Understanding the $B \to K^* \mu^+ \mu^-$ anomaly

Sébastien Descotes-Genon,[1] Joaquim Matias,[2] and Javier Virto[2]

[1]*Laboratoire de Physique Théorique, CNRS/Université Paris-Sud 11 (UMR 8627), 91405 Orsay Cedex, France*
[2]*Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain*

We present a global analysis of the $B \to K^*(\to K\pi)\mu^+\mu^-$ decay using the recent LHCb measurements of the primary observables $P_{1,2}$ and $P'_{4,5,6,8}$. Some of them exhibit large deviations with respect to the Standard Model (SM) predictions. We explain the observed pattern of deviations through a large new physics contribution to the Wilson coefficient of the semileptonic operator $\mathcal{O}_9$. This contribution has an opposite sign to the SM one, i.e., reduces the size of this coefficient significantly. A good description of data is achieved by allowing for new physics contributions to the Wilson coefficients $\mathcal{C}_7$ and $\mathcal{C}_9$ only. We find a 4.5σ deviation with respect to the SM prediction, combining the large-recoil $B \to K^*(\to K\pi)\mu^+\mu^-$ observables with other radiative processes. Once low-recoil observables are included the significance gets reduced to 3.9σ. We have tested different sources of systematics, none of them modifying our conclusions significantly. Finally, we propose additional ways of measuring the primary observables through new foldings.

The four-body $B \to K^*(\to K\pi)\mu^+\mu^-$ decay and its plethora of different observables [1–15] is becoming one of the key players not only in our search for new physics (NP) in the flavor sector but also to guide us in the construction of viable new models, which explains the remarkable experimental effort devoted to its precise measurement [16–20]. In the effective Hamiltonian approach used to analyze radiative decays at low energies, one of the most prominent virtues of this decay is the capacity to unveil NP contributions inside the short distance Wilson coefficients, denoted $\mathcal{C}_i = \mathcal{C}_i^{\rm SM} + \mathcal{C}_i^{\rm NP}$, not only for the Standard Model (SM) electromagnetic and dileptonic operators

$$\mathcal{O}_7 = e/(16\pi^2)m_b(\bar{s}\sigma_{\mu\nu}P_R b)F^{\mu\nu}, \qquad (1)$$

$$\mathcal{O}_9 = e^2/(16\pi^2)(\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \ell), \qquad (2)$$

$$\mathcal{O}_{10} = e^2/(16\pi^2)(\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \gamma_5 \ell), \qquad (3)$$

(with the usual $P_{L,R}$ chirality projection operators) but also for the chirally flipped operators $\mathcal{O}_{i'}$ as well as the scalar and pseudoscalar operators $\mathcal{O}_{S,P,S',P'}$. Among these, the only non-negligible Wilson coefficients in the SM are $\mathcal{C}_{7\text{eff},9,10}^{\rm SM}(\mu_b) = (-0.29, 4.07, -4.31)$ at $\mu_b = 4.8$ GeV. The correlations between the Wilson coefficients constitute a unique test ground to find consistent patterns pointing towards specific NP models.

However, the presence of hadronic effects can easily hide a NP signal. For this reason, it is essential to design an optimized basis of observables, easy to measure, with low hadronic and high NP sensitivities. In Refs. [15,21] we proposed such a basis, consisting of $P_{1,2,3}$ and $P'_{4,5,6}$ (primary observables with a low sensitivity to form-factor uncertainties at low dilepton invariant mass q^2), together with F_L (or $A_{\rm FB}$) and $d\Gamma/dq^2$ (containing large uncertainties but required to complete the basis).

There has been an evolution in the type of observables measured by LHCb. It started with the set of observables $A_{\rm FB}$, F_L and S_3 [19], all of them rather sensitive to hadronic uncertainties. The experimental results pointed towards a scenario consistent with the SM, but with small deviations in $A_{\rm FB}$ (in both the q^2 bin [2–4.3] GeV^2 and the position of the zero). The next generation of measurements included a theoretically controlled version of $A_{\rm FB}$ called $A_T^{\rm (re)}$ [6] or P_2 [7], and P_1, which are both less sensitive to hadronic effects and able to magnify deviations due to NP. Finally, LHCb has issued very recent results [20] completing the basis of P_i and P'_i primary observables [7,15,21]. These observables, with little sensitivity to hadronic uncertainties at low q^2, have unveiled a set of tensions with respect to the SM that have to be understood from the theoretical point of view. This paper aims at providing such a consistent picture, where the Wilson coefficient $\mathcal{C}_9$ plays an essential role.

In Sec. I we discuss the experimental evidence, i.e., the pattern of deviations observed at LHCb. In Sec. II, we present the main results and the details of our analysis of data using our basis of observables. Finally, in Sec. III we explore the robustness of these results analysing different sources of uncertainty, namely, their sensitivity to perturbative and nonperturbative charm effects, large power-suppressed corrections as well as the comparison between naive and next-to-leading-order (NLO) QCD factorizations. We discuss possible improvements on the control of the S-wave pollution in an appendix. We conclude by suggesting cross-checks of our findings and further prospects for similar analyses.

I. EXPERIMENTAL EVIDENCE

The recent LHCb breakthrough, leading to the measurement of most observables of the basis, namely, $P_{1,2}$ and $P'_{4,5,6,8}$ using folded distributions [19,20], actually exhibits a consistent pattern of deviations with respect to SM expectations. In Table I we summarize the experimental results expressed in our convention in Refs. [7,15]. Ordering the bins according to the dilepton invariant mass q^2, and focusing on the first three bins, comparing the data [19,20] with the NP scenarios discussed in Ref. [7] leads to the following comments:

(i)   There is no substantial deviation in P_1 (still with large error bars). From this observable, no definite conclusion can be drawn yet on the presence nor the absence of contributions from right-handed currents. One can notice a mild preference for negative values in the first two bins and positive ones in the third bin, also present in CDF measurements [16].

(ii) A slight preference for a lower value of the second and third bins of A_{FB} is consistent with a 2.9σ (1.7σ) deviation in the second (third) bin of P_2. One notices also a preference for a slightly higher q^2 value for the zero of A_{FB} (at the same position as the zero of P_2). Both effects can be accommodated with $\mathcal{C}_7^{NP} < 0$ and/or $\mathcal{C}_9^{NP} < 0$.

(iii) There is a striking 4.0σ (1.6σ) deviation in the third (second) bin of P'_5,[1] consistent with large negative contributions in $\mathcal{C}_7^{NP}$ and/or $\mathcal{C}_9^{NP}$.

(iv) P'_4 is in agreement with the SM, but within large uncertainties, and it has future potential to determine the sign of $\mathcal{C}_{10}^{NP}$.

(v) P'_6 and P'_8 show small deviations with respect to the SM, but such effect would require complex phases in $\mathcal{C}_9^{NP}$ and/or $\mathcal{C}_{10}^{NP}$.

A similar pattern of deviations can be observed when one considers the wider q^2 bin $[1, 6]$ GeV^2. Other mechanisms, involving contributions to chirally flipped operators $\mathcal{C}_{7',9',10'}^{NP}$, could yield similar effects on some of these observables, but they fail to provide a consistent picture, as will become clear in the following.

Indeed, deviations involving mainly P_2, P'_4 and P'_5 at low q^2 can be understood qualitatively if NP affects $\mathcal{C}_9$ and/or $\mathcal{C}_{10}$. First, at low q^2, this NP contribution will impact the transversity amplitudes $A_{\perp,\parallel}$ only mildly (as it will be hidden by the contribution proportional to $\mathcal{C}_7$ which is enhanced by the $1/q^2$ photon pole) but it will affect A_0 much more significantly (where $\mathcal{C}_7$ is not enhanced). At low q^2, this type of NP will thus mainly affect observables built from the A_0 amplitude such as P'_4 and P'_5. Second, as shown

---

[1]In Ref. [20], only a 3.7σ deviation was quoted for the third bin, due to the fact that the comparison was performed between the central experimental value and the 68.3% C.L. upper bound for the theoretical prediction. Here the deviation is computed comparing the two central values.

TABLE I.   Experimental averages and SM predictions for the observables used in the analysis. The dictionary between the different conventions used in Refs. [7,15,19,20] is $P_1^{LHCb} = P_1$, $P_2^{LHCb} = -P_2$, $P_4^{\prime LHCb} = -\frac{1}{2}P'_4$, $P_5^{\prime LHCb} = P'_5$, $P_6^{\prime LHCb} = -P'_6$ and $P_8^{\prime LHCb} = \frac{1}{2}P'_8$ (sign differences stem from $\theta_\ell^{LHCb} = \pi - \theta_\ell$).

Observable	Experiment	SM prediction	Pull
$\langle P_1 \rangle_{[0.1,2]}$	$-0.19^{+0.40}_{-0.35}$	$0.007^{+0.043}_{-0.044}$	-0.5
$\langle P_1 \rangle_{[2,4.3]}$	$-0.29^{+0.65}_{-0.46}$	$-0.051^{+0.046}_{-0.046}$	-0.4
$\langle P_1 \rangle_{[4.3,8.68]}$	$0.36^{+0.30}_{-0.31}$	$-0.117^{+0.056}_{-0.052}$	$+1.5$
$\langle P_1 \rangle_{[1,6]}$	$0.15^{+0.39}_{-0.41}$	$-0.055^{+0.041}_{-0.043}$	$+0.5$
$\langle P_2 \rangle_{[0.1,2]}$	$0.03^{+0.14}_{-0.15}$	$0.172^{+0.020}_{-0.021}$	-1.0
$\langle P_2 \rangle_{[2,4.3]}$	$0.50^{+0.00}_{-0.07}$	$0.234^{+0.060}_{-0.086}$	$+2.9$
$\langle P_2 \rangle_{[4.3,8.68]}$	$-0.25^{+0.07}_{-0.08}$	$-0.407^{+0.049}_{-0.037}$	$+1.7$
$\langle P_2 \rangle_{[1,6]}$	$0.33^{+0.11}_{-0.12}$	$0.084^{+0.060}_{-0.078}$	$+1.8$
$\langle P'_4 \rangle_{[0.1,2]}$	$0.00^{+0.52}_{-0.52}$	$-0.342^{+0.031}_{-0.026}$	$+0.7$
$\langle P'_4 \rangle_{[2,4.3]}$	$0.74^{+0.54}_{-0.60}$	$0.569^{+0.073}_{-0.063}$	$+0.3$
$\langle P'_4 \rangle_{[4.3,8.68]}$	$1.18^{+0.26}_{-0.32}$	$1.003^{+0.028}_{-0.032}$	$+0.6$
$\langle P'_4 \rangle_{[1,6]}$	$0.58^{+0.32}_{-0.36}$	$0.555^{+0.067}_{-0.058}$	$+0.1$
$\langle P'_5 \rangle_{[0.1,2]}$	$0.45^{+0.21}_{-0.24}$	$0.533^{+0.033}_{-0.041}$	-0.4
$\langle P'_5 \rangle_{[2,4.3]}$	$0.29^{+0.40}_{-0.39}$	$-0.334^{+0.097}_{-0.113}$	$+1.6$
$\langle P'_5 \rangle_{[4.3,8.68]}$	$-0.19^{+0.16}_{-0.16}$	$-0.872^{+0.053}_{-0.041}$	$+4.0$
$\langle P'_5 \rangle_{[1,6]}$	$0.21^{+0.20}_{-0.21}$	$-0.349^{+0.088}_{-0.100}$	$+2.5$
$\langle P'_6 \rangle_{[0.1,2]}$	$0.24^{+0.23}_{-0.20}$	$-0.084^{+0.034}_{-0.044}$	$+1.6$
$\langle P'_6 \rangle_{[2,4.3]}$	$-0.15^{+0.38}_{-0.36}$	$-0.098^{+0.043}_{-0.056}$	-0.1
$\langle P'_6 \rangle_{[4.3,8.68]}$	$0.04^{+0.16}_{-0.16}$	$-0.027^{+0.060}_{-0.063}$	$+0.4$
$\langle P'_6 \rangle_{[1,6]}$	$0.18^{+0.21}_{-0.21}$	$-0.089^{+0.042}_{-0.052}$	$+1.3$
$\langle P'_8 \rangle_{[0.1,2]}$	$-0.12^{+0.56}_{-0.56}$	$0.037^{+0.037}_{-0.030}$	-0.3
$\langle P'_8 \rangle_{[2,4.3]}$	$-0.30^{+0.60}_{-0.58}$	$0.070^{+0.045}_{-0.034}$	-0.6
$\langle P'_8 \rangle_{[4.3,8.68]}$	$0.58^{+0.34}_{-0.38}$	$0.020^{+0.054}_{-0.055}$	$+1.5$
$\langle P'_8 \rangle_{[1,6]}$	$0.46^{+0.36}_{-0.38}$	$0.063^{+0.042}_{-0.033}$	$+1.0$
$\langle A_{FB} \rangle_{[0.1,2]}$	$-0.02^{+0.13}_{-0.13}$	$-0.136^{+0.051}_{-0.048}$	$+0.8$
$\langle A_{FB} \rangle_{[2,4.3]}$	$-0.20^{+0.08}_{-0.08}$	$-0.081^{+0.055}_{-0.069}$	-1.1
$\langle A_{FB} \rangle_{[4.3,8.68]}$	$0.16^{+0.06}_{-0.05}$	$0.220^{+0.138}_{-0.113}$	-0.5
$\langle A_{FB} \rangle_{[1,6]}$	$-0.17^{+0.06}_{-0.06}$	$-0.035^{+0.037}_{-0.034}$	-2.0
$\langle P_1 \rangle_{[14.18,16]}$	$0.07^{+0.26}_{-0.28}$	$-0.352^{+0.697}_{-0.468}$	$+0.6$
$\langle P_1 \rangle_{[16,19]}$	$-0.71^{+0.36}_{-0.26}$	$-0.603^{+0.589}_{-0.315}$	-0.2
$\langle P_2 \rangle_{[14.18,16]}$	$-0.50^{+0.03}_{-0.00}$	$-0.449^{+0.136}_{-0.041}$	-1.1
$\langle P_2 \rangle_{[16,19]}$	$-0.32^{+0.08}_{-0.08}$	$-0.374^{+0.151}_{-0.126}$	$+0.3$
$\langle P'_4 \rangle_{[14.18,16]}$	$-0.18^{+0.54}_{-0.70}$	$1.161^{+0.190}_{-0.332}$	-2.1
$\langle P'_4 \rangle_{[16,19]}$	$0.70^{+0.44}_{-0.52}$	$1.263^{+0.119}_{-0.248}$	-1.1
$\langle P'_5 \rangle_{[14.18,16]}$	$-0.79^{+0.27}_{-0.22}$	$-0.779^{+0.328}_{-0.363}$	$+0.0$
$\langle P'_5 \rangle_{[16,19]}$	$-0.60^{+0.21}_{-0.18}$	$-0.601^{+0.282}_{-0.367}$	$+0.0$
$\langle P'_6 \rangle_{[14.18,16]}$	$0.18^{+0.24}_{-0.25}$	$0.000^{+0.000}_{-0.000}$	$+0.7$
$\langle P'_6 \rangle_{[16,19]}$	$-0.31^{+0.38}_{-0.39}$	$0.000^{+0.000}_{-0.000}$	-0.8
$\langle P'_8 \rangle_{[14.18,16]}$	$-0.40^{+0.60}_{-0.50}$	$-0.015^{+0.009}_{-0.013}$	-0.6
$\langle P'_8 \rangle_{[16,19]}$	$0.12^{+0.52}_{-0.54}$	$-0.008^{+0.005}_{-0.007}$	$+0.2$
$\langle A_{FB} \rangle_{[14.18,16]}$	$0.51^{+0.07}_{-0.05}$	$0.404^{+0.199}_{-0.191}$	$+0.5$
$\langle A_{FB} \rangle_{[16,19]}$	$0.30^{+0.08}_{-0.08}$	$0.360^{+0.205}_{-0.172}$	-0.3
$10^4 \mathcal{B}_{B \to X_s \gamma}$	3.43 ± 0.22	3.15 ± 0.23	$+0.9$
$10^6 \mathcal{B}_{B \to X_s \mu^+ \mu^-}$	1.60 ± 0.50	1.59 ± 0.11	$+0.0$
$10^9 \mathcal{B}_{B_s \to \mu^+ \mu^-}$	2.9 ± 0.8	3.56 ± 0.18	-0.8
$A_I(B \to K^* \gamma)$	0.052 ± 0.026	0.041 ± 0.025	$+0.3$
$S_{K^* \gamma}$	-0.16 ± 0.22	-0.03 ± 0.01	-0.6

in Ref. [7], the position of the zero of P_2 and P'_5 is affected by a contribution $\mathcal{C}_7^{\mathrm{NP}}$ and/or $\mathcal{C}_9^{\mathrm{NP}}$, whereas that of P'_4 depends on $\mathcal{C}_7^{\mathrm{NP}}$, $\mathcal{C}_9^{\mathrm{NP}}$ and $\mathcal{C}_{10}^{\mathrm{NP}}$. A NP contribution $\mathcal{C}_9^{\mathrm{NP}} < 0$ would shift the zero of P_2 (resp. P'_4 and P'_5) to higher (resp. lower and higher) q^2 values, leading to an increase of the third bin (resp. an increase of the first bin and an increase of the second and third bins). A contribution from $\mathcal{C}_{10}^{\mathrm{NP}} < 0$ would mainly affect the zero of P'_4 in a similar way to $\mathcal{C}_9^{\mathrm{NP}} < 0$, whereas a contribution from $\mathcal{C}_7^{\mathrm{NP}} < 0$ would have the same impact as $\mathcal{C}_9^{\mathrm{NP}} < 0$ in P_2 and P'_5.

II. ANALYSIS, RESULTS AND DISCUSSION

Qualitatively, all the comments of the previous section point towards a scenario where $\mathcal{C}_9^{\mathrm{NP}} < 0$ (with possible small contributions $\mathcal{C}_7^{\mathrm{NP}}$, $\mathcal{C}_{10}^{\mathrm{NP}} < 0$) in a consistent way. We will now proceed with a quantitative analysis of these measurements.

A. General analysis and overview of constraints

We start with a global analysis to the data, in a general scenario with simultaneous arbitrary (real) NP contributions to the Wilson coefficients $\mathcal{C}_{7,9,10}$ and $\mathcal{C}_{7',9',10'}$, writing $\mathcal{C}_i = \mathcal{C}_i^{\mathrm{SM}} + \mathcal{C}_i^{\mathrm{NP}}$ (we neglect scalar contributions). We use the predictions and uncertainty estimates described in Ref. [21], following NLO QCD factorization for the large-recoil (low-q^2) bins [22,23] and Heavy-Quark Effective Theory for the low-recoil (large-q^2) bins [24]. We follow a standard χ^2 frequentist approach, following Appendix C in Ref. [15]: we take into account the asymmetric errors on the experimental numbers, estimate theoretical uncertainties for each set of values for $\mathcal{C}_i^{\mathrm{NP}}$, and treat all uncertainties (experimental and theoretical) as statistical to combine them in quadrature. The correlations among the measurements are not available currently and are thus neglected. We consider the following observables:

(1) *Optimized observables in $B \to K^* \mu^+ \mu^-$*: We take P_1, P_2, P'_4, P'_5, P'_6 and P'_8, within the 3 large-recoil bins [0.1,2], [2,4.3] and [4.3, 8.68] GeV^2, and the 2 low-recoil bins [14.18,16] and [16, 19] GeV^2. We note that: (a) all these observables are independent, as P_3 is not measured (see Ref. [7]), (b) these observables have little hadronic sensitivity only at low q^2 [21], so we expect weak constraints from the low-recoil region, (c) we do not consider the [1,6] bin at this stage since these observables are not independent of the previous ones. We will consider this bin later on.

(2) *Forward-backward asymmetry in $B \to K^* \mu^+ \mu^-$*: Once one has chosen a maximal set of optimized observables, one has still to choose two independent observables sensitive to form-factor uncertainties. The differential branching ratio $d\mathcal{B}/dq^2$ is one of them, necessary to fix the overall normalization. We do not include this observable because of its large theoretical uncertainty derived from its significant

sensitivity to hadronic form factors. The other observable can be either A_{FB} or F_L. We choose A_{FB} because of its expected higher sensitivity to $\mathcal{C}_9^{\mathrm{NP}}$ and its complementarity with P_2 [7,25]. Again, we keep the [1,6] bin for a later stage of the analysis. We consider only LHCb measurements (the inclusion of the available results from other experiments has only a marginal impact on the data [26]).

(3) *Radiative and dileptonic B decays*: There are other important $b \to s$ penguin modes sensitive to magnetic and dileptonic operators. We consider the branching ratios $\mathcal{B}(B \to X_s \gamma)_{E_\gamma > 1.6\,\mathrm{GeV}}$, $\mathcal{B}(B \to X_s \mu^+ \mu^-)_{[1,6]}$ and $\mathcal{B}(B_s \to \mu^+ \mu^-)$, the isospin asymmetry $A_I(B \to K^* \gamma)$ and the $B \to K^* \gamma$ time-dependent CP asymmetry $S_{K^* \gamma}$. Relevant formulas for these observables can be found in Ref. [27], while updated experimental numbers are taken from Refs. [26,28,29] and Refs. [30–32] [where the average for $\mathcal{B}(B_s \to \mu^+ \mu^-)$ takes into account differences in the ratio of production fractions f_s/f_d and normalizations for CMS and LHCb]. We disregard other similar observables, either because their theoretical description is not ascertained, such as $A_{CP}(B \to X_s \gamma)$, or because of experimental issues, as is the case with $B \to K \mu^+ \mu^-$ due to the unclear status of the experimental separation of neutral and charge modes indicated by the measured isospin asymmetry [33].

For $B \to K \mu^+ \mu^-$, an additional issue was raised in Ref. [34], as an unexpected resonant structure $\psi(4160)$ has been observed in $B^+ \to K^+ \mu^+ \mu^-$ at low recoil. It remains to be seen how this resonant structure can impact the neutral mode around $q^2 \simeq 17.3$ GeV^2, and if it can modify the predictions for $B \to K^* \mu^+ \mu^-$ observables for the two bins in the low-recoil region. In the following analysis, we will always consider two data sets: one with only large-recoil data, the other one with both low- and large-recoil data.

Experimental averages and SM theoretical predictions for all these observables are summarized in Table I. As an outcome of the fit, all Wilson coefficients have 2σ C.L. intervals encompassing the SM value, except for $\mathcal{C}_9$ (below its SM value) with a best-fit point around $\mathcal{C}_9^{\mathrm{NP}} \sim -1.2$. The SM hypothesis ($\mathcal{C}_i^{\mathrm{NP}} = 0$ for all i) has a pull[2] of 2.9σ. The individual 1, 2 and 3σ intervals for the Wilson coefficients are given in Table II.[3] The most economical scenario

---

[2]When testing a hypothesis (e.g. the SM) in a given framework, we refer to its p value in units of σ as the "pull" of this hypothesis. Therefore, a lower p value corresponds to a larger pull in units of σ.

[3]Besides the region close to the SM point, there is an allowed but less favored region close to the flipped-sign solution for the Wilson coefficients ($\mathcal{C}_i \simeq -\mathcal{C}_i^{\mathrm{SM}}$). The appearance of this region is expected and well understood (e.g., Ref. [27]), but would correspond to a significant amount of NP. We thus disregard it for the time being, even though this should be kept in mind in future analyses.

TABLE II. 68.3% (1σ), 95.5% (2σ) and 99.7% (3σ) confidence intervals for the NP contributions to Wilson coefficients resulting from the global analysis.

Coefficient	1σ	2σ	3σ
$\mathcal{C}_7^{\mathrm{NP}}$	$[-0.05, -0.01]$	$[-0.06, 0.01]$	$[-0.08, 0.03]$
$\mathcal{C}_9^{\mathrm{NP}}$	$[-1.6, -0.9]$	$[-1.8, -0.6]$	$[-2.1, -0.2]$
$\mathcal{C}_{10}^{\mathrm{NP}}$	$[-0.4, 1.0]$	$[-1.2, 2.0]$	$[-2.0, 3.0]$
$\mathcal{C}_{7'}^{\mathrm{NP}}$	$[-0.04, 0.02]$	$[-0.09, 0.06]$	$[-0.14, 0.10]$
$\mathcal{C}_{9'}^{\mathrm{NP}}$	$[-0.2, 0.8]$	$[-0.8, 1.4]$	$[-1.2, 1.8]$
$\mathcal{C}_{10'}^{\mathrm{NP}}$	$[-0.4, 0.4]$	$[-1.0, 0.8]$	$[-1.4, 1.2]$

corresponds to a negative NP contribution to $\mathcal{C}_9$ with all the other Wilson coefficients close to their SM value. Even though the branching ratio is affected by very large uncertainties and is not considered in our analysis, it is also interesting to notice that $\mathcal{C}_9^{\mathrm{NP}} < 0$ would tend to decrease the differential branching ratio, improving the agreement with experimental data.

The large-recoil data favors $\mathcal{C}_9^{\mathrm{NP}} < 0$ more significantly than the low-recoil region. Removing the low-recoil $B \rightarrow K^* \mu^+ \mu^-$ observables from the fit enhances the effect, with a best-fit point around $\mathcal{C}_9^{\mathrm{NP}} \sim -1.6$. In this case also a negative contribution $\mathcal{C}_{9'}^{\mathrm{NP}} < 0$ is favored. This pattern is also obtained when considering only the [1,6] bin, though with larger uncertainties. We discuss these issues in more depth below.

B. Comparison of NP scenarios

In order to clarify the role played by $\mathcal{C}_9$ in explaining the $B \rightarrow K^* \mu^+ \mu^-$ anomaly and to discuss the role of the other coefficients, it is interesting to consider nested scenarios where NP is turned on for each Wilson coefficient one after the other, starting from the SM hypothesis. In a given scenario (where some Wilson coefficients $\mathcal{C}_{j_1,\dots,j_N}$ receive NP and the others do not), the improvement obtained by allowing one more Wilson coefficient $\mathcal{C}_i$ to receive NP contributions can be quantified by computing the pull of the $\mathcal{C}_i^{\mathrm{NP}} = 0$ hypothesis within the scenario where $\mathcal{C}_i^{\mathrm{NP}}$ and $\mathcal{C}_{j_1,\dots,j_N}^{\mathrm{NP}}$ are left free [35].

When considering the full set of large- and low-recoil data for $B \rightarrow K^* \mu^+ \mu^-$, we find that the larger pulls (around $\sim 4\sigma$) are obtained when adding $\mathcal{C}_9^{\mathrm{NP}}$, independently of which other Wilson coefficients are left free to receive NP contributions. The next-to-larger pulls are obtained by adding $\mathcal{C}_7^{\mathrm{NP}}$ (around $\sim 3\sigma$), in all cases except when $\mathcal{C}_9^{\mathrm{NP}}$ has been added beforehand; in such a case, the pull is $\sim 1.3\sigma$ (still the lowest after $\mathcal{C}_9^{\mathrm{NP}}$). The rest of the pulls are always around or below 1σ. These results are consistent with the fact that $\mathcal{C}_9$ plays a prominent role in explaining the $B \rightarrow K^* \mu^+ \mu^-$ anomaly; besides that, a NP contribution to $\mathcal{C}_7$ is also favored though less strongly.

It is also interesting to consider only the large-recoil bins, for which the theoretical description of the optimized

observables is more accurate. The main picture remains the same, although in some cases some other coefficients may play a (more modest) role in the discussion. For instance, pulls around $\sim 3\sigma$ can be obtained for $\mathcal{C}_{10}$ and $\mathcal{C}_{10'}$ if we insist that $\mathcal{C}_9^{\mathrm{NP}} = 0$ (not otherwise). Another example arises in a scenario with only $\mathcal{C}_9^{\mathrm{NP}} \neq 0$, for which the next most relevant NP contributions are $\mathcal{C}_7$ and $\mathcal{C}_{9'}$ (with pulls at the same level).

Finally, using both low- and large-recoil data for $B \rightarrow K^* \mu^+ \mu^-$, we can compute the pull corresponding to the $\mathcal{C}_{10,7',9',10'}^{\mathrm{NP}} = 0$ hypothesis in the scenario where all 6 Wilson coefficients are allowed to receive NP contributions. The resulting pull is below $\sim 0.2\sigma$, indicating that no other NP contribution is required by the data apart from $\mathcal{C}_9^{\mathrm{NP}}$ and $\mathcal{C}_7^{\mathrm{NP}}$. The same results occur when only large-recoil data is used.

Before focusing on this scenario in the following, a comment is in order concerning alternative scenarios with different sets of coefficients receiving NP contributions. In all scenarios considered the best fit corresponds to $\mathcal{C}_9^{\mathrm{NP}} \sim -1.2$ with a significant preference for negative values. In addition, a slight preference for negative values of $\mathcal{C}_{9'}^{\mathrm{NP}}$ or $\mathcal{C}_7^{\mathrm{NP}}$ occurs (with much less significance). It arises for $\mathcal{C}_{9'}^{\mathrm{NP}}$ when only large-recoil data is considered: $\mathcal{C}_{9'}^{\mathrm{NP}} < 0$ is favored to raise the value of $\langle P_5' \rangle_{[4.3,8.68]}$ without spoiling the agreement between theory and experiment in the first bin. This possibility is however weakened by the low-recoil data, and we will not consider this possibility any further. On the other hand, $\mathcal{C}_7^{\mathrm{NP}} < 0$ is also favored by the radiative decays (see e.g. Ref. [15]), and deserves further consideration.

C. ($\mathcal{C}_7^{\mathrm{NP}}, \mathcal{C}_9^{\mathrm{NP}}$) scenario

We focus on the implications of data for NP in $\mathcal{C}_7$ and $\mathcal{C}_9$. We perform a standard χ^2 fit to $\mathcal{C}_7^{\mathrm{NP}}, \mathcal{C}_9^{\mathrm{NP}}$, first including all the observables considered in Sec. II A, but taking only the first three large-recoil bins for $B \rightarrow K^* \mu^+ \mu^-$. The result is shown in Fig. 1, where 68.3% (red), 95.5% (green), and 99.7% (yellow) C.L. regions are shown. The best-fit point is obtained for $\mathcal{C}_9^{\mathrm{NP}} \sim -1.5$ and $\mathcal{C}_7^{\mathrm{NP}} \sim -0.02$. We stress that in this scenario, the SM hypothesis $\mathcal{C}_7^{\mathrm{NP}} = 0, \mathcal{C}_9^{\mathrm{NP}} = 0$ has a pull of 4.5σ.

The individual 68.3% C.L. (1σ) ranges for $\mathcal{C}_7^{\mathrm{NP}}$ and $\mathcal{C}_9^{\mathrm{NP}}$ are

$$\mathcal{C}_7^{\mathrm{NP}} = [-0.035, 0.000], \qquad \mathcal{C}_9^{\mathrm{NP}} = [-1.9, -1.3]. \quad (4)$$

Including the low-recoil bins decreases slightly the significance of the effect, since these observables are less constraining and compatible with the SM currently. The corresponding regions are indicated by the dashed curves in Fig. 1. The inclusion of the low-recoil data reduces the discrepancy with respect to the SM to 3.9σ. We notice that if the (first low-recoil) [14.18,16] bin is excluded from the analysis, we get similar significances analysing either
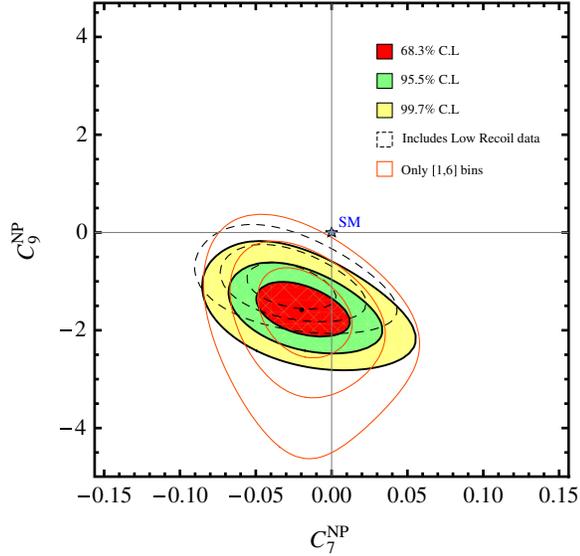
FIG. 1 (color online). Fit to $(\mathcal{C}_7^{NP}, \mathcal{C}_9^{NP})$, using the three large-recoil bins for $B \to K^* \mu^+ \mu^-$ observables, together with $B \to X_s \gamma$, $B \to X_s \mu^+ \mu^-$, $B \to K^* \gamma$ and $B_s \to \mu^+ \mu^-$. The dashed contours include both large- and low-recoil bins, whereas the orange (solid) ones use only the 1–6 GeV^2 bin for $B \to K^* \mu^+ \mu^-$ observables. The origin $\mathcal{C}_{7,9}^{NP} = (0, 0)$ corresponds to the SM values for the Wilson coefficients $\mathcal{C}_{7\text{eff},9}^{SM} = (-0.29, 4.07)$ at $\mu_b = 4.8$ GeV.

large-recoil data alone or large- and low-recoil data together, meaning that the reduction of significance stems mainly from the [14.18,16] bin, a region where the role of the newly found $\psi(4160)$ resonance has still to be understood [20]. In both analyses (with or without low-recoil data), P_2 and P_5' drive the fits away from the SM point, and fits including only one of these two observables for $B \to K^* \mu^+ \mu^-$, together with the other radiative and dileptonic decays, lead to contours in the $(\mathcal{C}_7^{NP}, \mathcal{C}_9^{NP})$ plane similar to Fig. 1.

We would like to understand whether this conclusion is due to peculiarities of individual bins. For this purpose we repeat the analysis restricting the input for the $B \to K^* \mu^+ \mu^-$ observables to $[1, 6]$ GeV^2 bins, exploiting several theoretical and experimental advantages. Such wider bins collect more events with larger statistics. Furthermore, some theoretical issues are less acute, such as the effect of low-mass resonances at very low $q^2 \lesssim 1$ GeV^2 [36], or the impact of charm loops above ~ 6 GeV^2 [37]. On the other hand, integrating over such a large bin washes out some effects related to the q^2 dependence of the observables, so that we expect this analysis to have less sensitivity to NP [15]. This can be seen in Fig. 1, where the regions in this case are indicated by the orange curves, and as expected the constraints get slightly weaker. In addition, due to the fact that theoretical uncertainties happen to increase moderately for large negative NP contributions to $\mathcal{C}_9$, the constraints are looser in the lower region of the $(\mathcal{C}_7^{NP}, \mathcal{C}_9^{NP})$ plane. We emphasize that even in this rather conservative

situation the main conclusion (a NP contribution $\mathcal{C}_9^{NP} \sim -1.5$) still prevails, whereas the SM hypothesis has still a pull of 3.2σ.

We illustrate the improvement gained by shifting $\mathcal{C}_9$ in Fig. 2, where we show the predictions for $\mathcal{C}_9^{NP} = -1.5$ (and other $\mathcal{C}_i^{NP} = 0$) for the observables P_2, P_4' and P_5', together with the experimental data and SM predictions. In particular, we observe how the various observables described in Sec. I change for $\mathcal{C}_9^{NP} < 0$. If the data is in general well reproduced in this scenario, there are still a few observables difficult to explain theoretically. Looking at Fig. 2, the most obvious cases are $\langle P_5' \rangle$ in the first and third bins. One can see there is a tension between these two bins: more negative values for $\mathcal{C}_9^{NP}$ reproduce better the third bin, but drive the first bin upwards, whose experimental value is consistent with the SM. A similar situation happens with the second and third bins of $\langle P_2 \rangle$, although in this case a good compromise is achieved.

Concerning the individual constraints to the fit, the large-recoil bins for P_2 and P_5' both favor the same large region away from the SM in the $(\mathcal{C}_7^{NP}, \mathcal{C}_9^{NP})$ plane, providing a negative correlation between $\mathcal{C}_7^{NP}$ and $\mathcal{C}_9^{NP}$. $B \to X_s \gamma$ selects values of $\mathcal{C}_7^{NP}$ close to the SM value, leading to the combined (smaller) region shown in Fig. 1. To be more quantitative, we have considered the pulls obtained by removing in turn one or two observables from the fit. We find that the largest pulls are associated to $\langle P_5' \rangle_{[4.3,8.68]}$, $B \to X_s \gamma$, $\langle P_2 \rangle_{[14.18,16]}$ and $\langle P_4' \rangle_{[14.18,16]}$. $B \to X_s \gamma$ has a large pull because it plays a very important role in disfavoring a scenario with large and negative $\mathcal{C}_7^{NP}$, which can mimic the $\mathcal{C}_9^{NP}$ scenario in $B \to K^* \mu^+ \mu^-$ observables. The observables $\langle P_5' \rangle_{[4.3,8.68]}$ and $\langle P_2 \rangle_{[14.18,16]}$ pull in different directions: the former favors more negative and the latter less negative values for $\mathcal{C}_9^{NP}$, while the best-fit point lies somewhat in the middle, with or without these observables. On the other hand $\langle P_4' \rangle_{[14.18,16]}$ has a marginal effect on the results of the fit.

The role of individual observables is confirmed by comparing our analysis with the preliminary results in Ref. [25], performed in the same framework, but with only P_1, P_2 and A_{FB} as inputs for $B \to K^* \mu^+ \mu^-$, leading to a 3σ deviation from the SM in the $(\mathcal{C}_7^{NP}, \mathcal{C}_9^{NP})$ plane (in our present analysis, this effect is magnified by the addition of $P_{4,5,6,8}'$ [20] among the observables). We emphasize the importance of choosing the right set of observables among the three correlated inputs A_{FB}, P_2, F_L: F_L has a very significant dependence on the choice of form factors (Fig. 5), which is less acute in the case of A_{FB} and P_2, so that the choices (F_L, P_2) or (F_L, A_{FB}) [38] lead to results that are more biased by the specific parametrization of form factors considered and less sensitive to NP compared to (A_{FB}, P_2) [25]. For this reason, we use A_{FB} instead of F_L in our analysis. We have checked by two different procedures (NLO QCD factorization and naive factorization) that the 3σ deviation reported in Ref. [25] using [1–6]
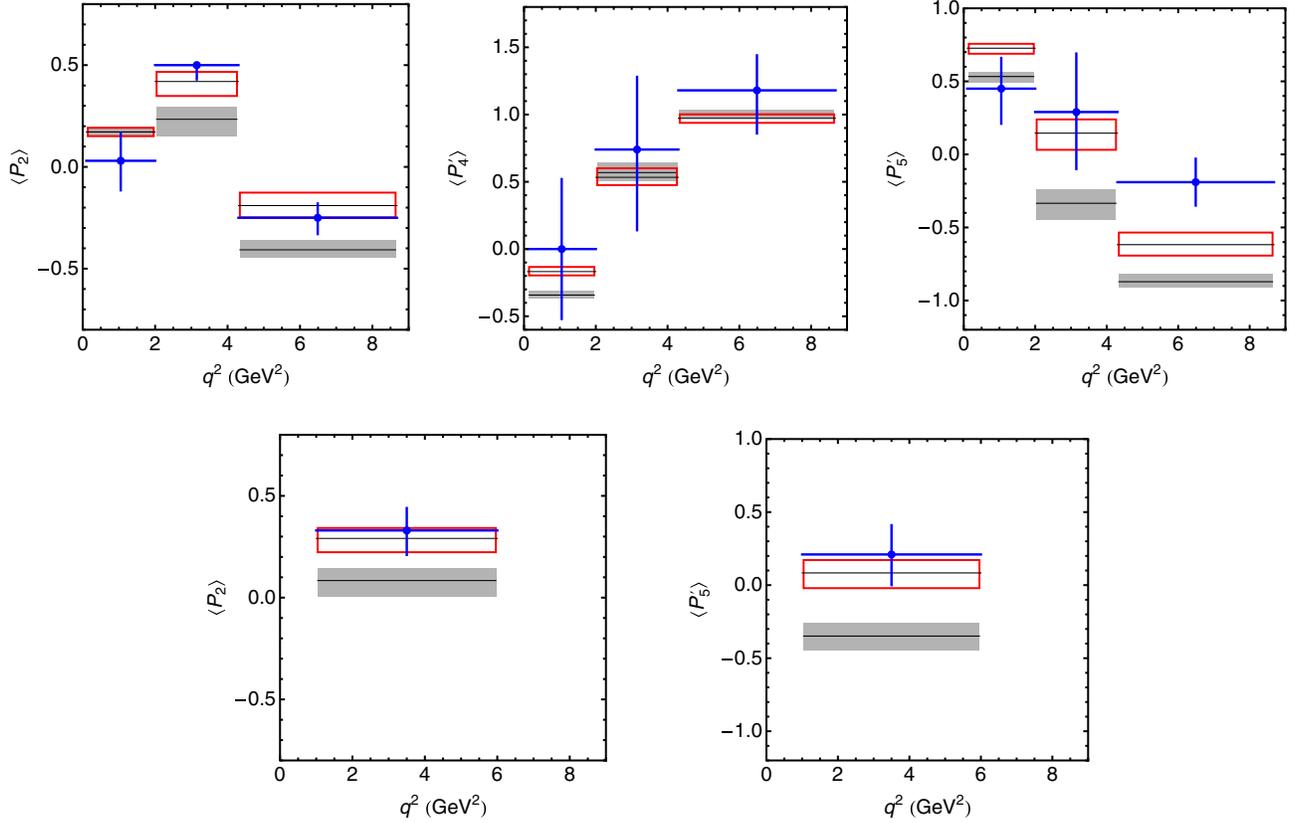
FIG. 2 (color online). Comparison between the SM predictions (gray boxes), the experimental measurements (blue) crosses and the predictions for the scenario with $\mathcal{C}_9^{\text{NP}} = -1.5$ and other $\mathcal{C}_i^{\text{NP}} = 0$ (red) empty squares.

bins gets reduced to around 1σ if F_L is used as an input instead of P_2 or A_{FB} (in agreement with Ref. [38], where F_L is used).

III. ROBUSTNESS OF THE RESULTS

In view of the results of the previous section, it is important to assess the robustness of the NP interpretation for the $B \rightarrow K^* \mu^+ \mu^-$ anomaly and how stable the conclusion $\mathcal{C}_9^{\text{NP}} < 0$ is, taking into account potential pollution from SM sources mimicking a negative $\mathcal{C}_9^{\text{NP}}$.

A. Charm loop

One of the key sources of uncertainty in the extraction of $\mathcal{C}_9$ from $B \rightarrow K^* \mu^+ \mu^-$ is related to the charm-loop contribution (subsequently decaying through a photon into a dilepton pair) coming from the insertion of 4-quark current-current ($\mathcal{O}_{1,2}^c$) or penguin operators ($\mathcal{O}_{3-6}$). The contributions from $\mathcal{O}_{1,2}^c$ are particularly important since the Wilson coefficients are numerically large and the processes are not Cabibbo-Kobayashi-Maskawa suppressed. This contribution can be described through a short-distance (perturbative) contribution, which exhibits a noticeable sensitivity to the value of m_c near the threshold of $c\bar{c}$ production, and a long-distance (nonperturbative) contribution which is difficult to assess.

The perturbative charm-loop contribution is usually absorbed into the definition of $\mathcal{C}_9^{\text{eff}}(q^2) = \mathcal{C}_9 + Y(q^2)$ [22] and is given at leading order by

$$Y^c(q^2, m_c) = -\frac{4}{27}(4\mathcal{C}_1 + 3\mathcal{C}_2 + 18\mathcal{C}_3 + 180\mathcal{C}_5)$$
$$\times \left[\ln\frac{m_c^2}{\mu^2} - \frac{2}{3} - z + (2+z)\right.$$
$$\left.\times \sqrt{|z-1|}\text{arccot}\sqrt{(z-1)}\right], \qquad (5)$$

where $z = 4m_c^2/q^2$. There is a threshold at $q^2 = 4m_c^2 \simeq 6\text{ GeV}^2$, above which Eq. (5) must be continued analytically and an imaginary part is generated. The real part exhibits a cusp at this threshold, whose exact position depends on m_c. There is a significant variety of choices in the literature concerning the value of m_c for such computation, for instance the pole mass (around 1.4 GeV) [22], the $\overline{\text{MS}}$ mass at the scale $\mu = m_c$ (around 1.27 GeV) [39] or the same mass at the scale $\mu = 2m_c$ (around 1 GeV) [37]. Following Ref. [21], we take the second option and perform the computation of $B \rightarrow K^* \mu^+ \mu^-$ observables with a reference value $\bar{m}_c = 1.27\text{ GeV}$. We can study the dependence on m_c by reinterpreting its change as a shift in the value of $\mathcal{C}_9$, given by

$$\delta \mathcal{C}_9^{c\bar{c},\text{pert}} = \text{Re}[Y^c(q^2, m_c) - Y^c(q^2, \bar{m}_c)]. \quad (6)$$

The same analysis can be performed for the imaginary part, which vanishes below the charm threshold, with similar conclusions.

The long-distance contribution is difficult to estimate, and we are not aware of a systematic approach able to compute this correction from first principles. Part of this effect is taken into account directly at the level of the experimental analysis, through the removal of charmonium contributions. From the theoretical side, one can exploit the light-cone sum-rule computation in Ref. [37], performed below the charm threshold and extrapolated up to the J/ψ pole at $q^2 \sim 9.4$ GeV^2. These long-distance contributions can be recast as an effective contribution to $\mathcal{C}_9$, but they depend on the transversity amplitude considered (contrary to the perturbative case). Even though one could in principle distinguish this effect from a "universal" NP contribution $\mathcal{C}_9^{\text{NP}}$ by comparing observables with different sensitivities to the three transversity amplitudes, this cannot be achieved with the current uncertainties. In particular, the results quoted in Ref. [37] for each amplitude are compatible within errors. Therefore we consider a universal correction to $\mathcal{C}_9$ arising from the long-distance charm-loop contribution, that we parametrize as

$$\delta \mathcal{C}_9^{c\bar{c},\text{LD}} = \frac{a + bq^2(c - q^2)}{q^2(c - q^2)}. \quad (7)$$

The parameters a, b, c can be obtained by imposing that the sum of the perturbative contribution Y^c and the non-perturbative $\delta \mathcal{C}_9^{c\bar{c},\text{LD}}$ contributions reproduce the results in Ref. [37] below the charm threshold [correcting for the different reference value of the charm-quark mass $\bar{m}_c(m_c) = 1.27$ GeV used here]. The parameter c is chosen to recover the J/ψ pole, leading to $a \sim 5$, $b \sim 0.25$ and $c \sim 9.5$.

In Fig. 3 we show the size and q^2 dependence of the perturbative and nonperturbative charm-loop contributions in Eqs. (6) and (7). The results of Sec. II were obtained taking $Y^c(\bar{m}_c)$ and neglecting nonperturbative contributions. To keep a good agreement with data, an additional $\delta \mathcal{C}_9^{c\bar{c}}$ correction must be compensated by a shift in the NP contribution $\mathcal{C}_9^{\text{NP}}$ compared to the values obtained previously. A positive $\delta \mathcal{C}_9^{c\bar{c}}$ means that $\mathcal{C}_9^{\text{NP}}$ should be enhanced (more negative), whereas a negative $\delta \mathcal{C}_9^{c\bar{c}}$ requires $\mathcal{C}_9^{\text{NP}}$ to be reduced in magnitude (less negative). Several comments are in order. First, one can see that the region between 1 and 6 GeV^2 is little affected by the long-distance contribution, and gets either an enhancement of NP effects in $\mathcal{C}_9^{\text{NP}}$ (for $m_c = 1$ GeV) or a very small depletion (for $m_c = 1.4$ GeV) from the short-distance part. In the case of the [4.3,8.68] bin, one has to remember that the negative contribution $\delta \mathcal{C}_9^{c\bar{c},\text{pert}}$ when varying m_c from 1.27 to 1.4 GeV has to be integrated over the whole bin, leading to a shift of at most $+0.3$ for $\mathcal{C}_9^{\text{NP}}$ from this bin (one finds a contribution
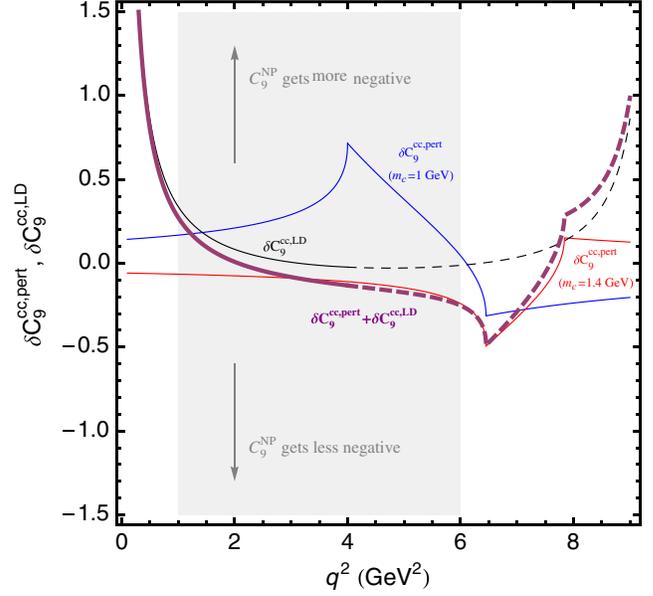
FIG. 3 (color online). Estimate of possible effective corrections to $\mathcal{C}_9$ from the m_c dependence of short-distance ($\delta \mathcal{C}_9^{c\bar{c},\text{pert}}$) and long-distance charm-loop effects ($\delta \mathcal{C}_9^{c\bar{c},\text{LD}}$). The thick line shows a (worst-case scenario) combination of both effects. A positive contribution here enhances a negative NP contribution to $\mathcal{C}_9$, while a negative contribution here means $\mathcal{C}_9^{\text{NP}}$ is less negative by the same amount. Note that these effects cannot mimic a NP contribution as large as $\mathcal{C}_9^{\text{NP}} \sim -1.5$.

of similar size for the imaginary part of $\mathcal{C}_9$, which is however more difficult to interpret directly in terms of a shift of $\mathcal{C}_9^{\text{NP}}$, taken real here).

We see therefore that (a) charm-loop effects affect very marginally our analysis using $B \to K^* \mu^+ \mu^-$ observables in the [1,6] bin only, and (b) these long-distance charm contributions will tend to enhance (rather than reduce) the need for a negative $\mathcal{C}_9^{\text{NP}}$ for the analysis including thinner bins, closer to the photon and J/ψ poles.

We have checked the sensitivity of our analysis to these effects (charm-loop effects and charm-mass dependence) by considering the least favorable situation for NP, where the charm-quark mass is shifted from 1.27 to 1.4 GeV and long-distance contributions are neglected. The need for $\mathcal{C}_9^{\text{NP}} \neq 0$ remains, with a slight decrease in significance (by less than 1σ) and a small shift of the best-fit value for $\mathcal{C}_9^{\text{NP}}$ (by $\simeq +0.3$, in agreement with our previous discussion).

B. Alternative approaches to factorization

A second issue consists in the sensitivity to the framework used to describe $B \to K^* \mu^+ \mu^-$ form factors and amplitudes. Following Refs. [7,21,22], we have adopted the QCD factorization framework. One issue still under discussion is the uncertainty coming from Λ/m_b corrections. As a check of the robustness of our results with respect to this issue, we have increased the uncertainties coming from Λ/m_b corrections by 3, resulting in a slight

decrease of the pull for the SM hypothesis in the $(\mathcal{C}_7, \mathcal{C}_9)$ scenario from 4.5σ to 4σ for the analysis based on the 3 large-recoil bins of $B \to K^* \mu^+ \mu^-$ (multiplying the same uncertainties by 6 decreases the significance down to 3σ).

If our uncertainties stemming from Refs. [7,21,22] are fairly standard, one can find alternative estimates for instance in Ref. [36], where larger uncertainties for P_i' at large recoil are obtained due to significantly larger power-suppressed corrections. The size of these power corrections and their q^2 dependence was obtained from the spread between different estimates of the form factors (several sum-rule computations in different settings and Dyson-Schwinger equations). If this approach is certainly conservative, it mixes frameworks with very different levels of accuracy, q^2 ranges of validity, and correlations between the various form factors. Indeed, as noticed in Ref. [36], the form factors are estimated in a conventional basis which is different from the helicity basis needed for the computation of the transversity amplitudes. Our lack of knowledge concerning the correlations among these form factors increases the uncertainties on the helicity form factors very significantly. In our mind, averaging different parametrizations (some of them not being updated to their latest values) given in this conventional basis tends to overestimate the power-suppressed corrections in the helicity form factors. However, we agree with Ref. [36] that a first-principle computation of the helicity form factors would be the best way to improve the accuracy on this type of systematics.

One could also consider the alternative approach of naive factorization following, e.g., Ref. [4]: the $B \to K^* \mu^+ \mu^-$ form factors are then treated as independent and not reduced to two form factors $\xi_{\perp,\parallel}$, and contributions to the transversity amplitudes from hard-gluon interactions (in particular, spectator interaction) are neglected. As shown in Ref. [4], these two approaches yield slightly shifted results for the q^2 dependence of some observables, but they agree on their qualitative variations. We have performed fits identical to the ones presented in Sec. II C, with very similar results (preference for $\mathcal{C}_9^{\mathrm{NP}} < 0$, SM hypothesis with a large pull but less significance for the large-recoil data), indicating that our results are robust with respect to the use of naive or NLO QCD factorization.

IV. CONCLUSIONS AND PERSPECTIVES

We have combined the recent LHCb measurements of $B \to K^* \mu^+ \mu^-$ observables [19,20] with other radiative modes in a fit to Wilson coefficients, using the framework of our previous works [15,21]. We have found a strong indication for a negative NP contribution to the coefficient $\mathcal{C}_9$, at 4.5σ using large-recoil data (3.9σ using both large- and low-recoil data). Our results correspond to $\mathcal{C}_9$ inside a 68% C.L. range $2.2 \leq \mathcal{C}_9 \leq 2.8$ to be compared with $\mathcal{C}_9^{\mathrm{SM}} = 4.07$ at the scale $\mu_b = 4.8$ GeV. This is the main conclusion of our analysis of LHCb $B \to K^* \mu^+ \mu^-$ measurements.

We also observe a slight preference for negative values in $\mathcal{C}_7^{\mathrm{NP}}$ (mainly driven by radiative constraints), but no clear-cut evidence for $\mathcal{C}_{7',9',10'}^{\mathrm{NP}} \neq 0$. The situation for $\mathcal{C}_{7'}$ could be clarified with a substantial reduction of error bars in P_1, whereas the case for $\mathcal{C}_{9'}^{\mathrm{NP}} < 0$ is supported by the data at large recoil, but not at low recoil (an illustration of the improvement brought by $\mathcal{C}_{9'}^{\mathrm{NP}} < 0$ at large recoil is shown in Fig. 4).

We emphasise that the same mechanism $\mathcal{C}_9^{\mathrm{NP}} < 0$ ($\mathcal{C}_7^{\mathrm{NP}} < 0$) is remarkably efficient and economical in explaining the whole pattern of deviations indicated by the recent LHCb results [19,20]: small discrepancies in A_{FB}, tensions in P_2 and P_5', preference for a higher position of the zero of A_{FB} and P_2.

If this pattern is confirmed, it remains to determine what kind of NP could accommodate a situation where $\mathcal{C}_9^{\mathrm{NP}} < 0$ is large, with the possibility of additional contributions to $\mathcal{C}_7^{\mathrm{NP}}$ and $\mathcal{C}_{9'}^{\mathrm{NP}}$. A natural possibility would consist in a Z' gauge boson [40], coupling to left-handed quarks only but equally to left- and right-handed charged leptons, and with flavor-changing couplings to down-type quarks. This category of models and the flavor constraints set on them were discussed in Ref. [41], even though the case discussed here (i.e., in their notation, $\Delta_L^{\mu\mu} = \Delta_R^{\mu\mu}$, $\Delta_R^{sb} \simeq 0$ and Δ_L^{sb} with the same phase as $V_{tb}V_{ts}^*$) was not considered specifically in this reference. In the minimal scenario, the Z' gauge boson would contribute to Δm_s, but not to the $B_s - \bar{B}_s$ mixing phase. As an illustration, such a boson with $M_{Z'} = 1$ TeV could yield $\mathcal{C}_9^{\mathrm{NP}} \sim -1.5$ and be compatible with Δm_s (using the values in Ref. [41]), provided that its coupling to muons $\Delta_L^{\mu\mu} = \Delta_R^{\mu\mu}$ is of order 0.1. It would
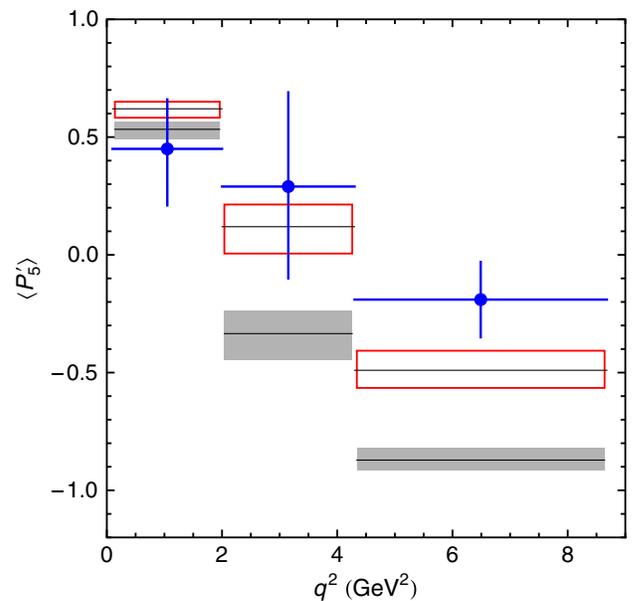
FIG. 4 (color online). Improvement in the q^2 dependence of P_5' in the illustrative case $\mathcal{C}_9^{\mathrm{NP}} = \mathcal{C}_{9'}^{\mathrm{NP}} = -1.5$ (and NP contributions to the other Wilson coefficients set to zero).

be worth checking if such a scenario could be framed in a specific model showing a good agreement with all the available $\Delta F = 2$ measurements.

Obviously, this analysis is only a starting point, illustrating the potential of $B \to K^* \mu^+ \mu^-$ observables to unveil interesting patterns of NP through rare radiative decays, and many improvements and confirmations should be gathered before reaching a definitive conclusion on the situation discussed here.

The main improvement for the analysis concerns the reduction of uncertainties. On the experimental side, the LHCb results [19,20] are based on 1/fb of data collected in 2011. Adding the 2012 data (another 2/fb already on disk) will constitute a big improvement concerning statistical uncertainties. In addition, one should be able to improve on some of the primary observables P_i, as well as on the determination of the S-wave pollution, by using new foldings (see the Appendix). In order to avoid the potential pollution of C_9 from charm-loop effects, it is essential that the LHCb experiment provides future results for $B \to K^* \mu^+ \mu^-$ with a finer q^2 binning, with several narrow bins between 1 and 6 GeV^2, and a summary of the correlations among the various measurements. On the theoretical side, since C_9 seems to be the main Wilson coefficient affected by NP, charm-loop effects become a very important issue, with several questions left open. It would be very useful for the theorists to converge on the scheme and the scale of the perturbative charm-quark contribution, as well as to provide alternative and/or improved estimates of the long-distance contribution obtained in Ref. [37], in particular above the charm threshold. Another significant source of uncertainties comes from the form factors, for which new lattice results should bring more control on the low-recoil region [42,43]. In order to decrease the uncertainty attached to the form factors, it will also become essential that their estimates are provided including correlations, or in the basis of helicity form factors discussed in Refs. [36,44].

An essential aspect consists in cross-checking and confirming the results from $B \to K^* \mu^+ \mu^-$ on C_9^{NP} through other channels accessible to LHCb and with good prospects of improving on our knowledge of the form factors. The $B \to K \mu^+ \mu^-$ decay [45] gives a linear constraint between C_7 and C_9 involving pseudoscalar-to-pseudoscalar form factors well suited for lattice simulations [42,46,47]. The $B_s \to \phi \mu^+ \mu^-$ [48] has the same potential as $B \to K^* \mu^+ \mu^-$ in terms of angular observables, with the added interest of a very narrow ϕ final state, avoiding the difficult simulation of wide resonances on the lattice. Finally, the $\Lambda_b \to \Lambda \mu^+ \mu^-$ decay [49,50] is also a useful cross-check, with a different angular structure [51] and very recent estimates for the relevant form factors [52–55].

In all these aspects, a deep interplay between experimental and theoretical analyses will prove essential to confirm the pattern of NP suggested by the $B \to K^* \mu^+ \mu^-$ anomaly.

APPENDIX: TAMING *S*-WAVE POLLUTION

In this section we focus on the S-wave pollution coming from the companion decay $B \to K_0^* \mu^+ \mu^-$ [56–58] with a threefold aim. First, we provide the explicit form of the S-wave polluting terms entering the different foldings allowing one to extract the primary observables. This might help in estimating the systematics if the bounds found in Ref. [21] are used. Second, we present three new foldings to disentangle the contributions from P_1 alone, $P_{1,2}$ and $P_{1,3}$. Separating P_1 from the other observables can be useful to sharpen the size of its error bars. Third, we show that certain combinations of folded distributions can be sensitive to the interesting observables free from S-wave pollution. This could be combined with other approaches to the S-wave contribution[4] to reduce systematic uncertainties of the experimental analysis.

Our starting point is the distribution given in Eq. (43) of Ref. [21] using the definitions for the angular variables defined there and working in the massless lepton limit (see Ref. [58] for the massive case). We determine the S-wave pollution accompanying the relevant folded distributions that can be used to extract the primary observables:

(i) $P_{1,2,3}$: Identifying $\phi \leftrightarrow \phi + \pi$ (when $\phi < 0$) amounts to the folded distribution $d\hat{\Gamma} = d\Gamma(\hat{\phi}) + d\Gamma(\hat{\phi} - \pi)$. Once normalized to the full distribution the result (see Sec. 7 of [21] for definitions) is $d\hat{\Gamma}/\Gamma_{\mathrm{full}} = \frac{x}{2} + \frac{9}{16\pi} F_T \sin^2 \hat{\theta}_K (2P_2 \cos \hat{\theta}_\ell - P_3 \sin 2\hat{\phi} \sin^2 \hat{\theta}_\ell)(1 - F_S) + \delta_{sw}^1$ where $x = \frac{9}{32\pi} (8F_L \cos^2 \hat{\theta}_K \sin^2 \hat{\theta}_\ell + F_T \sin^2 \hat{\theta}_K (3 + \cos 2\hat{\theta}_\ell) + 2F_T P_1 \sin^2 \hat{\theta}_K \cos 2\hat{\phi} \sin^2 \hat{\theta}_\ell)(1 - F_S)$ and $\delta_{sw}^{(1)} = \frac{3}{8\pi}(F_S + A_S \cos \hat{\theta}_K)\sin^2 \hat{\theta}_\ell$ stands for the S-wave contribution. The folded angle $\hat{\phi}$ is defined in the range $\hat{\phi} \in [0, \pi]$ and all other angles are inside their usual range $\hat{\theta}_l$, $\hat{\theta}_K \in [0, \pi]$. Other foldings providing subsets of $P_{1,2,3}$ can be found in Table III.

(ii) P_4': A double folded distribution is necessary $d\hat{\Gamma} = d\Gamma(\hat{\phi}) + d\Gamma(-\hat{\phi}) + d\Gamma(\pi - \hat{\phi}, \pi - \hat{\theta}_\ell) + d\Gamma(\hat{\phi} - \pi, \pi - \hat{\theta}_\ell)$ and once normalized $d\hat{\Gamma}/\Gamma_{\mathrm{full}} = x + \frac{9}{16\pi} \sqrt{F_T F_L} P_4' \cos \hat{\phi} \sin 2\hat{\theta}_K \sin 2\hat{\theta}_\ell$

---

[4]See Ref. [59] for an approach to the S-wave problem using chiral perturbation theory and dispersion relations and Ref. [60] for an experimental analysis of the S-wave impact on $B \to K^* \mu^+ \mu^-$.
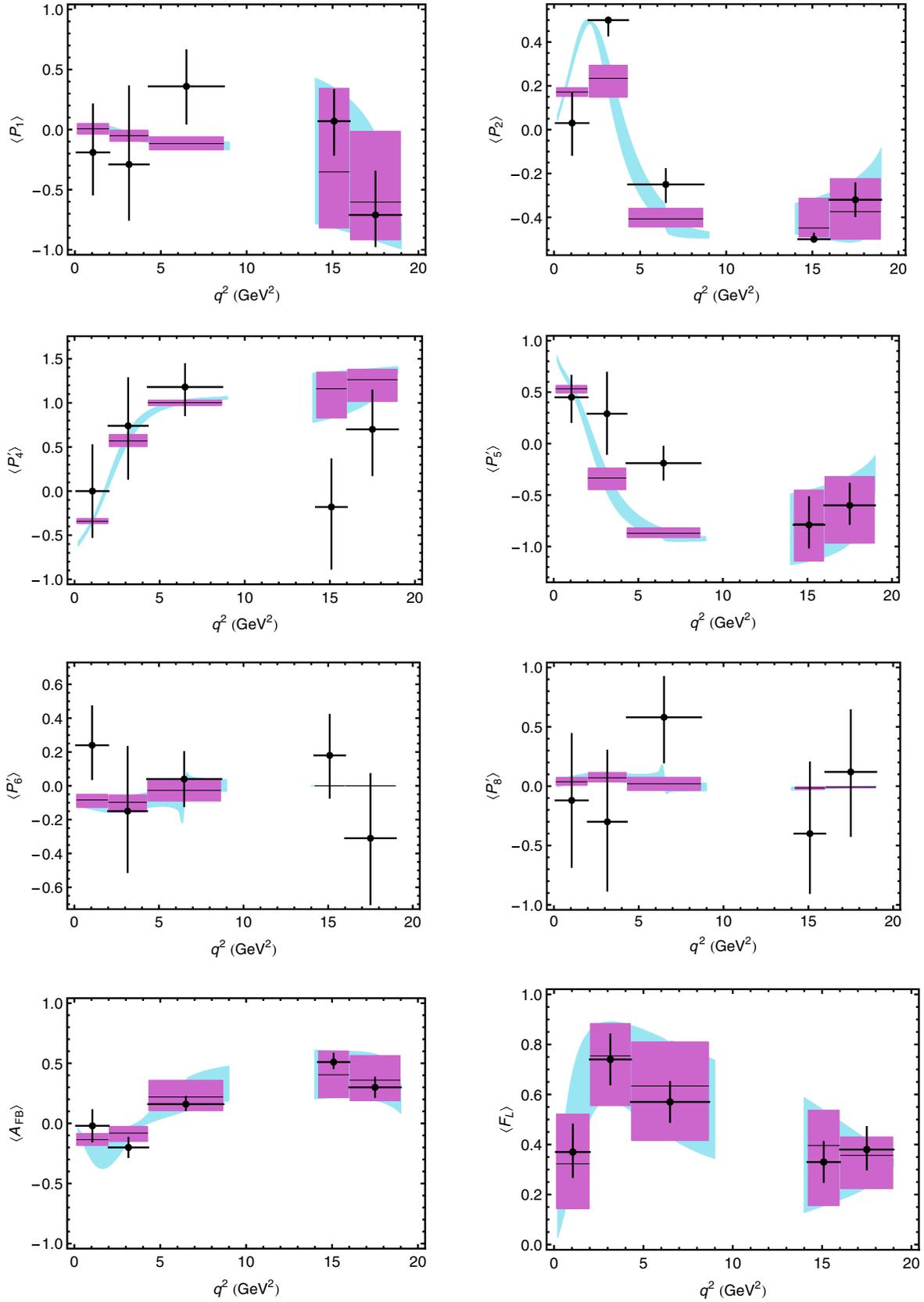
FIG. 5 (color online). Experimental measurements and SM predictions for some $B \to K^* \mu^+ \mu^-$ observables. The black crosses are the experimental LHCb data. The blue band corresponds to the SM predictions for the differential quantities, whereas the purple boxes indicate the corresponding binned observables.

TABLE III. Foldings needed to single out the interesting observables, with the corresponding remaining S-wave pollution. For all foldings, $\hat{\theta}_\ell$ and $\hat{\theta}_K$ lie within $[0, \pi/2]$, whereas $\hat{\phi}$ has different ranges depending on the observables considered.

Observables	S-wave	Folding	$\hat{\phi}$ range
$P_{1,2,3}$	A_s	$d\Gamma(\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\hat{\phi} - \pi, \hat{\theta}_l, \hat{\theta}_K)$	$[0, \pi]$
P_1	A_{s5}, A_{s8}	$d\Gamma(\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\hat{\phi}, \hat{\theta}_l, \pi - \hat{\theta}_K) + d\Gamma(-\hat{\phi}, \pi - \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(-\hat{\phi}, \pi - \hat{\theta}_l, \pi - \hat{\theta}_K)$	$[0, \pi]$
P_1 and P_2	A_{s4}, A_{s5}	$d\Gamma(\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\hat{\phi}, \hat{\theta}_l, \pi - \hat{\theta}_K) + d\Gamma(-\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(-\hat{\phi}, \hat{\theta}_l, \pi - \hat{\theta}_K)$	$[0, \pi]$
P_1 and P_3	A_{s5}, A_{s7}	$d\Gamma(\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\hat{\phi}, \hat{\theta}_l, \pi - \hat{\theta}_K) + d\Gamma(\hat{\phi}, \pi - \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\hat{\phi}, \pi - \hat{\theta}_l, \pi - \hat{\theta}_K)$	$[0, \pi]$
P_1 and P_4'	A_{s5}	$d\Gamma(\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(-\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\hat{\phi}, \pi - \hat{\theta}_l, \pi - \hat{\theta}_K) + d\Gamma(-\hat{\phi}, \pi - \hat{\theta}_l, \pi - \hat{\theta}_K)$	$[0, \pi]$
P_1 and P_5'	A_s, A_{s5}	$d\Gamma(\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(-\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\hat{\phi}, \pi - \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(-\hat{\phi}, \pi - \hat{\theta}_l, \hat{\theta}_K)$	$[0, \pi]$
P_1 and P_6'	A_s, A_{s7}	$d\Gamma(\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\pi - \hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\hat{\phi}, \pi - \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\pi - \hat{\phi}, \pi - \hat{\theta}_l, \hat{\theta}_K)$	$[-\pi/2, \pi/2]$
P_1 and P_8'	A_{s7}	$d\Gamma(\hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\pi - \hat{\phi}, \hat{\theta}_l, \hat{\theta}_K) + d\Gamma(\hat{\phi}, \pi - \hat{\theta}_l, \pi - \hat{\theta}_K) + d\Gamma(\pi - \hat{\phi}, \pi - \hat{\theta}_l, \pi - \hat{\theta}_K)$	$[-\pi/2, \pi/2]$

$(1 - F_S) + \delta_{sw}^{(2)}$ where $\delta_{sw}^{(2)} = 2\delta_{sw}^{(1)} + \frac{3}{4\pi}A_S^4 \cos\hat{\phi}\sin\hat{\theta}_K \sin2\hat{\theta}_\ell$ with $\hat{\phi} \in [0, \pi]$, $\hat{\theta}_\ell \in [0, \pi/2]$, $\hat{\theta}_K \in [0, \pi]$. Another possible folding leading to P_4' can be found in Table III.

(iii) P_5': A double folded distribution is also required $d\hat{\Gamma} = d\Gamma(\hat{\phi}) + d\Gamma(-\hat{\phi}) + d\Gamma(\hat{\phi}, \pi - \hat{\theta}_\ell) + d\Gamma(-\hat{\phi}, \pi - \hat{\theta}_\ell)$ leading to $d\hat{\Gamma}/\Gamma_{\text{full}} = x + \frac{9}{8\pi}\sqrt{F_T F_L}P_5' \cos\hat{\phi}\sin2\hat{\theta}_K \sin\hat{\theta}_\ell(1 - F_S) + \delta_{sw}^{(3)}$ where $\delta_{sw}^{(3)} = 2\delta_{sw}^{(1)} + \frac{3}{4\pi}A_S^5 \cos\hat{\phi}\sin\hat{\theta}_K \sin\hat{\theta}_\ell$. Here $\hat{\phi} \in [0, \pi]$, $\hat{\theta}_\ell \in [0, \pi/2]$ and $\hat{\theta}_K \in [0, \pi]$.

The explicit form of the folding in terms of distributions for the last two primary observables $P_{6,8}'$ depends on the region in parameter space $(\phi, \theta_\ell, \theta_K)$ chosen. We provide here the folded distributions for P_6' and P_8' in the region $\hat{\phi} \in [0, \pi/2]$ (a similar folding can be obtained for $\hat{\phi} \in [-\pi/2, 0]$).

(i) P_6': The associated folded distribution is $d\hat{\Gamma} = d\Gamma(\hat{\phi}) + d\Gamma(\pi - \hat{\phi}) + d\Gamma(\hat{\phi}, \pi - \hat{\theta}_\ell) + d\Gamma(\pi - \hat{\phi}, \pi - \hat{\theta}_\ell)$ corresponding to $d\hat{\Gamma}/\Gamma_{\text{full}} = x - \frac{9}{8\pi} \times \sqrt{F_T F_L}P_6' \sin\hat{\phi}\sin2\hat{\theta}_K \sin\hat{\theta}_\ell(1 - F_S) + \delta_{sw}^{(4)}$ where $\delta_{sw}^{(4)} = 2\delta_{sw}^{(1)} + \frac{3}{4\pi}A_S^7 \sin\hat{\phi}\sin\hat{\theta}_K \sin\hat{\theta}_\ell$ and $\hat{\theta}_\ell \in [0, \pi/2]$, $\hat{\theta}_K \in [0, \pi]$.

(ii) P_8': The folded distribution is in this case $d\hat{\Gamma} = d\Gamma(\hat{\phi}) + d\Gamma(\pi - \hat{\phi}) + d\Gamma(\pi - \hat{\theta}_\ell, \pi - \hat{\theta}_K) + d\Gamma(\pi - \hat{\phi}, \pi - \hat{\theta}_\ell, \pi - \hat{\theta}_K) = x + \frac{9}{16\pi}\sqrt{F_T F_L}P_8' \sin\hat{\phi}\sin2\hat{\theta}_K \sin2\hat{\theta}_\ell(1 - F_S) + \delta_{sw}^5$ where $\delta_{sw}^5 = \frac{3}{4\pi}F_S\sin^2\hat{\theta}_\ell + \frac{3}{4\pi}A_S^7 \sin\hat{\phi}\sin\hat{\theta}_K \sin\hat{\theta}_\ell$ with $\hat{\theta}_\ell \in [0, \pi/2]$, $\hat{\theta}_K \in [0, \pi]$.

We summarize this discussion in Table III presenting some of the foldings already discussed, and other possibilities, with their sensitivity to primary observables and S-wave polluting terms (besides F_S).

Given that the present statistics does not allow one to fit all S-wave coefficients, we suggest combinations of folded distributions sensitive to the interesting observables P_2 and $P_{4,5}'$ free from S-wave pollution (besides a global $1 - F_S$ factor), but at the price of reducing the experimental sensitivity. We provide in the following two examples of this approach, one for P_2 and one for $P_{4,5}'$.

First, it was found in Ref. [58] that the identification $\phi \leftrightarrow \phi + \pi$ and $\theta_\ell \leftrightarrow \pi - \theta_\ell$ leads to a folding similar to the first folding in Table III but with twice the S-wave term. In this way the combination of distributions $d\hat{\Gamma} = d\Gamma(\hat{\phi}, \pi - \hat{\theta}_\ell) + d\Gamma(\hat{\phi} - \pi, \pi - \hat{\theta}_\ell) - d\Gamma(\hat{\phi}) - d\Gamma(\hat{\phi} - \pi)$ once normalized allows a direct measurement of P_2, i.e., $d\hat{\Gamma}/\Gamma_{\text{full}} = -\frac{9}{4\pi}F_T P_2 \cos\hat{\theta}_\ell\sin^2\hat{\theta}_K(1 - F_S)$.

The second example combines the two distributions for $P_{4,5}'$ described at the beginning of this section with three distributions given by Eqs. (28), (41) and (42) in Ref. [58]. In this case the resulting combination is $d\hat{\Gamma} = d\Gamma(-\hat{\phi}) + d\Gamma(-\hat{\phi} + \pi, \pi - \hat{\theta}_\ell) + d\Gamma(\hat{\phi} - \pi, \pi - \hat{\theta}_K) + d\Gamma(-\hat{\phi}, \pi - \hat{\theta}_\ell) - d\Gamma(\hat{\phi} - \pi) - d\Gamma(-\hat{\phi}, \pi - \hat{\theta}_K) - d\Gamma(\pi - \hat{\theta}_\ell) - d\Gamma(\hat{\phi} - \pi, \pi - \hat{\theta}_\ell)$ leading to $d\hat{\Gamma}/\Gamma_{\text{full}} = \frac{9}{4\pi}\sin\hat{\theta}_K \cos\hat{\phi} \sin\hat{\theta}_\ell(\sqrt{F_L F_T}(P_5' + P_4'\cos\hat{\theta}_\ell)\cos\hat{\theta}_K + F_T P_3 \sin\hat{\phi}\sin\hat{\theta}_K \sin\hat{\theta}_\ell)(1 - F_S)$.

Finally, we would like to point out that one can also tame the S-wave contribution using the bounds on the A_S^i coefficients presented in Ref. [21], which however require a measurement of F_S. This can be achieved through the folded distribution $d\hat{\Gamma} = d\Gamma(\hat{\phi}) + d\Gamma(\pi - \hat{\theta}_K) + d\Gamma(\hat{\phi} - \pi) + d\Gamma(\hat{\phi} - \pi, \pi - \hat{\theta}_K)$ where only F_S enters as an S-wave pollution $d\hat{\Gamma}/\Gamma_{\text{full}} = x + \frac{1}{4\pi} \times [9F_T\sin^2\hat{\theta}_K(P_2 \cos\hat{\theta}_\ell - P_3 \cos\hat{\phi}\sin\hat{\phi}\sin^2\hat{\theta}_\ell)](1 - F_S) + \frac{3}{4\pi}F_S\sin^2\hat{\theta}_\ell$.

[1] F. Kruger and J. Matias, Phys. Rev. D **71**, 094009 (2005).

[2] U. Egede, T. Hurth, J. Matias, M. Ramon, and W. Reece, J. High Energy Phys. 11 (2008) 032.

[3] E. Lunghi and J. Matias, J. High Energy Phys. 04 (2007) 058.

[4] W. Altmannshofer, P. Ball, A. Bharucha, A. J. Buras, D. M. Straub, and M. Wick, J. High Energy Phys. 01 (2009) 019.

[5] U. Egede, T. Hurth, J. Matias, M. Ramon, and W. Reece, J. High Energy Phys. 10 (2010) 056.

[6] D. Becirevic and E. Schneider, Nucl. Phys. **B854**, 321 (2012).

[7] J. Matias, F. Mescia, M. Ramon, and J. Virto, J. High Energy Phys. 04 (2012) 104.

[8] C. Bobeth, G. Hiller, and D. van Dyk, J. High Energy Phys. 07 (2010) 098.

[9] C. Bobeth, G. Hiller, and D. van Dyk, Phys. Rev. D **87**, 034016 (2013).

[10] A. K. Alok, A. Datta, A. Dighe, M. Duraisamy, D. Ghosh, and D. London, J. High Energy Phys. 11 (2011) 121.

[11] A. K. Alok, A. Datta, A. Dighe, M. Duraisamy, D. Ghosh, and D. London, J. High Energy Phys. 11 (2011) 122.

[12] W. Altmannshofer, P. Paradisi, and D. M. Straub, J. High Energy Phys. 04 (2012) 008.

[13] W. Altmannshofer and D. M. Straub, J. High Energy Phys. 08 (2012) 121.

[14] F. Beaujean, C. Bobeth, D. van Dyk, and C. Wacker, J. High Energy Phys. 08 (2012) 030.

[15] S. Descotes-Genon, J. Matias, M. Ramon, and J. Virto, J. High Energy Phys. 01 (2013) 048.

[16] CDF Collaboration, Phys. Rev. Lett. **108**, 081807 (2012).

[17] Belle Collaboration, Phys. Rev. Lett. **103**, 171801 (2009).

[18] *BABAR* Collaboration, Phys. Rev. D **86**, 032012 (2012).

[19] LHCb Collaboration, J. High Energy Phys. 08 (2013) 131.

[20] Report No. LHCb-PAPER-2013-037 (to be published); see also N. Serra, "Studies of Electroweak Penguin Transitions of $b \to s\mu\mu$" EPS-HEP Conference, Stockholm, 2013 (unpublished).

[21] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, J. High Energy Phys. 05 (2013) 137.

[22] M. Beneke, T. Feldmann, and D. Seidel, Nucl. Phys. **B612**, 25 (2001).

[23] M. Beneke, T. Feldmann, and D. Seidel, Eur. Phys. J. C **41**, 173 (2005).

[24] B. Grinstein and D. Pirjol, Phys. Rev. D **70**, 114005 (2004).

[25] J. Matias, "Optimizing the Basis of $B \to K^*\ell^+\ell^-$ Observables and Understanding its Tensions," EPS-HEP Conference, Stockholm, 2013 (unpublished).

[26] Y. Amhis *et al.* (Heavy Flavor Averaging Group Collaboration), arXiv:1207.1158.

[27] S. Descotes-Genon, D. Ghosh, J. Matias, and M. Ramon, J. High Energy Phys. 06 (2011) 099.

[28] T. Huber, T. