Semi-inclusive b→s$\bar{\ell}\ell$ transitions at high q^2

Isidori, Gino ; Polonsky, Zachary ; Tinari, Arianna

DOI: https://doi.org/10.1103/physrevd.108.093008

Semi-inclusive $b \to s\bar{\ell}\ell$ transitions at high q^2

Gino Isidori, Zachary Polonsky, and Arianna Tinari
Physik-Institut, Universität Zürich, CH-8057 Zürich, Switzerland

(Received 2 June 2023; accepted 28 October 2023; published 21 November 2023)

We present an updated Standard Model (SM) estimate of the inclusive $b \to s\bar{\ell}\ell$ rate at a high dilepton invariant mass ($q^2 \geq 15$ GeV^2). We show that this estimate is in good agreement with the result obtained summing the SM predictions for the leading one-body modes (K and K^*) and the subleading nonresonant $K\pi$ channel (for which we also present an updated estimate). On the contrary, the semi-inclusive sum based on data exhibits a deficit compared to the inclusive SM prediction in the muon modes. The statistical significance of this deficit does not exceed 2σ but is free from uncertainties on hadronic form factors and fully compatible with the deficit observed at low-q^2 on the exclusive modes. The implications of these results in conjunction with other SM tests on $b \to s\bar{\mu}\mu$ modes are briefly discussed.

DOI: 10.1103/PhysRevD.108.093008

I. INTRODUCTION

The ultimate goal of studying $b \to s\bar{\ell}\ell$ decays is to probe the short-distance structure of the corresponding flavor-changing neutral-current (FCNC) amplitudes. By doing so, we perform precise tests of the Standard Model (SM) probing, at the same time, motivated beyond-the-SM (BSM) theories. The presence of narrow charmonium resonances poses challenges in extracting short-distance information for both exclusive and inclusive $b \to s\bar{\ell}\ell$ decays if the invariant mass of the dilepton pair, $q^2 = (p_{\bar{\ell}} + p_{\ell})^2$, is close to the resonance masses. This is why precise SM tests are confined to $q^2 \lesssim 6$–8 GeV^2 (low-q^2 region) and $q^2 \gtrsim 14$–15 GeV^2 (high-q^2 region). It is important to study both these regions as they are sensitive to different short-distance physics and, most importantly, they experience a different interplay between short- and long-distance dynamics. For a similar reason, it is important to study $b \to s\bar{\ell}\ell$ transitions both at the exclusive and inclusive levels.

In the last few years, measurements of rates and angular distributions of the exclusive $B \to K^{(*)}\bar{\mu}\mu$ decays by LHCb [1–3] have shown significant tensions with the corresponding SM predictions, especially in the low-q^2 region (see, e.g., Refs. [4–9] for recent analyses). All the attempts to compute the decay amplitudes from QCD agree on the observed tension. However, using a more agnostic data-driven approach, some doubts about the reliability of the theory errors have been raised in Refs. [10,11]. The goal of this paper is to attempt to shed light on this issue by looking at the inclusive B-meson decay rate, $\Gamma(B \to X_s\bar{\ell}\ell)$, in the high-$q^2$ region. This observable provides complementary information on $b \to s\bar{\ell}\ell$ amplitudes, being affected by qualitatively different uncertainties with respect to those appearing in the exclusive modes in the low-q^2 region.

The heavy-quark expansion in the high-q^2 region is an expansion in $O(\Lambda_{\rm QCD}/(m_b - \sqrt{q^2}))$ [12], which converges less rapidly with respect to the $O(\Lambda_{\rm QCD}/m_b)$ expansion at work in the low-q^2 region (see Refs. [13–15]). However, as pointed out by Ligeti and Tackmann [16], nonperturbative uncertainties in the high-q^2 region can be greatly reduced by computing the ratio of the FCNC transition and the $b \to u$ charged-current decay,

$$R^{(\ell)}_{\rm incl}(q^2_0) = \frac{\int_{q^2_0}^{m^2_B} dq^2 \frac{d\Gamma(B\to X_s\bar{\ell}\ell)}{dq^2}}{\int_{q^2_0}^{m^2_B} dq^2 \frac{d\Gamma(B\to X_u\bar{\ell}\nu)}{dq^2}}, \tag{1}$$

where q^2_0 is the lower cut on q^2. The hadronic structure of the two transitions is very similar ($b \to q_{\rm light}$ left-handed current), leading to a significant cancellation of nonperturbative uncertainties when taking the ratio on an approximately equal phase space. Thanks to the recent experimental measurement of the $B \to X_u\bar{\ell}\nu$ inclusive rate as function of q^2 by Belle [17], the procedure proposed in [16] of computing the ratio (1) to predict $\Gamma(B \to X_s\bar{\ell}\ell)$ can finally be put in place.

On the experimental side, the $\Gamma(B \to X_s\bar{\ell}\ell)$ rate at high-q^2 is not fully available. However, in this kinematic region, only a few decay modes are relevant, and we can replace the

inclusive sum with the sum over a limited set of exclusive modes. To this end, we update the prediction for the nonresonant $B \to K\pi$ mode at high q^2 presented in Ref. [12]. By doing so, we show that for $q_0^2 = 15\ \mathrm{GeV}^2$, the inclusive rate is largely dominated by the two leading one-body modes ($B \to K$ and $B \to K^*$), with $B \to K\pi$ representing an $O(10\%)$ correction and additional multibody modes being further suppressed. We also show that the semi-inclusive rate obtained by summing the SM predictions for the leading one-body modes and the $B \to K\pi$ channel is in good agreement with the fully inclusive SM prediction obtained by means of $R^{(\ell)}_{\mathrm{incl}}(q_0^2)$ in Eq. (1). In other words, on the one hand, we cross-check the SM inclusive prediction; on the other, we validate the procedure to extract the (semi-)inclusive rate from data (while waiting for a fully inclusive measurement). As we shall show, the analysis of present data confirms the tension between data and SM prediction in the muon modes.

The paper is organized as follows: in Sec. II, we review the $b \to s\bar{\ell}\ell$ effective Lagrangian, pointing out the usefulness of a change of basis for the FCNC operators compared to the standard choice. In Sec. III, we present an updated estimate of

$$\Gamma\big(B \to X_s \bar{\ell}\ell\big)_{[15]} \equiv \Gamma\big(B \to X_s \bar{\ell}\ell, q^2 \geq 15\ \mathrm{GeV}^2\big), \tag{2}$$

by means of (1), following the analysis of Ref. [16]. In Sec. IV, we present the updated estimate of the $B \to K\pi\bar{\ell}\ell$ rate for $q^2 \geq 15\ \mathrm{GeV}^2$. In Sec. V, we compare inclusive vs semi-inclusive predictions within the SM, and the inclusive SM rate vs data (semi-inclusive, muon modes only). Finally, in Sec. VI, we discuss the implications for the Wilson coefficients, encoding short-distance physics, inferred by the comparison with data. The results are summarized in the Conclusions.

