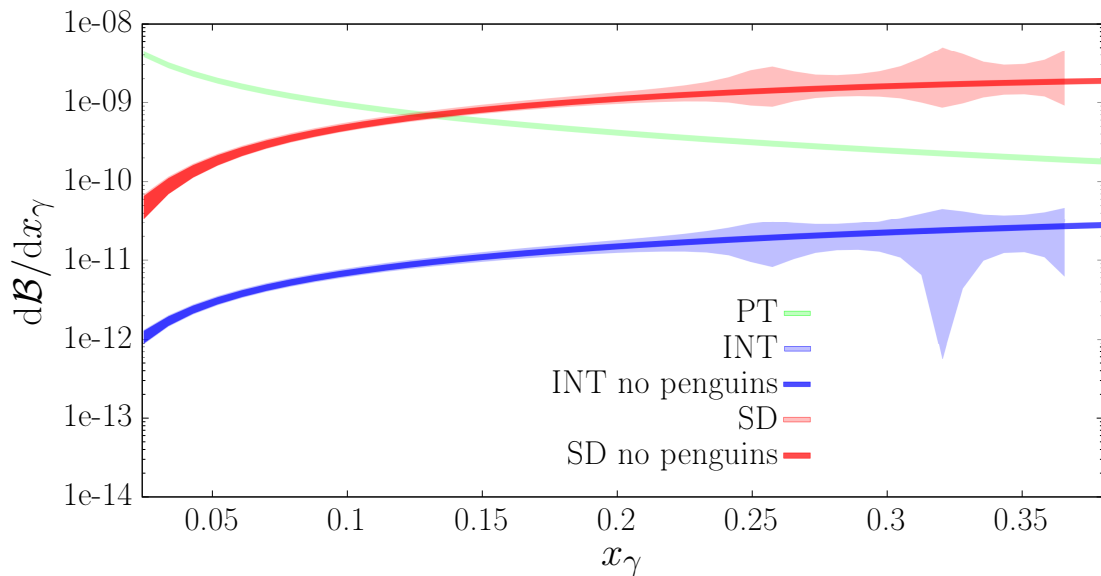
FIG. 21: *Our determination of the differential branching $\mathrm{d}\mathcal{B}/\mathrm{d}x_\gamma$ for $x_\gamma \in [0.025, 0.4]$. We give separately the point-like (light-green band), interference (light-blue band), and structure-dependent (light-red band) contributions. The blue and red bands correspond respectively to the determination of the interference and structure-dependent contribution obtained neglecting the charming-penguin diagrams.*

one adopted in Ref. [5]. We assume for each resonance the value $|k_V| = 1.75(0.75)^{15}$, and that the phases δ_V are completely unknown. Furthermore we assume that the resonance parameters are completely uncorrelated. To correctly propagate the uncertainty on $|k_V|$ and δ_V, as well as the one coming from all other input parameters (e.g. from the CKM matrix elements $|V_{tb}|$ and $|V_{ts}|$), we generate a large bootstrap sample of size $N_b = \mathcal{O}(1000)$ (and we assume that the parameters δ_V are uniformly distributed in the interval $[0, 2\pi)$), and repeat the calculation of the rate for each bootstrap value of the input parameters. Central values and standard errors are then obtained from the usual bootstrap average and dispersion formulae.

For the Wilson coefficients we take the values from Ref [9], which in the basis of operators which we use, correspond to

$$C_1(5\,\mathrm{GeV}) = 0.147 \ , \quad C_2(5\,\mathrm{GeV}) = -1.053 \ , \quad C_7(5\,\mathrm{GeV}) = 0.330 \ , \quad C_9(5\,\mathrm{GeV}) = -4.327 \ , \quad C_{10}(5\,\mathrm{GeV}) = 4.262 \ , \tag{148}$$

and for the remaining input parameters we take [10]

$$|V_{tb}| = 1.014(29) \ , \qquad |V_{ts}| = 4.15(9) \times 10^{-2} \ , \qquad \tau_{B_s} = 1.521(5) \times 10^{-12} \ \mathrm{s} \ , \tag{149}$$

where τ_{B_s} is the average between the lifetimes of the B_{sH} and B_{sL} mesons, which are the mass eigenstates of the $B_s - \bar{B}_s$ system. The Wilson coefficients computed in Ref. [9] include next-to-leading logarithm corrections. This has a particularly large relative effect on C_1 which is reduced by approximately 40% compared to the leading logarithmic result [43, 44], and subsequently on the magnitude of the combination $C_1 + C_2/3$ entering the charming-penguin parameterization in Eq. 147, which is increased by more than 60%. In the plot of Figure 21 we provide our determination of the differential branching fraction

$$\frac{\mathrm{d}\mathcal{B}}{\mathrm{d}x_\gamma} \equiv \tau_{B_s} \frac{\mathrm{d}\Gamma}{\mathrm{d}x_\gamma} \ , \tag{150}$$

as a function of $x_\gamma \in [0.025, 0.4]$. We give separately the point-like, interference, and structure-dependent contributions. As the figure shows the point-like contribution becomes subleading for $x_\gamma \gtrsim 0.15$, while the interference contribution turns out to be orders of magnitude smaller than the structure-dependent one on the entire range of x_γ explored. At

---

[15] It has been found [42] that $|k_V| \simeq 2.5$ well describes the $B \to K\mu^+\mu^-$ experimental data. The 1σ interval we choose for $|k_V|$ thus spans the region between $|k_V| = 2.5$ and $|k_V| = 1$ which corresponds to the value obtained in the factorization approximation.
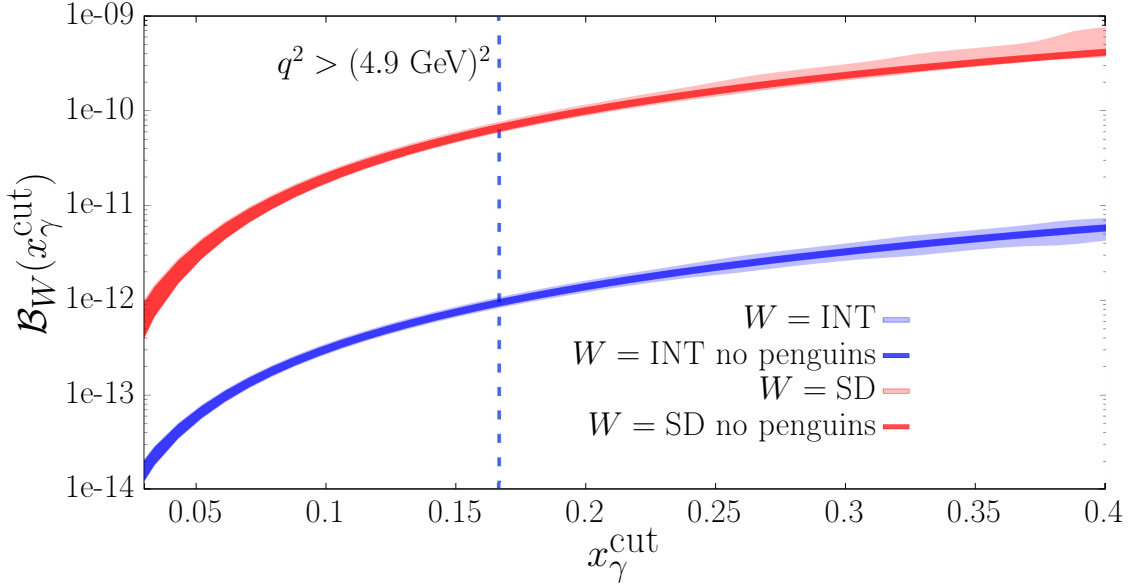
FIG. 22: *Our determination of the IR-finite structure-dependent ($\mathcal{B}_{\mathrm{SD}}(x_\gamma^{\mathrm{cut}})$, red band) and interference ($\mathcal{B}_{\mathrm{INT}}(x_\gamma^{\mathrm{cut}})$, blue band) contributions to the partial branching fractions $\mathcal{B}(x_\gamma^{\mathrm{cut}})$. The vertical blue line corresponds to the experimental cut imposed on the photon energy by the LHCb collaboration in Refs. [1, 2].*

large $x_\gamma \gtrsim 0.2$, the uncertainty stemming from the missing charming-penguin contributions is dominant over all other sources of uncertainties, and therefore in order to improve the precision of the differential branching at large x_γ a rigorous treatment of the charming penguin diagrams is necessary.

We now proceed to discuss the determination of the total branching fraction

$$\mathcal{B}(x_\gamma^{\mathrm{cut}}) = \int_0^{x_\gamma^{\mathrm{cut}}} \mathrm{d}x_\gamma \, \frac{\mathrm{d}\mathcal{B}}{\mathrm{d}x_\gamma} \,, \tag{151}$$

as a function of the upper bound $E_\gamma^{\mathrm{cut}} = m_{B_s} x_\gamma^{\mathrm{cut}}/2$ on the measured photon energy. As is well known, $\mathcal{B}(x_\gamma^{\mathrm{cut}})$ suffers from an infrared divergence generated by the point-like contribution to $\mathrm{d}\mathcal{B}/\mathrm{d}x_\gamma$ which at small x_γ behaves as $1/x_\gamma$. The infrared divergence appearing in the decay rate with a real photon in the final state is then cancelled by the $\mathcal{O}(\alpha_{\mathrm{em}})$ virtual photon contribution to the $B_s \to \mu^+\mu^-$ decay amplitude, through the usual Block-Nordsieck mechanism [45]. The interference ($\mathcal{B}_{\mathrm{INT}}(x_\gamma^{\mathrm{cut}})$) and structure-dependent ($\mathcal{B}_{\mathrm{SD}}(x_\gamma^{\mathrm{cut}})$) contributions are instead IR finite. In the experimental analysis made by the LHCb Collaboration in Refs. [1, 2] the point-like contribution (called the final-state-radiation (FSR) contribution in Refs. [1, 2]) has been included in the analysis of the $B_s \to \mu^+\mu^-$ invariant-mass distribution, as a radiative tail. For the IR-finite structure-dependent contribution (called the initial-state-radiation (ISR) contribution in Refs. [1, 2]) LHCb quotes the following upper-bound

$$\mathcal{B}_{\mathrm{SD}}^{\mathrm{LHCb}}(0.166) < 2 \times 10^{-9} \,. \tag{152}$$

In agreement with our results, the interference contribution has been instead considered negligible in Refs. [1, 2] on the basis of the results of Ref. [4] obtained using the relativistic dispersion approach. In Figure 22 we provide our determination of the IR-finite structure-dependent ($\mathcal{B}_{\mathrm{SD}}(x_\gamma^{\mathrm{cut}})$) and interference ($\mathcal{B}_{\mathrm{INT}}(x_\gamma^{\mathrm{cut}})$) contributions to $\mathcal{B}(x_\gamma^{\mathrm{cut}})$. The blue vertical line corresponds to the experimental cut $x_\gamma^{\mathrm{cut}} \simeq 0.166$ adopted in the experimental result of Eq. (152). For $x_\gamma^{\mathrm{cut}} = 0.166$ we obtain

$$\mathcal{B}_{\mathrm{SD}}(0.166) = 6.9(9) \times 10^{-11} \,, \tag{153}$$

while the interference contribution is completely negligible. Our result is well within the bound set by the LHCb Collaboration (Eq. (152)). In Table VI we collect our results for the sum $\mathcal{B}_{\mathrm{SD+INT}}(x_\gamma^{\mathrm{cut}}) \equiv \mathcal{B}_{\mathrm{SD}}(x_\gamma^{\mathrm{cut}}) + \mathcal{B}_{\mathrm{INT}}(x_\gamma^{\mathrm{cut}})$ of the interference and structure dependent contribution to the partial branching fraction, for different values of x_γ^{cut}.

We can further compare our results with the ones obtained using the model-dependent determination of the form factors F_V, F_A, F_{TV} and F_{TA} from Refs. [3–5]. The results of the comparison are shown in Figure 23. As the figure shows, our results for $\mathcal{B}_{\mathrm{INT+SD}}$ are smaller than those obtained using the form factors from Ref. [3], and larger than

	$\sqrt{q^2_{\rm cut}}$ [GeV] = $m_{B_s}\sqrt{1 - x^{\rm cut}_\gamma}$					
	4.1	4.2	4.3	4.4	4.5	4.6
$\mathcal{B}_{\rm SD+INT} \times 10^{10}$	6.1(2.1)	5.3(1.7)	3.99(88)	3.31(74)	2.57(50)	2.02(39)
	4.7	4.8	4.9	5.0	5.1	5.2
$\mathcal{B}_{\rm SD+INT} \times 10^{10}$	1.47(22)	1.04(14)	0.685(90)	0.399(55)	0.188(29)	0.057(12)

TABLE VI: *Our results for the partial branching fraction $\mathcal{B}_{\rm SD+INT}(x^{\rm cut}_\gamma) \equiv \mathcal{B}_{\rm SD}(x^{\rm cut}_\gamma) + \mathcal{B}_{\rm INT}(x^{\rm cut}_\gamma)$ for different values of $x^{\rm cut}_\gamma$. The interference contribution $\mathcal{B}_{\rm INT}(x^{\rm cut}_\gamma)$ is orders of magnitude smaller than $\mathcal{B}_{\rm SD}(x^{\rm cut}_\gamma)$ and completely negligible within uncertainties.*
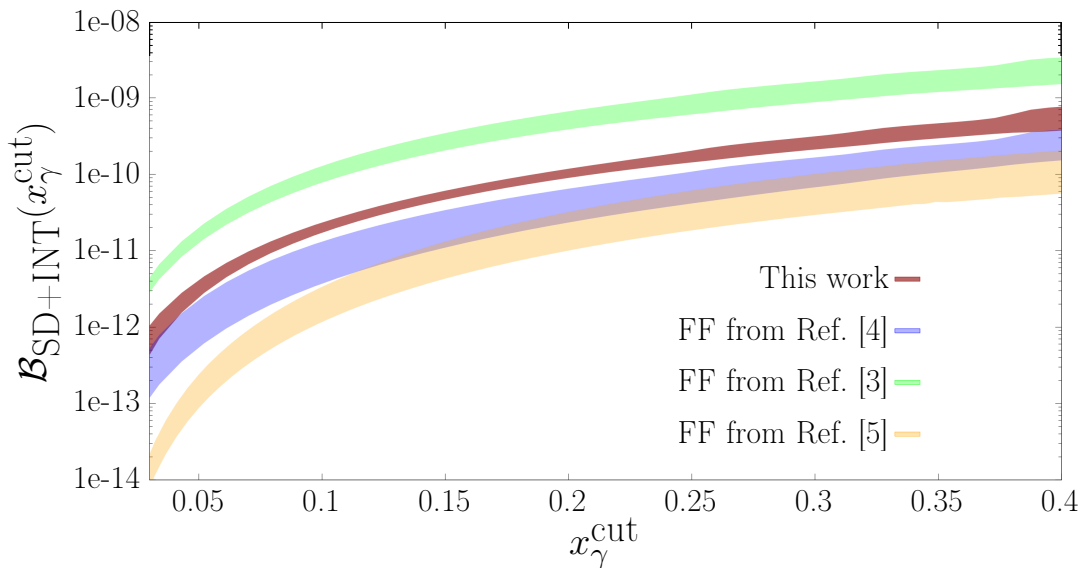
FIG. 23: *Comparison between our results for $\mathcal{B}_{\rm SD+INT}(x^{\rm cut}_\gamma)$ (shown in the figure by the red band), the ones obtained in Ref. [5] (shown by the orange band), and the ones we obtained using the form factors F_V, F_A, F_{TV} and F_{TA} from Ref. [3] (shown in the figure by the green band), and Ref. [4] (shown in the figure by the blue band). We used the estimate of $\bar{F}_T$ given in Ref. [4] to produce the results corresponding to the blue band, while the green band has been obtained setting $\bar{F}_T = 0$. The impact of $\bar{F}_T$ on the branching fraction is however extremely modest, and negligible within uncertainties.*

those obtained using the form factors from Refs. [4, 5] (w.r.t. Ref. [4] the difference is however less pronounced). This is not surprising given that the same trend is observed for the form factors (see Figure 11). Finally, we repeat that in order to obtain a more accurate theoretical prediction for $\mathcal{B}_{\rm SD}(x^{\rm cut}_\gamma)$ at large values of $x^{\rm cut}_\gamma$, a first-principles calculation of the charming-penguin contributions is needed, since our model-dependent estimate presently represents the main source of uncertainty for large $x^{\rm cut}_\gamma$.

VI. CONCLUSIONS AND FUTURE PERSPECTIVES

The rare radiative leptonic decay $B_s \to \mu^+\mu^-\gamma$ is a flavour-changing-neutral current transition which is forbidden at tree level in the Standard Model and is therefore particularly sensitive to potential New Physics contributions. Although there is an additional factor of $\alpha_{\rm em}$ in the amplitude for this process compared to that for the widely-studied $B_s \to \mu^+\mu^-$ decay, the presence of the final state photon removes the helicity suppression making the rates for the two processes comparable. On the other hand, while the leading hadronic effects in the $B_s \to \ell^+\ell^-$ ($\ell = e, \mu, \tau$) decay amplitude depend only on the B_s-meson decay constant f_{B_s}, which is known to sub-percent precision from lattice computations, the determination of the amplitude for the $B_s \to \mu^+\mu^-\gamma$ decay is much more complex. In this case the non-perturbative hadronic effects depend not only on local form factors, but also on resonance (including "charming penguin") and other long-distance contributions. In this paper, we have presented a first-principles calculation of the local form factors F_V, F_A, F_{TV}, F_{TA} and $\bar{F}_T$, which provide the main contributions to the amplitude for the $B_s \to \mu^+\mu^-\gamma$ decay at large di-muon invariant masses $\sqrt{q^2} > 4.16$ GeV, above the peaks of the lowest charmonium resonances. In order to

determine the amplitude, we combine our results for the form factors with previous phenomenological estimates of the remaining contributions, in particular those from charming penguins. Whilst we find that the dominant contribution to the differential branching fraction is given by the well-determined form factor F_V, the largest contribution to the uncertainty for $q^2 \lesssim (4.8\,\text{GeV})^2$ at present comes from the charming penguins.

The fitted results for the form factors are plotted as functions of x_γ in Figure 11, where they are also compared to earlier estimates obtained using different techniques [3–5]. It can be seen that, with a few exceptions, our results for the form factors differ significantly from the earlier estimates (which also differ from each other). In particular our results for the form factor F_V, which gives the largest contribution to the amplitude, are significantly smaller than that obtained in Ref. [3] and larger than those in Refs. [4, 5].

In evaluating $\bar{F}_T^s$, the contribution to the form factor $\bar{F}_T$ in which the virtual photon is emitted from the strange anti-quark, one encounters the difficulty of performing the analytic continuation to Euclidean spacetime due to the presence of intermediate vector $s\bar{s}$ states with masses below $\sqrt{q^2}$. As explained in detail in Sec. IV, in order to overcome this problem, we have employed the novel spectral-density reconstruction technique developed in Ref. [8]. Since the contribution of $\bar{F}_T$ to the differential rate is small, and in view of the computational expense of implementing the spectral representation technique, we have evaluated it at the same four values of x_γ as the four other form factors, but only on two ensembles and at three values of the heavy-quark mass ($m_h/m_c = 1, 1.5, 2.5$). We do not observe any significant dependence on x_γ in $\bar{F}_T^s(x_\gamma)$ and present our results for its real and imaginary parts in Eq. (131). There is no difficulty in the continuation to Euclidean space for $\bar{F}_T^b$, the contribution to the form factor $\bar{F}_T$ in which the virtual photon is emitted from the b-quark, and we find that $\bar{F}_T^b$ is an order of magnitude smaller than F_{TV} and F_{TA}. The results at the four values $x_\gamma = 0.1, 0.2, 0.3$ and 0.4 are presented in Table IV.

We use our results for the local form factors to evaluate the $B_s \to \mu^+\mu^-\gamma$ amplitude for $q^2 > (4.16\,\text{GeV})^2$, taking into account the systematic uncertainties due to the contributions that we have not computed in the present work, in particular those from the charming-penguin diagrams. We present our results for the partial branching fractions as a function of the upper cut-off on x_γ (or equivalently on the lower cut-off on $\sqrt{q^2}$) in Table VI. Imposing the same cut on the photon energy ($q^2 > (4.9\,\text{GeV})^2$, i.e. $x_\gamma < 0.166$) as adopted by the LHCb Collaboration, we obtain a value for the structure-dependent contribution to the branching fraction $\mathcal{B}_{\text{SD}}(0.166) = 7.0(8) \times 10^{-11}$, which is well within the bound set by the LHCb collaboration $\mathcal{B}_{\text{SD}}(0.166) < 2.0 \times 10^{-9}$ [1, 2]. However, as illustrated in Figures 11 and 23, our results disagree with the LCSR and model/effective-theory determinations of the branching fractions from Refs. [3–5]; in particular they are smaller than the result in Ref. [3] and larger than those in Refs. [4, 5]. The difference can be traced back to the fact that our result for the form factor F_V, which is the dominant contribution to the rate, is larger (smaller) than those obtained in Refs. [4, 5] ([3]) by about a factor of $1.5 - 2$.

At present our results for the branching fractions have uncertainties ranging from $\mathcal{O}(15\%)$ for $\sqrt{q_{\text{cut}}^2} = 4.9$ GeV, to $\mathcal{O}(30\%)$ for $\sqrt{q_{\text{cut}}^2} = 4.2$ GeV. Our uncertainties should already be at the level of precision that can be obtained in the future experimental measurements of $\mathcal{B}(B_s \to \mu^+\mu^-\gamma)$ at LHCb. Our analysis shows that in order to further improve the accuracy of the theoretical predictions in the low-q^2 region, it is necessary to obtain a first-principles determination of the (currently missing) charming-penguin contributions, which presently constitute the main source of uncertainty in the differential branching fraction for $q^2 \lesssim (4.8\,\text{GeV})^2$. This can be seen as the difference between the light-red band in Figure 21, which is our full result, and the dark-red curve in which the charming penguin contributions have been neglected.

VII. ACKNOWLEDGEMENTS

[1] LHCb collaboration, *Measurement of the $B_s^0 \to \mu^+\mu^-$ decay properties and search for the $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-\gamma$ decays*, *Phys. Rev. D* **105** (2022) 012010 [2108.09283].

[2] LHCb collaboration, *Analysis of Neutral B-Meson Decays into Two Muons*, *Phys. Rev. Lett.* **128** (2022) 041801 [2108.09284].

[3] T. Janowski, B. Pullin and R. Zwicky, *Charged and neutral $\overline{B}_{u,d,s} \to \gamma$ form factors from light cone sum rules at NLO*, *JHEP* **12** (2021) 008 [2106.13616].

[4] A. Kozachuk, D. Melikhov and N. Nikitin, *Rare FCNC radiative leptonic $B_{s,d} \to \gamma l^+ l^-$ decays in the standard model*, *Phys. Rev. D* **97** (2018) 053007 [1712.07926].

[5] D. Guadagnoli, C. Normand, S. Simula and L. Vittorio, *From $D_s \to \gamma$ in lattice QCD to $B_s \to \mu\mu\gamma$ at high q^2*, *JHEP* **07** (2023) 112 [2303.02174].

[6] D. Guadagnoli, M. Reboud and R. Zwicky, *$B_s^0 \to \ell^+ \ell^-\gamma$ as a test of lepton flavor universality*, *JHEP* **11** (2017) 184 [1708.02649].

[7] R. Frezzotti, N. Tantalo, G. Gagliardi, F. Sanfilippo, S. Simula, V. Lubicz et al., *Lattice calculation of the Ds meson radiative form factors over the full kinematical range*, *Phys. Rev. D* **108** (2023) 074505 [2306.05904].

[8] R. Frezzotti, N. Tantalo, G. Gagliardi, F. Sanfilippo, S. Simula and V. Lubicz, *Spectral-function determination of complex electroweak amplitudes with lattice QCD*, *Phys. Rev. D* **108** (2023) 074510 [2306.07228].

[9] M. Beneke, C. Bobeth and Y.-M. Wang, *$B_{d,s} \to \gamma\ell\bar{\ell}$ decay with an energetic photon*, *JHEP* **12** (2020) 148 [2008.12494].

[10] Particle Data Group collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.

[11] F. Kruger and L.M. Sehgal, *Lepton polarization in the decays $b \longrightarrow X(s)$ mu+ mu- and $B \longrightarrow X(s)$ tau+ tau-*, *Phys. Lett. B* **380** (1996) 199 [hep-ph/9603237].

[12] A. Desiderio et al., *First lattice calculation of radiative leptonic decay rates of pseudoscalar mesons*, *Phys. Rev. D* **103** (2021) 014502 [2006.05358].

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

[14] J.M. Flynn, A. Juttner, C.T. Sachrajda, P.A. Boyle and J.M. Zanotti, *Hadronic form factors in Lattice QCD at small and vanishing momentum transfer*, *JHEP* **05** (2007) 016 [hep-lat/0703005].

[15] 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].

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

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

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

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

[20] Extended Twisted Mass collaboration, *Quark masses using twisted-mass fermion gauge ensembles*, *Phys. Rev. D* **104** (2021) 074515 [2104.13408].

[21] Extended Twisted Mass collaboration, *Ratio of kaon and pion leptonic decay constants with Nf=2+1+1 Wilson-clover twisted-mass fermions*, *Phys. Rev. D* **104** (2021) 074520 [2104.06747].

[22] Extended Twisted Mass collaboration, *Lattice calculation of the short and intermediate time-distance hadronic vacuum polarization contributions to the muon magnetic moment using twisted-mass fermions*, *Phys. Rev. D* **107** (2023) 074506 [2206.15084].

[23] C. Alexandrou et al., *Simulating twisted mass fermions at physical light, strange and charm quark masses*, *Phys. Rev. D* **98** (2018) 054518 [1807.00495].

[24] Extended Twisted Mass collaboration, *Non-perturbative renormalisation of quark bilinear operators with $N_f = 2 + 1 + 1$ Wilson-clover twisted-mass fermions*, *in preparation*. We are grateful to the authors for providing us with the preliminary values of the renormalization constant Z_T.

[25] H. Akaike, *A new look at the statistical model identification*, *IEEE Transactions on Automatic Control* **19** (1974) 716.

[26] G. Parisi, *The Strategy for Computing the Hadronic Mass Spectrum*, *Phys. Rept.* **103** (1984) 203.

[27] G.P. Lepage, *The Analysis of Algorithms for Lattice Field Theory*, in *Theoretical Advanced Study Institute in Elementary Particle Physics*, 6, 1989.

[28] M. Beneke and J. Rohrwild, *B meson distribution amplitude from $B \to \gamma l \nu$*, *Eur. Phys. J. C* **71** (2011) 1818 [1110.3228].

[29] G.P. Korchemsky, D. Pirjol and T.-M. Yan, *Radiative leptonic decays of B mesons in QCD*, *Phys. Rev. D* **61** (2000) 114510 [hep-ph/9911427].

[30] S. Descotes-Genon and C.T. Sachrajda, *Factorization, the light cone distribution amplitude of the B meson and the radiative decay $B \longrightarrow$ gamma l nu(l)*, *Nucl. Phys. B* **650** (2003) 356 [hep-ph/0209216].

[31] E. Lunghi, D. Pirjol and D. Wyler, *Factorization in leptonic radiative $B \to \gamma e\nu$ decays*, *Nucl. Phys. B* **649** (2003) 349 [`hep-ph/0210091`].

[32] S.W. Bosch, R.J. Hill, B.O. Lange and M. Neubert, *Factorization and Sudakov resummation in leptonic radiative B decay*, *Phys. Rev. D* **67** (2003) 094014 [`hep-ph/0301123`].

[33] M. Beneke, V.M. Braun, Y. Ji and Y.-B. Wei, *Radiative leptonic decay $B \to \gamma \ell \nu_\ell$ with subleading power corrections*, *JHEP* **07** (2018) 154 [`1804.04962`].

[34] M. Neubert, *Heavy quark effective theory*, *Subnucl. Ser.* **34** (1997) 98 [`hep-ph/9610266`].

[35] G.C. Donald, C.T.H. Davies, J. Koponen and G.P. Lepage, *Prediction of the D_s^* width from a calculation of its radiative decay in full lattice QCD*, *Phys. Rev. Lett.* **112** (2014) 212002 [`1312.5264`].

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

[37] Extended Twisted Mass Collaboration (ETMC) collaboration, *Probing the Energy-Smeared R Ratio Using Lattice QCD*, *Phys. Rev. Lett.* **130** (2023) 241901 [`2212.08467`].

[38] Extended Twisted Mass collaboration, *Inclusive hadronic decay rate of the τ lepton from lattice QCD*, *Phys. Rev. D* **108** (2023) 074513 [`2308.03125`].

[39] C. Bonanno, F. D'Angelo, M. D'Elia, L. Maio and M. Naviglio, *Sphaleron rate of $N_f = 2 + 1$ QCD*, `2308.01287`.

[40] Flavour Lattice Averaging Group (FLAG) collaboration, *FLAG Review 2021*, *Eur. Phys. J. C* **82** (2022) 869 [`2111.09849`].

[41] K.G. Chetyrkin, *Quark mass anomalous dimension to O (alpha-s**4)*, *Phys. Lett. B* **404** (1997) 161 [`hep-ph/9703278`].

[42] J. Lyon and R. Zwicky, *Resonances gone topsy turvy - the charm of QCD or new physics in $b \to s\ell^+\ell^-$?*, `1406.0566`.

[43] A.J. Buras and M. Munz, *Effective Hamiltonian for B —> X(s) e+ e- beyond leading logarithms in the NDR and HV schemes*, *Phys. Rev. D* **52** (1995) 186 [`hep-ph/9501281`].

[44] G. Buchalla, A.J. Buras and M.E. Lautenbacher, *Weak decays beyond leading logarithms*, *Rev. Mod. Phys.* **68** (1996) 1125 [`hep-ph/9512380`].

[45] F. Bloch and A. Nordsieck, *Note on the radiation field of the electron*, *Phys. Rev.* **52** (1937) 54.

[46] European Twisted Mass collaboration, *Up, down, strange and charm quark masses with $N_f = 2+1+1$ twisted mass lattice QCD*, *Nucl. Phys. B* **887** (2014) 19 [`1403.4504`].

[47] K.G. Chetyrkin and A.G. Grozin, *Three loop anomalous dimension of the heavy light quark current in HQET*, *Nucl. Phys. B* **666** (2003) 289 [`hep-ph/0303113`].

[48] D.J. Broadhurst and A.G. Grozin, *Matching QCD and HQET heavy - light currents at two loops and beyond*, *Phys. Rev. D* **52** (1995) 4082 [`hep-ph/9410240`].

[49] TUMQCD collaboration, *Relations between Heavy-light Meson and Quark Masses*, *Phys. Rev. D* **97** (2018) 034503 [`1712.04983`].

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

Appendix A: Determination of the decay constant f_{B_s}

In this appendix we discuss our determination of the decay constant of the B_s meson, which enters in the extrapolation formulae of Section III C. For this calculation, in order to account for any correlations in the determination of the decay constants and form factors, we use the same configurations and masses as in the determination of the form factors F_W, $W = \{V, A, TV, TA\}$ discussed in Section III A.

For each simulated value of the heavy-light meson mass m_{H_s}, and for each lattice spacing, we use two different estimators of the decay constant f_{H_s}. The first determination of f_{H_s} is obtained from the pseudoscalar-pseudoscalar two-point correlation functions

$$C_{\text{PS}}^{\text{sm-loc}}(t) \equiv \sum_{\vec{x}} \langle P_5^{\text{loc}}(t, \vec{x}) P_5^{\dagger \text{sm}}(0) \rangle, \qquad C_{\text{PS}}^{\text{sm-sm}}(t) \equiv \sum_{\vec{x}} \langle P_5^{\text{sm}}(t, \vec{x}) P_5^{\dagger \text{sm}}(0) \rangle , \tag{A1}$$

where the labels sm and loc indicate "smeared" and "local" respectively,

$$P_5^{\text{loc}}(x) = \bar{s}(x)\gamma^5 h(x) , \qquad P_5^{\text{sm}}(x) = \bar{s}(x)\gamma^5 H_k^N(x, y) h(y) , \tag{A2}$$

and where

$$H_k(x, y) = \frac{1}{1 + 6k} \left(1 + k \sum_{\hat{j}} U_j(x)\delta(x + a\hat{j}, y) \right) \tag{A3}$$

is the gauge-field-dependent Gaussian-smearing operator that has been used to construct the interpolating operator of the $\bar{H}_s$ meson in the Euclidean three-point correlation function $B_W^{\mu\nu}(t; k, p)$ in Eq. (21). The Wilson parameter of the

$\bar{s}$ and h valence quarks entering the bilinears are always chosen to be opposite. In the large time limit $t \gg a$, $t \ll T$, the correlation functions $C_{\text{PS}}^{\text{sm-loc}}(t)$ and $C_{\text{PS}}^{\text{sm-sm}}(t)$ behave as

$$C_{\text{PS}}^{\text{sm-loc}}(t) = \frac{|Z_{\text{PS}}^{\text{sm-loc}}|^2}{2m_{H_s}} \left(e^{-m_{H_s}t} + e^{-m_{H_s}(T-t)} \right) + \dots ,$$

$$C_{\text{PS}}^{\text{sm-sm}}(t) = \frac{|Z_{\text{PS}}^{\text{sm-sm}}|^2}{2m_{H_s}} \left(e^{-m_{H_s}t} + e^{-m_{H_s}(T-t)} \right) + \dots , \tag{A4}$$

where the ellipsis indicate terms that are subleading in the limit of large time separations. From the knowledge of $Z_{\text{PS}}^{\text{sm-loc}}$ and $Z_{\text{PS}}^{\text{sm-sm}}$ it is possible to determine the decay constant f_{H_s}, without the need of any renormalization constant, making use of [46]

$$Z_{\text{PS}} \equiv \langle \bar{H}_s | \bar{s}\gamma^5 h | 0 \rangle = \frac{|Z_{\text{PS}}^{\text{sm-loc}}|^2}{|Z_{\text{PS}}^{\text{sm-sm}}|} , \qquad f_{H_s} = a(m_s^{\text{bare}} + m_h^{\text{bare}}) \frac{|Z_{\text{PS}}|}{m_{H_s} \sinh(am_{H_s})} , \tag{A5}$$

where m_s^{bare} and m_h^{bare} are the simulated values of the bare strange and heavy quark mass. The use of smeared interpolating operators is essential in order to improve the signal for f_{H_s}, as it allows the ground state contribution to be isolated at much smaller times (as compared to the standard local interpolating operator), where the correlation function is generally more precise.

The second estimator of f_{H_s} that we use is obtained exploiting the fact that in the zero photon-momentum limit $\boldsymbol{k} = 0$, the spatial part $H_A^{ii}(\mathbf{0}, \mathbf{0})$ of the axial hadronic tensor is equal to if_{H_s} if we choose fictitious values q'_s and q'_h for the electric charges of the strange and heavy quark entering J_{em}^μ in Eq. (22), in such a way that $q'_s + q'_h = 1$ (see Appendix C of Refs. [12] for more details on this point). In the following we denote the determination of f_{H_s} from Eqs. (A1)-(A5) as $f_{H_s}^{\text{2pt}}$, and the one from $H_A^{ii}(\mathbf{0},\mathbf{0})$ as $f_{H_s}^{\text{3pt}}$. The two estimates of the the decay constants only differ by cut-off effects and allow us to better constraint the result of the continuum fits. Our strategy is to perform the continuum limit extrapolation at fixed values of m_{H_s}, fitting simultaneously $f_{H_s}^{\text{2pt}}$ and $f_{H_s}^{\text{3pt}}$ using the following Ansatz

$$\phi_{H_s}^{\text{2pt}} \equiv f_{H_s}^{\text{2pt}} \sqrt{m_{H_s}} = A + B^{\text{2pt}}a^2 + D^{\text{2pt}}a^4, \qquad \phi_{H_s}^{\text{3pt}} \equiv f_{H_s}^{\text{3pt}} \sqrt{m_{H_s}} = A + B^{\text{3pt}}a^2 + D^{\text{3pt}}a^4 , \tag{A6}$$

where $A, B^{\text{2pt/3pt}}$ and $D^{\text{2pt/3pt}}$ are free fit parameters. Note that a common continuum value is enforced. In order to avoid overfitting, given the limited number of gauge ensembles that we employ, we do not include fits with both D^{2pt} and D^{3pt} as fit parameters, i.e. we set either D^{2pt} or D^{3pt} or both to zero. The fits are performed minimizing a correlated χ^2-function which takes into account the correlation between $\phi_{H_s}^{\text{2pt}}$ and $\phi_{H_s}^{\text{3pt}}$ evaluated on the same gauge ensemble. The results of the extrapolation to the continuum limit for the five simulated values of m_{H_s} are illustrated in Figure 24. We have performed a total of six different fits for each m_{H_s}, which differ depending on whether the ensemble with the coarsest lattice spacing is included or not, and on whether we include or neglect the terms proportional to a^4, i.e. whether we set D^{2pt} and/or D^{3pt} to zero. The results obtained from the different fits are then combined using the AIC, which has already been discussed in Section III A. To be conservative, we add linearly the systematic and statistical errors from AIC. Having extrapolated the decay constants f_{H_s} for each of the five simulated values of m_{H_s} to the continuum limit, in order to obtain f_{B_s} we need to perform the extrapolation in the mass. To do so, we make use of the heavy-quark scaling relation

$$\phi(m_{H_s}) = \text{const.} \times \left(1 + \frac{B}{m_{H_s}} + \mathcal{O}(m_{H_s}^{-2}) \right) . \tag{A7}$$

As is well known, Eq. (A7) is valid exactly only in the effective theory at the bare level. Logarithmic corrections to Eq. (A7) are generated by the non-zero anomalous dimension of the axial current in the HQET, as well as from its matching to the axial current in full QCD. Let $J_\Gamma(\mu') = Z_J^{-1}(\alpha_s(\mu))\bar{\ell}\Gamma h$ b, where Γ is one of the Dirac matrices and ℓ and h denote light and heavy quark fields in QCD, be a generic renormalized heavy-light current in QCD and let $\tilde{J}_\Gamma(\mu) = Z_{\tilde{J}}^{-1}(\alpha_s(\mu'))\bar{\ell}\Gamma h_v$ be its counterpart in the HQET. The relation between the two currents is given, at leading order in the heavy-quark mass m_h, by [47]

$$J_\Gamma(\mu') = C_\Gamma(m_h, m_h) \exp \left\{ \int_{\alpha_s(m_h)}^{\alpha_s(\mu')} \frac{\gamma_J(\alpha_s)}{2\beta(\alpha_s)} \frac{d\alpha_s}{\alpha_s} - \int_{\alpha_s(m_h)}^{\alpha_s(\mu')} \frac{\gamma_{\tilde{J}}(\alpha_s)}{2\beta(\alpha_s)} \frac{d\alpha_s}{\alpha_s} \right\} \tilde{J}_\Gamma(\mu') , \tag{A8}$$

where $\beta(\alpha_s)$ is the QCD β-function, and $\gamma_{J(\tilde{J})}$ is the anomalous dimension of the current $J(\tilde{J})$. In the case of the QCD axial current one has $\gamma_J = 0$. $C_\Gamma(m_h, m_h)$ is the matching coefficient and is obtained by imposing the equality
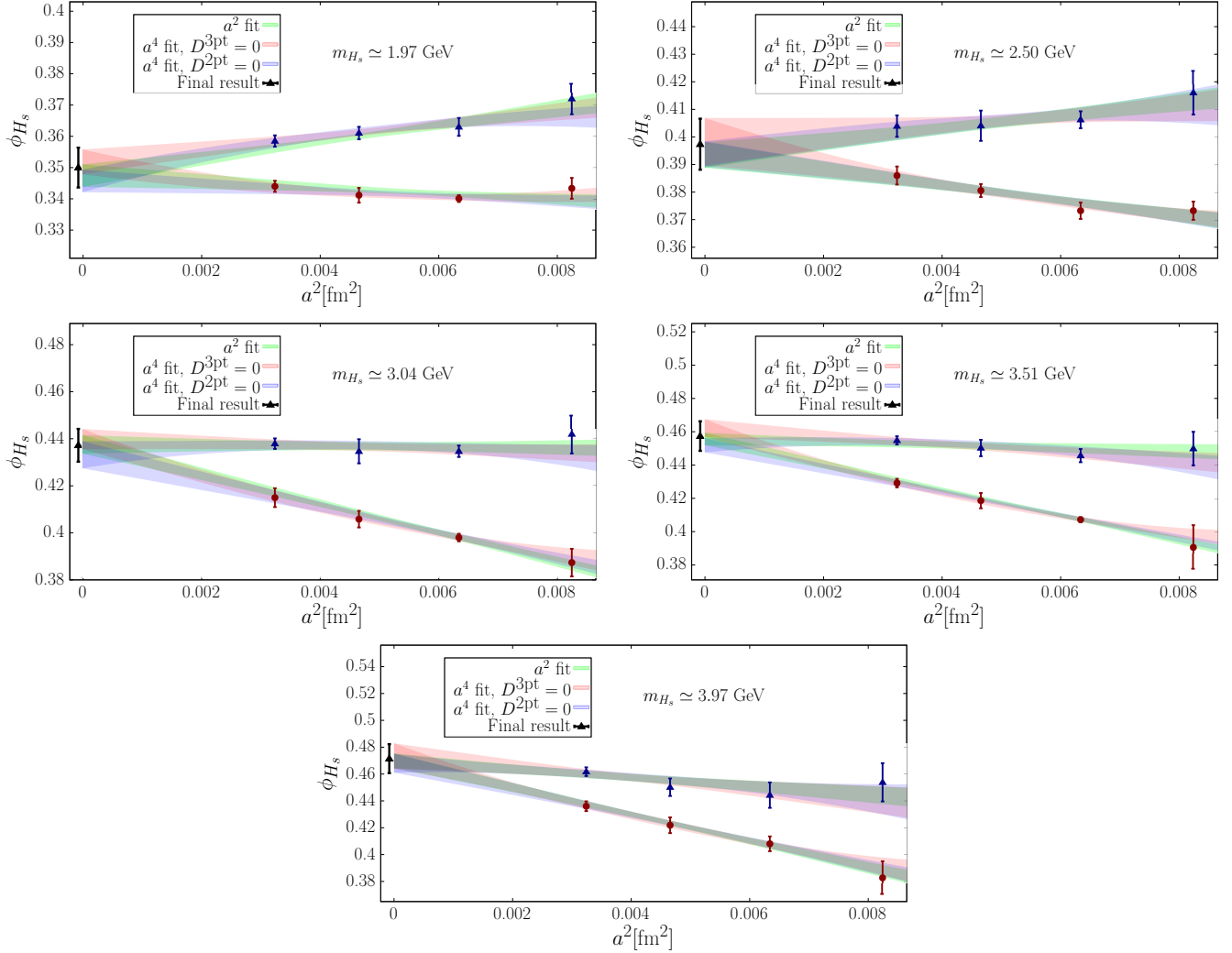
FIG. 24: *Continuum limit extrapolation of ϕ_{H_s} for the five different simulated values of the heavy-strange meson mass m_{H_s}. For each mass we show the continuum limit fits obtained employing all four lattice spacings. The blue and red data points correspond respectively to the estimator ϕ^{2pt} and ϕ^{3pt}. The green, red, and blue bands correspond respectively to the pure a^2 fit, and to the a^4 fit with $D^{3pt} = 0$ and $D^{2pt} = 0$. Finally, the black data point at $a^2 \simeq 0$ correspond to our final determination, which is obtained combining the six different fits we performed using the AIC.*

between the renormalized proper vertices of the quark-bilinear in question, evaluated in QCD and in the HQET at the heavy quark scale m_h. Its two-loop expression for the axial current has been computed in Ref. [48] and is given by

$$C_{\gamma^0\gamma^5}(m_h, m_h) = 1 - \frac{2}{3}\frac{\alpha_s(m_h)}{\pi} - 2.95\left(\frac{\alpha_s(m_h)}{\pi}\right)^2 . \qquad (A9)$$

The anomalous dimension $\gamma_{\tilde{J}}$ in the HQET (which does not depend on the specific Γ considered) has been computed at three-loops in Ref. [47]. In summary, the relation in Eq. (A7) gets modified by the matching and by the running of $\tilde{J}_\Gamma(\mu')$ into

$$\phi(m_{H_s}) = C_{\gamma^0\gamma^5}(m_h, m_h) \exp\left\{\int_0^{\alpha_s(m_h)} \frac{\gamma_{\tilde{J}}(\alpha_s)}{2\beta(\alpha_s)}\frac{d\alpha_s}{\alpha_s}\right\} \times \text{const}' \times \left(1 + \frac{B}{m_{H_s}} + \mathcal{O}(m_{H_s}^{-2})\right) , \qquad (A10)$$

and we have reabsorbed the μ' dependence in Eq. (A8) into the constant factor, const'. A delicate point of the analysis is the determination of the heavy quark mass m_h of the HQET, which should be identified with the pole mass. Notoriously, the pole mass is affected by renormalon ambiguities and its perturbative expansion in terms of the $\overline{\text{MS}}$ mass $m_h^{\overline{\text{MS}}}(m_h^{\overline{\text{MS}}})$ is asymptotically divergent. To avoid the use of the pole mass, we follow two different strategies and use, in
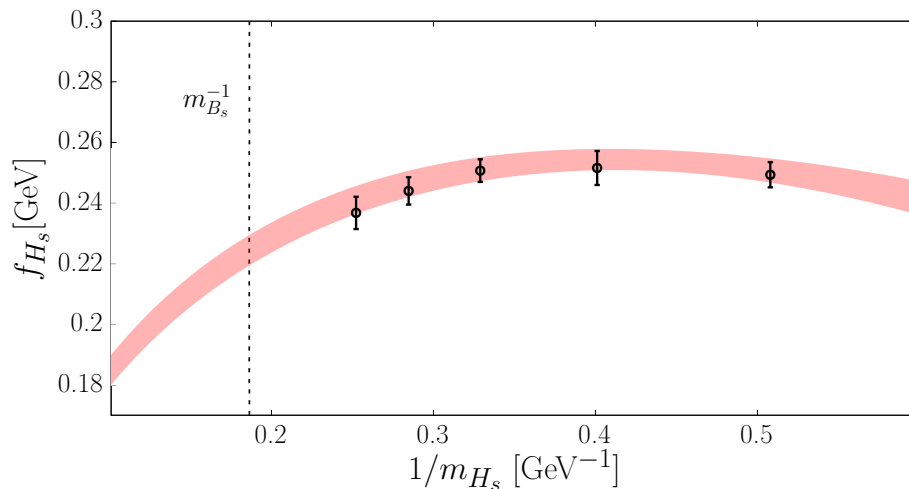
FIG. 25: *Extrapolation of the decay constant f_{H_s} to the physical B_s meson mass $m_{B_s} \simeq 5.367$ GeV using the Ansatz in Eq. (A10) and with m_h replaced by m_{H_s}. The red curve corresponds to the best-fit function, while the dashed vertical line corresponds to the inverse mass of the B_s meson. The reduced χ^2 of the fit is smaller than unit. Our final determination of f_{B_s} is given in Eq. (A11).*

place of m_h, either the heavy-strange meson mass m_{H_s} ($m_{H_s} - m_h \simeq \mathcal{O}(\Lambda_{QCD})$), or the minimal-renormalon-subtracted mass advocated in Ref. [49] (see also Ref. [50] for another alternative to the use of the pole mass).

The result of the extrapolation obtained replacing m_h with m_{H_s} in Eq. (A10), and with the constant term and the B coefficient considered as free fit parameters, is shown in Figure 25. The fit takes fully into account the correlation between the values of the decay constant obtained at the different values of m_{H_s}. The resulting χ^2/dof of the fit is very good and smaller than unity. By repeating the fit employing the minimal-renormalon-subtracted mass, we find that the value of f_{B_s} changes by less than 0.3σ, and therefore we do not add any additional systematic error. We obtain the value

$$f_{B_s} = 224.5\,(5.0)\,\mathrm{MeV}\,, \tag{A11}$$

which agrees with the $N_f = 2 + 1 + 1$ FLAG average $f_{B_s}^{\mathrm{FLAG}} = 230.3\,(1.3)$ MeV at the level of 1.1σ. Our final uncertainty is however much larger, which reflects the fact that our calculation is not tailored to be a precise determination of f_{B_s}. Our main interest is in the calculation of the local form factors contributing to the $B_s \to \mu^+\mu^-\gamma$ decay amplitude. The main limitation preventing a more precise determination of f_{B_s} comes from the systematic uncertainty associated to the continuum limit extrapolation, performed here with a rather limited number of lattice spacings. The new gauge ensembles that the Extended Twisted Mass (ETM) Collaboration will produce at smaller values of the lattice spacing, will allow for significantly reduced uncertainties on f_{B_s}.