II. THE $b \to s\bar{\ell}\ell$ EFFECTIVE LAGRANGIAN

The effective Lagrangian valid below the electroweak scale relevant to $b \to s\bar{\ell}\ell$ transitions is conventionally written as

$$\mathcal{L}^{b\to s\bar{\ell}\ell}_{\mathrm{eff}} = \frac{4G_F}{\sqrt{2}}\frac{\alpha_e}{4\pi}\left(V^*_{ts}V_{tb}\sum_i C_i\mathcal{O}_i + \mathrm{H.c.}\right) + \mathcal{L}^{N_f=5}_{\mathrm{QCD}\times\mathrm{QED}}, \tag{3}$$

where we have used CKM unitarity, and neglected the tiny $O(V^*_{us}V_{ub})$ terms, to normalize all the flavor-changing operators in terms of a single CKM coefficient.

The only $\mathcal{O}_i$ with $b \to s\bar{\ell}\ell$ matrix elements, which are nonvanishing at tree level, are the electric-dipole operator,

$$\mathcal{O}_7 = \frac{m_b}{e}(\bar{s}_L\sigma_{\mu\nu}b_R)F^{\mu\nu}, \tag{4}$$

and the two FCNC semileptonic operators,

$$\mathcal{O}_9 = (\bar{s}_L\gamma_\mu b_L)(\bar{\ell}\gamma^\mu\ell), \quad \mathcal{O}_{10} = (\bar{s}_L\gamma_\mu b_L)(\bar{\ell}\gamma^\mu\gamma_5\ell). \tag{5}$$

For reasons that will be clear in the following, we find it convenient to perform a change of basis $\{\mathcal{O}_9, \mathcal{O}_{10}\} \to \{\mathcal{O}_V, \mathcal{O}_L\}$, where

$$\mathcal{O}_V = (\bar{s}_L\gamma_\mu b_L)(\bar{\ell}\gamma^\mu\ell), \quad \mathcal{O}_L = (\bar{s}_L\gamma_\mu b_L)(\bar{\ell}_L\gamma^\mu\ell_L), \tag{6}$$

such that

$$C_V = C_9 + C_{10}, \qquad C_L = -2C_{10}. \tag{7}$$

The new basis allows us to separate effective interactions, which originate by different underlying dynamics and behave differently in the evolution from high scales ($\mu_0 \sim m_t$) down to low scales ($\mu_b \sim m_b$). To better understand the different structures of these two operators, it is worth looking at the corresponding Wilson coefficients at the lowest nontrivial order.

A. The $\mathcal{O}_L$ operator

The purely left-handed operator is completely dominated by short-distance dynamics: it is generated at high scales by the top-quark Yukawa and $SU(2)_L$ interactions and, to a large extent, it does not evolve in the effective theory or mix with any other effective operator.

In the case of $\mathcal{O}_L$, the presence of α_e in the normalization of $\mathcal{L}^{b\to s\bar{\ell}\ell}_{\mathrm{eff}}$ is rather misleading: this is evident if we look at the overall coefficient of $\mathcal{O}_L$ (modulo the CKM factor), namely [18],

$$\mathcal{C}_L = \frac{4G_F}{\sqrt{2}}\frac{\alpha_e}{4\pi}C_L. \tag{8}$$

The corresponding one-loop expression is

$$\mathcal{C}^{(0)}_L = \frac{2G_F^2m_W^2}{\pi^2}Y_0(x_t) = \frac{y_t^2}{16\pi^2v^2}\left[1 + O(g^2/y_t^2)\right], \tag{9}$$

where $Y_0(x_t = m_t^2/m_W^2) \approx 0.98$ is the (finite) one-loop function [19], defined as in [20]. As can be seen, the expression of $\mathcal{C}^{(0)}_L$ depends only on the top-quark Yukawa coupling (y_t) and the $SU(2)_L$ coupling (g), once we normalize the effective interaction via the Higgs vacuum expectation value $v = (2\sqrt{2}G_F)^{1/2} \approx 246$ GeV. Moreover, this coefficient is nonzero in the so-called gaugeless limit of the SM (i.e., in the limit $g \to 0$ and $y_t \neq 0$, see, e.g., [21]).

The separation of $\mathcal{O}_L$ from all the other operators in $\mathcal{L}^{b\to s\bar{\ell}\ell}_{\mathrm{eff}}$ is guaranteed by the fact that, in the limit where we neglect light Yukawa couplings, the RG evolution arises only from QCD and QED, which are vectorlike theories. Note also that the bilinear quark current in $\mathcal{O}_L$ is a

TABLE I. Input parameters used for the computation of the inclusive rate.

Parameter	Value	Reference
m_t	172.7(5) GeV	[23]
$\alpha(m_b)^{-1}$	132.306(9)	[13,23]
$\lvert V_{ub}\rvert$	$3.82(20) \times 10^{-3}$	[23]
$\lvert V_{ts}\rvert$	$4.15(9) \times 10^{-2}$	[23]
$\lvert V_{tb}\rvert$	0.990903(64)	[23]

conserved current; hence, there is no mixing and no contribution in the Renormalization Group (RG) evolution of C_L at any order in QCD. A small anomalous dimension and a tiny mixing with other operators arise only from higher-order QED corrections.

The stability of $\mathcal{C}_L$ under quantum corrections is reflected by the small numerical difference between the leading (one-loop) result, $\mathcal{C}_L^{(0)} \approx 1.7 \times 10^{-7}$ GeV^{-2}, and the precise value estimated in [18,22], taking into account NNLO QCD and EW corrections (which play an important role in reducing the scale uncertainty in the high-scale matching). From the analysis of Ref. [22], taking into account the updated input for m_t (see Table I), we deduce

$$\mathcal{C}_L(m_b) = (1.662 \pm 0.008) \times 10^{-7}\ \text{GeV}^{-2}, \quad (10)$$

$$C_L(m_b) = 8.38 \pm 0.04, \quad (11)$$

where in (11) we have used the normalization (3) and, correspondingly, the value of $\alpha_e(m_b)$ in Table I.

B. The $\mathcal{O}_V$ operator

This effective operator receives contributions from all scales, vanishes in the limit $\alpha_e \to 0$ (i.e., in the limit $s_W \to 0$ at fixed g), and it mixes with the four-quark operators already at the one-loop level.

The one-loop expression, obtained without resumming large logarithms, is

$$C_V^{(0)}(\mu) = -4Z_0(x_t) + \frac{4}{9} - \frac{4}{9}\ln\left(\frac{\mu^2}{M_W^2}\right) \approx 0.22 - 0.89 \times \ln\left(\frac{\mu^2}{m_b^2}\right), \quad (12)$$

with $Z_0(x_t) \approx 0.67$ defined as in [20]. The numerical result in (12) is only qualitative (given we have not resummed the large logarithms), but it illustrates well the main features of C_V. There are two competing contributions that tend to cancel each other: (i) the finite and largely scale-independent short-distance contribution, encoded in $Z_0(x_t)$, and (ii) the charm-loop contribution generating large logarithms in the RG evolution from high scales to low scales. The cancellation becomes even more effective when the sizable QCD corrections are resummed via a proper treatment of the RG evolution. At NNLO accuracy [24], using the numerical results in [25], we find

$$C_V(\mu_b) = -0.01 \pm 0.14, \qquad \mu_b \in [2,5]\ \text{GeV}, \quad (13)$$

with a significant residual (low) scale dependence, which is a remnant of the μ dependence in Eq. (12).

C. Four-quark operators

Beside $\mathcal{O}_L$ and $\mathcal{O}_V$, an important role in $b \to s\bar{\ell}\ell$ amplitudes is played by the four-quark operators, whose matrix elements are nonvanishing beyond tree level. To high accuracy (i.e., to first order in α_e and arbitrary order in α_s), the contribution of four-quark operators can be expressed via a (process-dependent, nonlocal) modification of C_V,

$$C_{V,X_s}^{\text{eff}}(q^2) = \frac{\sum_{H_s \in X_s}\left(C_V\langle H_s\bar{\ell}\ell|\mathcal{O}_V|B\rangle + \sum_i C_i\langle H_s\bar{\ell}\ell|\mathcal{O}_i|B\rangle\right)}{\sum_{H_s \in X_s}\langle H_s\bar{\ell}\ell|\mathcal{O}_V|B\rangle}. \quad (14)$$

The sum over H_s denotes the sum over all the hadronic states belonging to the final state $|X_s\rangle$. Due to quark-hadron duality, we expect that for a sufficiently inclusive $|X_s\rangle$, the hadronic sum can be replaced by a partonic sum.

Evaluating the matrix elements of the $\mathcal{O}_i$ in Eq. (14) in perturbation theory at lowest order in α_s, leads to a process-independent expression that we denote $C_V^{\text{eff}}(q^2)|_{\text{pert}}^{(0)}$. More precisely, the coefficient thus obtained does not depend on $|X_s\rangle$, provided this state has the valence-quark content of $\mathcal{O}_V|B\rangle$. The expression of $C_V^{\text{eff}}(q^2)|_{\text{pert}}^{(0)}$ is the same for the fully inclusive mode or for an exclusive decay, such as $B \to K\bar{\ell}\ell$. Considering only the leading four-quark charm-quark operators[1] $\mathcal{O}_{1,2}^c$, which have $O(1)$ Wilson coefficients, one finds

$$C_V^{\text{eff}}(q^2)|_{\text{pert}}^{(0)} \approx C_V + \left(C_2^c + \frac{4}{3}C_1^c\right) \times h(m_c^2, q^2), \quad (15)$$

where $h(m_c^2, q^2)$ is given in [26] and $h(0, q^2) = (4/9) \times \log(\mu_b^2/m_b^2)$. Using the numerical expressions for the Wilson coefficients in [25], we find

$$\text{Re}\left[C_V^{\text{eff}}(q^2 = 15\ \text{GeV}^2)|_{\text{pert}}^{(0)}\right] = 0.43 \pm 0.26, \quad (16)$$

$$\text{Re}\left[C_V^{\text{eff}}(q^2 = 1\ \text{GeV}^2)|_{\text{pert}}^{(0)}\right] = 0.13 \pm 0.13. \quad (17)$$

[1]We define Wilson coefficients of four-quark operators as in [25] (the definition of the operators differs by an overall factor due to the different normalizations of $\mathcal{L}_{\text{eff}}^{b\to s\bar{\ell}\ell}$).

The error, due to the scale dependence, is closely connected to the scale variation of $C_V(\mu_b)$ in (13).

Going beyond this approximation, we can decompose $C^{\rm eff}_{V,X_s}$ for the inclusive case as

$$C^{\rm eff}_{V,X_s}(q^2) = C^{\rm eff}_{V,X_s}(q^2)|_{\rm pert} + C^{\rm eff}_{V,X_s}(q^2)|_{\rm n.p.} \quad (18)$$

The two terms on the rhs of Eq. (18) denote the result obtained in perturbation theory, considering partonic states, and possible additional nonperturbative contributions, respectively. The NLO corrections to the perturbative term, evaluated for the first time in [27] in the high-q^2 region, are within the error band of the leading contribution in Eq. (16).

On general grounds, nonperturbative contributions are expected to be smaller than perturbative ones. The latter start to lowest-order in α_s and are not power suppressed in the heavy-quark limit. The only notable exception is the q^2 region of the narrow charmonium resonances, where large *local* violations of quark-hadron duality do occur. We defer a more detailed discussion of possible nonperturbative contributions to Sec. VI. In the following, we limit ourselves to consider perturbative contributions only.

III. INCLUSIVE RATE AT HIGH q^2

The comparison between the numerical value of C_V in (13) and C_L in (11) indicates that, within the SM, the local part of the $b \to s\bar{\ell}\ell$ interaction has an approximate left-handed structure, as in the $b \to u\ell\bar{\nu}$ case. In both processes, we deal with a $b \to q_{\rm light}$ transition; hence, nonperturbative effects in sufficiently inclusive distributions are expected to be very similar. In $b \to s\bar{\ell}\ell$ transitions, corrections to a pure local left-handed interaction are generated by the matrix elements of $\mathcal{O}_7$ and those of the four-quark operators (discussed in Sec. II C). However, both these effects are quite small in the high-q^2 region. This is why the ratio (1) provides a very interesting observable to perform precise SM tests, as pointed out first in Ref. [16].

In order to compare this ratio with experiments, it is important to define the treatment of electromagnetic corrections below m_b. As pointed out first in [13], these give rise to $\log(m_\ell)$-enhanced terms in the q^2 spectrum, which implies a sizeable suppression of the rate in the high-q^2 region and a corresponding enhancement at low q^2. The origin of this effect is the migration of events to low q^2 due to real photon emissions by the dilepton system. As shown in [28], this effect is absent (and the electromagnetic corrections related to scales below m_b become tiny) if the cut employed to define the relevant kinematical region is $q_0^2 = (p_B - p_X)^2$. The experiments whose data we use for comparison effectively utilize this definition when analyzing exclusive modes [29]. We therefore do not include electromagnetic corrections (real or virtual) at scales below m_b in our prediction of $\mathcal{B}(B \to X_s\bar{\ell}\ell)^{\rm SM}_{[15]}$. Neglecting such corrections, the ratio (1) becomes m_ℓ independent (within the SM), and we can therefore drop the corresponding lepton label.

To provide an updated numerical prediction of $R_{\rm incl}(q_0^2)$ within the SM we reexpress the result of Ref. [16] in the $C_{L,V}$ basis, rather than in $C_{9,10}$ one. This way, we can write

$$R_{\rm incl}(q_0^2) = \frac{|V_{tb}V_{ts}^*|^2}{|V_{ub}|^2}\left[\mathcal{R}_L + \Delta\mathcal{R}_{[q_0^2]}\right], \quad (19)$$

where

$$\mathcal{R}_L = \frac{\alpha_e^2 C_L^2}{16\pi^2} = \frac{\mathcal{C}_L^2}{8G_F^2}, \quad (20)$$

and, for $q^2 = 15\ {\rm GeV}^2$,

$$\Delta\mathcal{R}_{[15]} = \frac{\alpha_e^2}{8\pi^2}\big[C_V^2 + C_V C_L + 0.485 C_L + 0.97 C_V + 0.93 + \Delta_{\rm n.p.} + C_7(1.91 + 2.05 C_L + 4.27 C_7 + 4.1 C_V)\big]. \quad (21)$$

The $\mathcal{R}_L$ term is the result obtained in the limit of purely left-handed interactions and identical hadronic distributions, while $\Delta\mathcal{R}_{[q_0^2]}$ describes all the deviations from this ideal limit. The numerical coefficients in $\Delta\mathcal{R}_{[15]}$ take into account the (perturbative) matrix elements of the four-quark operators, integrated over q^2, while C_V is the (q^2 independent) Wilson coefficient. We denote by $\Delta_{\rm n.p.}$, the nonperturbative effects estimated in [16]. The latter do not include $SU(3)$-breaking corrections due to light-quark masses. A naive estimate of these effects, from the phase space differences on the leading hadronic modes, indicates corrections up to 8% for $q_0^2 = 15\ {\rm GeV}^2$. This is very similar in size to the error associated with $\Delta_{\rm n.p.}$ that, as we shall see, does not represent the dominant source of uncertainty in the final estimate of $\mathcal{B}(B \to X_s\bar{\ell}\ell)^{\rm SM}_{[15]}$. Similarly, possible nonfactorizable contributions related to the broad charmonium resonances have not been explicitly included (we will revisit this point in Sec. VI).

The numerical expressions in the SM are

$$\mathcal{R}_L^{\rm SM} = (2.538 \pm 0.024) \times 10^{-5},$$
$$\Delta\mathcal{R}^{\rm SM}_{[15]} = \left(-0.03 \pm 0.14_{C_i} \pm 0.17_{\rm n.p.}\right) \times 10^{-5}. \quad (22)$$

As it can be seen, $\Delta\mathcal{R}^{\rm SM}_{[15]}$ is fully compatible with zero, but largely dominates the theoretical uncertainty in (1). The first error is due to the values of C_V and C_7,[2] while the second one is due to nonperturbative effects.

[2]Since the scale variation in Eq. (16) is larger with respect to the one in Eq. (13), we assume the former as a conservative estimate of the scale uncertainty on C_V.

Using the numerical results in (22), together with the experimental measurement of the $B \to X_u \bar{\ell}\nu$ inclusive rate for $q^2 \geq 15$ GeV^2 [17],[3]

$$\mathcal{B}(B \to X_u \bar{\ell}\nu)^{\text{exp}}_{[15]} = (1.50 \pm 0.24) \times 10^{-4}, \quad (23)$$

and the CKM inputs in Table I, we finally obtain

$$\mathcal{B}(B \to X_s \bar{\ell}\ell)^{\text{SM}}_{[15]} = (4.5 \pm 1.0) \times 10^{-7} \quad (24)$$

$$= 4.5 \times 10^{-7}[1 \pm 0.16_{\text{exp}} \pm 0.11_{\text{CKM}} \pm 0.09_{\Delta\mathcal{R}}]. \quad (25)$$

Note that the leading uncertainties are due to the experimental result in (23) and the CKM inputs. We could therefore expect a significant reduction of the total uncertainty in (24) in the near future.

Our estimate of $\mathcal{R}_{[15]}$ in (22), and the corresponding SM prediction in (25), are about 20% higher with respect to the results obtained in Ref. [15] for the $\ell = e$ case. A large fraction of this difference can be attributed to the different treatment of the real electromagnetic radiation. The discrepancy indeed reduces to about 5% when comparing with the results of Ref. [15] in the absence of long-distance electromagnetic corrections.[4]

IV. THE $B \to K\pi$ RATE AT HIGH q^2

As anticipated, our goal is twofold. First, we cross-check the SM prediction in (24) with the corresponding semi-inclusive result, namely the sum of the SM predictions of the leading exclusive modes. Second, we compare (24) with the experimental results in the high-q^2 region. A necessary ingredient to achieve both goals is the SM prediction of the $B \to K\pi$ rate, which we present in this section.

The $B \to K\pi$ process must be treated with some care since it receives resonant contributions from $\mathcal{B}(B \to (K^* \to K\pi)\bar{\ell}\ell)$, which are at least partially accounted for in the $B \to K^*\bar{\ell}\ell$ branching fraction. To avoid double counting these terms, we assume K^* dominance for the p-wave $B \to K\pi\bar{\ell}\ell$ decay amplitude. In other words, we describe this part of the amplitude via the exchange of the K^* resonance (also in the off shell region, assuming a q^2-independent K^* width). The K^* resonance cannot contribute to the s-wave part of the amplitude, and the interference between s and p waves cancels when integrated over the phase space at fixed q^2. We therefore compute separately the s-wave component of the total branching fraction, $\mathcal{B}(B \to (K\pi)_s \bar{\ell}\ell)$, that we treat as an independent decay channel.

[3]The result for $q^2 \geq 15$ GeV^2 is obtained by means of the q^2 differential data in https://doi.org/10.17182/hepdata.131599.

[4]See Table 5 in the Appendix of Ref. [15] (note that the ratio $\mathcal{R}$ defined in Ref. [15] includes also the CKM factors).

At high q^2, the light mesons have very low recoil energies, $E_{\text{had}} \ll \Lambda_{\text{QCD}}$, and heavy hadron chiral perturbation theory (HHChPT) is valid in this region. We calculate the leading s-wave contribution to the total $B \to K\pi$ rate by computing the $B \to K\pi$ matrix element in HHChPT, and subtracting the corresponding $B \to K^* \to K\pi$ contribution, evaluated at the $K\pi$ threshold. The HHChPT calculation was performed in Ref. [12], and we independently verified the result. In order to simplify the comparison to the $B \to K^* \to K\pi$ matrix element obtained using the lattice results of Ref. [30], we parametrize the matrix elements as

$$\begin{aligned}
&\langle K(p_K)\pi(p_\pi)|\bar{s}\gamma^\mu(1-\gamma_5)b|B(p)\rangle \\
&\quad = -i\big(w_+ P^\mu + w_- Q^\mu + c q^\mu + ih\epsilon^{\mu\nu\rho\sigma} q_\nu P_\rho Q_\sigma\big), \\
&\frac{iq_\nu}{q^2}\langle K(p_K)\pi(p_\pi)|\bar{s}\sigma^{\nu\mu}(1+\gamma_5)b|B(p)\rangle \\
&\quad = -i\big(w'_+ P^\mu + w'_- Q^\mu + c' q^\mu + ih'\epsilon^{\mu\nu\rho\sigma} q_\nu P_\rho Q_\sigma\big), \quad (26)
\end{aligned}$$

where $P^\mu = p_K^\mu + p_\pi^\mu$ and $Q^\mu = p_K^\mu - p_\pi^\mu$. The relevant form factors are related to those in Ref. [12], by

$$w_+^{(\prime)} = \frac{a^{(\prime)} + b^{(\prime)}}{2} + c^{(\prime)}, \qquad w_-^{(\prime)} = \frac{b^{(\prime)} - a^{(\prime)}}{2}. \quad (27)$$

Defining

$$w_1 = m_B\left(w_+ + \frac{d}{t} w_-\right), \qquad w_2 = m_B w_-, \quad (28)$$

the leading term of the differential decay width in the expansion around the $K\pi$ threshold is given by

$$\begin{aligned}
\frac{d\Gamma}{ds_\ell} &= \frac{G_F^2 M_B^5}{192\pi^3}|V_{ts}^* V_{tb}|^2 \frac{\alpha^2}{4\pi^2}\frac{1}{32\pi^2}\frac{\pi}{4}\frac{\sqrt{t^2 x_1 x_2}}{(1-t)^{3/2}} \\
&\times \left[\left(\frac{1}{2}|C_L|^2 + |C_V|^2 + \text{Re}(C_V^* C_L)\right) F_9 \right. \\
&\left. + 4m_b^2 |C_7|^2 F_7 + 4m_b \text{Re}\left\{C_7\left(\frac{1}{2}C_L^* + C_V^*\right) F_{97}\right\}\right] \\
&\times \left(s_\ell^{K\pi} - s_\ell\right)^3 + O\left[\left(s_\ell^{K\pi} - s_\ell\right)^{7/2}\right], \quad (29)
\end{aligned}$$

where the nondimensional parameters are $x_1 = m_K/m_B$, $x_2 = m_\pi/m_B$, $t = x_1 + x_2$, $d = x_1 - x_2$, $s_\ell = q^2/m_B^2$, and $s_\ell^{K\pi} \approx 0.775$ is the value of s_ℓ at the $K\pi$ threshold. The expression (29) does not take into account the matrix elements of the four quark operators. The latter can easily be incorporated by the replacement $C_V \to C_{V,K\pi}^{\text{eff}}(s_\ell m_B^2)$. Approximating $C_{V,K\pi}^{\text{eff}}$ with $C_V^{\text{eff}}|_{\text{pert}}^{(0)}$ in (15), this replacement has a negligible numerical impact given the additional sources of uncertainty.

TABLE II. Input parameters used for the computation of the exclusive branching fractions.

Parameter	Value	Reference
f_B	0.1900(13) GeV	[31]
f_π	0.13041(20) GeV	[32]
m_π	0.137(3) GeV	[23]
m_K	0.495(3) GeV	[23]
$m_{B^\pm}$	5.27925(26) GeV	[23]
$m_{K^{*\pm}}$	0.89547(77) GeV	[23]
$\Gamma_{K^{*\pm}}$	0.0462(13) GeV	[23]
$m_{B^{*\pm}} - m_{B^\pm}$	0.04537(21) GeV	[23]
$m_{B_s} - m_B$	0.08742(24) GeV	[23]

The F_i factors are given by

$$F_7 = |w_1'|^2 + \frac{4x_1x_2}{t^2}(1-t)|w_2'|^2,$$
$$F_9 = |w_1|^2 + \frac{4x_1x_2}{t^2}(1-t)|w_2|^2,$$
$$F_{97} = w_1'w_1^* + \frac{4x_1x_2}{t^2}(1-t)w_2'w_2^*. \tag{30}$$

The rate is largely dominated by the terms proportional to $w_+^{(\prime)}$ (i.e., the form factors associated with the total hadron momentum in the matrix elements), which are the only ones relevant to the s-wave transition. At the $K\pi$ threshold, where we can still trust the HHChPT result, we find

$$w_+ = 79.46, \qquad w_+' = 16.49, \tag{31}$$

in units of GeV^{-1}. Conversely, evaluating the $B \to K^* \to K\pi$ contribution at the $K\pi$ threshold, we find

$$w_+|_{\mathrm{res}} = 10.23 + 1.05i,$$
$$w_+'|_{\mathrm{res}} = 2.47 + 0.25i. \tag{32}$$

Comparing these two results, we determine

$$w_+^s = 69.23(62) - 1.05(6)i,$$
$$w_+'^s = 14.02(11) - 0.25(1)i, \tag{33}$$

again in units of GeV^{-1}. The uncertainties arise from both parametric inputs as well as lattice form factors.

Using the s-wave form factors in Eq. (29), and integrating for $q^2 \geq 15\ \mathrm{GeV}^2$, we find[5]

$$\mathcal{B}(B \to (K\pi)_s \overline{\ell}\ell)^{\mathrm{SM}}_{[15]} = (5.8 \pm 2.5) \times 10^{-8}, \tag{34}$$

where the errors are estimated from the fact that the NLO behavior in the expansion around the $K\pi$ threshold scales like $(s_\ell^{K\pi} - s_\ell)^{7/2}$. This error hugely dominates the parametric error, so the latter is not included in Eq. (34). Input parameters are given in Tables I and II. Additionally, we use $g_\pi \sim 0.5$ for the HHChPT coupling constant and, along with C_V and C_L given in Sec. II, the remaining C_i $(i = 1, \ldots, 8)$ are taken from Ref. [25].

V. INCLUSIVE HIGH-q^2 RATE AS SUM OF EXCLUSIVE MODES

The exclusive $B \to K\overline{\ell}\ell$ and $B \to K^*\overline{\ell}\ell$ branching fractions can be computed using the form factors calculated in Refs. [30,33]. Again, integrating for $q^2 \geq 15$ GeV, we find

$$\mathcal{B}(B \to K\overline{\ell}\ell)^{\mathrm{SM}}_{[15]} = \left(1.31 \pm 0.08_{\mathrm{lat}} \pm 0.09_{\mathrm{par}}\right) \times 10^{-7}$$
$$\mathcal{B}(B \to K^*\overline{\ell}\ell)^{\mathrm{SM}}_{[15]} = \left(3.19 \pm 0.21_{\mathrm{lat}} \pm 0.22_{\mathrm{par}}\right) \times 10^{-7}, \tag{35}$$

where "lat" refers to the uncertainty induced by the lattice form factors and "par" to the one from parametric inputs.

In the above, we have used the narrow-width approximation to estimate the K^* contribution. We can estimate the error due to a nonvanishing decay width by using the following double differential branching ratio, which is valid in the limit of constant width (or Breit-Wigner resonance):

$$\frac{d^2\mathcal{B}(B \to (K^* \to K\pi)\overline{\ell}\ell)}{dq^2 dp_{K\pi}^2} = \frac{d\mathcal{B}(B \to K^*\overline{\ell}\ell)}{dq^2} \frac{1}{\pi} \frac{m_{K^*}\Gamma_{K^*}\mathcal{B}(K^* \to K\pi)}{(p_{K\pi}^2 - m_{K^*}^2)^2 + m_{K^*}\Gamma_{K^*}}. \tag{36}$$

Expanding in powers of Γ_{K^*}/m_{K^*} and using $\mathcal{B}(K^* \to K\pi) \approx 1$ [23], gives

$$\mathcal{B}(B \to K^*\overline{\ell}\ell) \approx \left(1 - \frac{1}{\pi}\frac{\Gamma_{K^*}}{m_{K^*}}\right)[\mathcal{B}(B \to K^*\overline{\ell}\ell)]_{\Gamma=0}. \tag{37}$$

This implies an additional $O(1\%)$ correction that we can safely neglect due to the size of the form factor and parametric errors in Eq. (35).

The results in (34) and (35) can be combined to define a "correction factor" from the two-body final state relative to the one-body modes,

$$\Delta_{K\pi}^{[15]} = \frac{\mathcal{B}(B \to (K\pi)_s\overline{\ell}\ell)_{[15]}}{\mathcal{B}(B \to K\overline{\ell}\ell)_{[15]} + \mathcal{B}(B \to K^*\overline{\ell}\ell)_{[15]}} = 0.13 \pm 0.06. \tag{38}$$

This correction factor is largely independent of the values of the Wilson coefficients, which cancel in the ratio; hence,

[5]This value is obtained by summing the two isospin-related final states and is equivalent to 3/2 times the branching fraction featuring a charged pion in the final state.

it can be applied both in the SM and in a wide class of SM extensions.[6] In principle, multibody modes such as $B \to K\pi\pi\bar{\ell}\ell$ can also contribute to the total inclusive rate. However, these modes are suppressed even further by phase space factors and are expected to give a correction well within the current uncertainties of one- and two-body modes.

Using $\Delta^{[15]}_{K\pi}$, the semi-inclusive branching fraction obtained summing over one- and two-body modes can be written as

$$\sum_i \mathcal{B}\big(B \to X_s^i \bar{\ell}\ell\big)_{[15]} = \big(1 + \Delta^{[15]}_{K\pi}\big)\big[\mathcal{B}\big(B \to K\bar{\ell}\ell\big)_{[15]} + \mathcal{B}\big(B \to K^*\bar{\ell}\ell\big)_{[15]}\big], \tag{39}$$

both within and beyond the SM. Combining (39) and (35), we arrive at the following SM estimate of the semi-inclusive branching fraction ($q^2 \geq 15$ GeV):

$$\sum_i \mathcal{B}\big(B \to X_s^i \bar{\ell}\ell\big)^{\rm SM}_{[15]} = (5.07 \pm 0.42) \times 10^{-7}. \tag{40}$$

As can be seen, this result is well-compatible with the truly inclusive estimate presented in Eq. (24). The compatibility of these two results can be viewed both as a consistency check of the form factor calculations [30,33] or, alternatively, as a consistency check of the inclusive result in Eq. (24).

A. Comparison with data

The experimental determinations of the two leading modes in the high-q^2 region, and in the $\ell = \mu$ case, can be extracted from the results of the LHCb Collaboration, Ref. [2],

$$\begin{aligned} \mathcal{B}\big(B \to K\bar{\mu}\mu\big)^{\rm exp}_{[15]} &= (8.47 \pm 0.50) \times 10^{-8}, \\ \mathcal{B}\big(B \to K^*\bar{\mu}\mu\big)^{\rm exp}_{[15]} &= (1.58 \pm 0.35) \times 10^{-7}. \end{aligned} \tag{41}$$

Applying the correction factor in Eq. (38), we determine the following result for the measured semi-inclusive branching fraction:

$$\sum_i \mathcal{B}\big(B \to X_s^i \bar{\mu}\mu\big)^{\rm exp}_{[15]} = (2.74 \pm 0.41) \times 10^{-7}. \tag{42}$$

As summarized in Fig. 1, this result is significantly below the (consistent) SM predictions in Eqs. (24) and (40).

[6]We refer here to the motivated class of SM extensions where nonstandard contributions do not introduce sizable new local operators different from those present in (3).

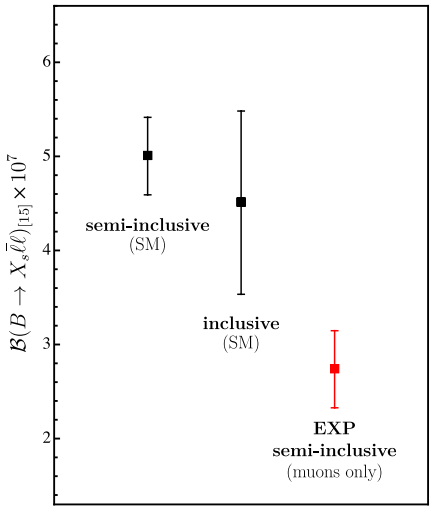

FIG. 1. SM predictions vs experimental data for the inclusive branching ratio, $\mathcal{B}(B \to X_s\bar{\ell}\ell)$, in the region $q^2 \geq 15$ GeV^2.

VI. DISCUSSION

The difference between the experimental result in (42) and the SM predictions in Eqs. (24) and (40) confirms the finding of several groups of a sizable suppression of the observed $b \to s\bar{\mu}\mu$ rates compared to SM expectations (see, e.g., [4,6,8] for recent analyses). The novel aspect of our analysis is that we support this conclusion, despite with a lower significance, by means of the inclusive rate at high q^2 in (24), which is *insensitive* to hadronic form factors. We thus provide an important independent verification of this phenomenon.

The inclusive rate in the high-q^2 region also has a different sensitivity to nonperturbative effects associated with charm rescattering, compared to exclusive observables (rates and angular distributions) in the low-q^2 region. We stress this point given that nonperturbative effects induced by charm rescattering have been invoked as a possible SM explanation for the (lepton-universal) anomalies observed in the low-q^2 region [11].

The fact that the observed discrepancy is hardly explained by charm rescattering, especially for the inclusive rate, can be better appreciated by looking at the size of the effect in the operator basis defined in Sec. II. In Fig. 2, we plot the region in the C_V–C_L plane favored by present data, i.e., treating C_V and C_L in Eq. (19) as free parameters and fitting the experimental result in (42). As already discussed in Sec. II C, both perturbative and nonperturbative contributions due to charm rescattering can be accounted for via an effective (q^2-dependent) modification of C_V. Assuming $C_L = C_L^{\rm SM}$, the modification of C_V necessary to describe the data is very large: it is larger, and opposite in sign, with respect to the perturbative estimate of charm rescattering contributions leading to Eq. (16).

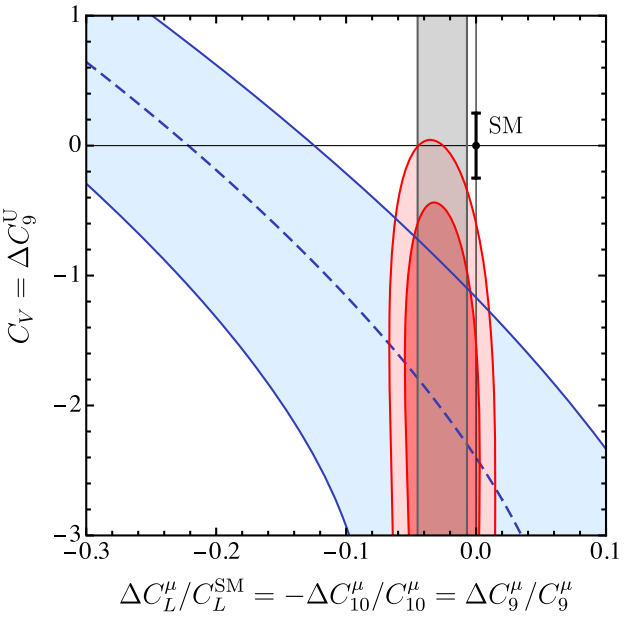

FIG. 2. Regions for the Wilson coefficients favored by experimental data. Here, $\Delta C_L^\mu = C_L^\mu - C_L^{\rm SM}$ is the correction to the SM value of C_L for the muon modes. The blue area is the 1σ compatibility region between the inclusive computation of $\mathcal{B}(B \to X_s \bar{\ell}\ell)$ and the experimental sum of exclusive modes (the dashed line indicates the best fit). The vertical gray band shows the 1σ value of C_L^μ determined by $\mathcal{B}(B_s \to \bar{\mu}\mu)$ and the LFU ratios (assuming a lepton-universal C_V). The dark and light red regions give the combined compatibility at 68% and 90% confidence level, respectively. To ease the comparison with previous studies, we also show on both axes the notation in the standard operator basis.

The central value of the discrepancy is beyond any realistic estimate of nonperturbative charm rescattering far from the narrow-charmonium region. The latter are not enhanced by RG logarithms, and in the high-q^2 region are expected to be of $O(\Lambda_{\rm QCD}^2/q^2)$. Explicit estimates of these effects for the inclusive rate [14] and the leading exclusive modes [34] lead to modifications of C_V of $O(1\%)$. Even assuming an order of magnitude enhancement (which would be hard to justify [14,34]), these contributions are within the error band shown in Fig. 2 (SM point), that we deduce from the scale variation of the perturbative contribution [as already stated, we assume the scale variation of $C_V^{\rm eff}(q^2)|_{\rm pert}^{(0)}$ in (16) as uncertainty for the SM estimate of C_V].

In exclusive modes, and specific values of q^2, large violations of quark-hadron duality are certainly possible. For instance, the large violations of naive factorization observed in [35], also in the high-q^2 region, are a manifestation of this statement. However, we stress that we are considering an inclusive quantity, where such effects are expected to be much smaller [36]. In conclusion, although we are unable to provide a rigorous upper bound on charm rescattering contributions, we believe that the size of $C_V^{\rm eff}(q^2)|_{\rm pert}^{(0)}$ and the explicit estimates of nonperturbative effects presented in [34,36], indicate that such effects cannot account for the bulk of the difference between SM and experimental points in Fig. 1.

In Fig. 2, we also show the impact of a possible change in C_L, which can occur only beyond the SM. More precisely, we consider the motivated case (see, e.g., [7,37]) of a lepton nonuniversal modification of C_L, affecting the muon modes only, versus a lepton-universal shift in C_V.[7] The value of ΔC_L^μ is strongly constrained by $\mathcal{B}(B_s \to \bar{\mu}\mu)$ and the lepton flavor universality (LFU) ratios (R_K and R_{K^*}). Updating the analysis of Ref. [37] taking into account the superseded values of the LFU ratios in [38,39], together with the R_{K_S} and $R_{K^{*+}}$ results in [40], and adding to the $\mathcal{B}(B_s \to \bar{\mu}\mu)$ data the CMS result in [41], we find $\Delta C_L^\mu = -0.026 \pm 0.019$. This result, taken alone, does not indicate a significant deviation from the SM; however, combining it with the constraint from $\mathcal{B}(B \to X_s \bar{\mu}\mu)_{[15]}$ leads to a preferred region in the C_V–ΔC_L^μ plane, which does not include the SM point at the 90% C.L.

On general grounds, beyond-the-SM contributions to the Wilson coefficients are expected to be small corrections over the SM ones (evaluated at the electroweak scale). Figure 2 shows that this condition cannot be satisfied if we assume nonstandard contributions to C_V only. On the other hand, this condition can be satisfied for both C_V and ΔC_L^μ, but only if $|\Delta C_L^\mu| \neq 0$, hence in the presence of a small but non-negligible LFU-violating amplitude.

VII. CONCLUSIONS

The inclusive $B \to X_s \bar{\ell}\ell$ rate at high dilepton invariant mass provides a clean and sensitive probe of $b \to s\bar{\ell}\ell$ amplitudes. In this paper, we have presented an updated estimate of this rate within the SM, by means of the ratio (1) and Belle's data on $\Gamma(B \to X_u \bar{\ell}\nu)$ [17]. The result, shown in Fig. 1, is in good agreement with the semi-inclusive estimate obtained summing the leading one-body modes (K and K^*) and the subleading nonresonant $K\pi$ channel in the relevant kinematical region. The uncertainty on the fully inclusive prediction is sizable, but it is dominated by the experimental error on $\Gamma(B \to X_u \bar{\ell}\nu)$; hence, it could be significantly improved in the near future.

The good compatibility between inclusive and semi-inclusive SM predictions confirms the expectation that the inclusive rate, in the kinematical region $q^2 \geq 15\ \text{GeV}^2$, is dominated by few exclusive modes. This opens up the possibility of a precise comparison with data for the rare

[7]The lepton universal nature of C_V is a natural consequence of its vectorlike structure: this effective operator can appear naturally (i.e., without a tuning between left-handed and right-handed components) via an effective short-distance interaction of the type $(\bar{s}_L \gamma_\mu b_L) D_\nu F^{\mu\nu}$, which is necessarily lepton universal.

decays with muon modes collected by LHCb. This comparison, also shown in Fig. 1, confirms the finding of several groups of a sizable suppression of the observed $b \to s\bar{\mu}\mu$ rates compared to SM expectation. The evidence of this effect from the inclusive high-q^2 rate does not exceed 2σ, but it is not based on hadronic form factors, hence providing an important independent verification of this phenomenon.

As a by-product of our analysis, we have shown that it is more convenient to describe short-distance contributions to $b \to s\bar{\ell}\ell$ amplitudes, both within and beyond the SM, via the effective operators $\mathcal{O}_V$–$\mathcal{O}_L$ in Eq. (6), rather than in the standard $\mathcal{O}_9$–$\mathcal{O}_{10}$ basis. In this basis, BSM contributions to $\mathcal{O}_V$ are naturally lepton universal, whereas those to $\mathcal{O}_L$ can be lepton-flavor dependent. The best-fit values in the C_V–C_L^μ plane, following from the present analysis, combined with recent data on LFU ratios and $\mathcal{B}(B_s \to \bar{\mu}\mu)$, are shown in Fig. 2. This analysis indicates that explaining the bulk of the present discrepancy via a modification to C_V only would require an unnaturally large correction to C_V, which we cannot justify via underestimated nonperturbative effects (and is also unlikely to appear in realistic BSM theories). By contrast, relatively small BSM effects to both C_V and C_L^μ can describe the data well.

As already stressed, uncertainties at present are still large, but the theory errors play a subleading role. We thus expect that the method outlined in this paper can have a significant impact in the near future in shedding light on the interesting puzzle of $b \to s\bar{\ell}\ell$ transitions.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 833280 (FLAY), and by the Swiss National Science Foundation (SNF) under Contract No. 200020_204428.

[1] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **111**, 191801 (2013).
[2] R. Aaij *et al.* (LHCb Collaboration), J. High Energy Phys. 06 (2014) 133.
[3] R. Aaij *et al.* (LHCb Collaboration), J. High Energy Phys. 02 (2016) 104.
[4] N. Gubernari, M. Reboud, D. van Dyk, and J. Virto, J. High Energy Phys. 09 (2022) 133.
[5] N. Gubernari, D. van Dyk, and J. Virto, J. High Energy Phys. 02 (2021) 088.
[6] M. Algueró, A. Biswas, B. Capdevila, S. Descotes-Genon, J. Matias, and M. Novoa-Brunet, Eur. Phys. J. C **83**, 648 (2023).
[7] M. Algueró, B. Capdevila, S. Descotes-Genon, P. Masjuan, and J. Matias, Phys. Rev. D **99**, 075017 (2019).
[8] W. Altmannshofer and P. Stangl, Eur. Phys. J. C **81**, 952 (2021).
[9] T. Hurth, F. Mahmoudi, and S. Neshatpour, Phys. Rev. D **103**, 095020 (2021).
[10] M. Ciuchini, A. M. Coutinho, M. Fedele, E. Franco, A. Paul, L. Silvestrini, and M. Valli, Eur. Phys. J. C **79**, 719 (2019).
[11] M. Ciuchini, M. Fedele, E. Franco, A. Paul, L. Silvestrini, and M. Valli, Phys. Rev. D **107**, 055036 (2023).
[12] G. Buchalla and G. Isidori, Nucl. Phys. **B525**, 333 (1998).
[13] T. Huber, E. Lunghi, M. Misiak, and D. Wyler, Nucl. Phys. **B740**, 105 (2006).
[14] T. Huber, T. Hurth, J. Jenkins, E. Lunghi, Q. Qin, and K. K. Vos, J. High Energy Phys. 10 (2019) 228.
[15] T. Huber, T. Hurth, J. Jenkins, E. Lunghi, Q. Qin, and K. K. Vos, J. High Energy Phys. 10 (2020) 088.
[16] Z. Ligeti and F. J. Tackmann, Phys. Lett. B **653**, 404 (2007).
[17] L. Cao *et al.* (Belle Collaboration), Phys. Rev. Lett. **127**, 261801 (2021).
[18] C. Bobeth, M. Gorbahn, and E. Stamou, Phys. Rev. D **89**, 034023 (2014).
[19] T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981); **65**, 1772(E) (1981).
[20] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996).
[21] R. Barbieri, G. Isidori, and D. Pappadopulo, J. High Energy Phys. 02 (2009) 029.
[22] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou, and M. Steinhauser, Phys. Rev. Lett. **112**, 101801 (2014).
[23] R. L. Workman *et al.* (Particle Data Group Collaboration), Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).
[24] C. Bobeth, P. Gambino, M. Gorbahn, and U. Haisch, J. High Energy Phys. 04 (2004) 071.
[25] T. Blake, G. Lanfranchi, and D. M. Straub, Prog. Part. Nucl. Phys. **92**, 50 (2017).
[26] W. G. Parrott, C. Bouchard, and C. T. H. Davies, Phys. Rev. D **107**, 014511 (2023).
[27] A. Ghinculov, T. Hurth, G. Isidori, and Y. P. Yao, Nucl. Phys. **B685**, 351 (2004).
[28] G. Isidori, S. Nabeebaccus, and R. Zwicky, J. High Energy Phys. 12 (2020) 104.
[29] G. Isidori, D. Lancierini, S. Nabeebaccus, and R. Zwicky, J. High Energy Phys. 10 (2022) 146.
[30] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, Proc. Sci. LATTICE2014 (**2015**) 372 [arXiv:1501.00367].
[31] Y. Aoki *et al.* (Flavour Lattice Averaging Group (FLAG) Collaboration), Eur. Phys. J. C **82**, 869 (2022).
[32] K. A. Olive (Particle Data Group Collaboration), Chin. Phys. C **38**, 090001 (2014).

[33] W. G. Parrott, C. Bouchard, and C. T. H. Davies (HPQCD and HPQCD Collaborations), Phys. Rev. D **107**, 014510 (2023).
[34] M. Beylich, G. Buchalla, and T. Feldmann, Eur. Phys. J. C **71**, 1635 (2011).
[35] J. Lyon and R. Zwicky, arXiv:1406.0566.
[36] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, Eur. Phys. J. C **61**, 439 (2009).
[37] C. Cornella, D. A. Faroughy, J. Fuentes-Martin, G. Isidori, and M. Neubert, J. High Energy Phys. 08 (2021) 050.
[38] The CMS Collaboration, Phys. Rev. Lett. **131**, 051803 (2023).
[39] The CMS Collaboration, Phys. Rev. D **108**, 032002 (2023).
[40] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **128**, 191802 (2022).
[41] The CMS Collaboration, Phys. Lett. B **842**, 137955 (2023).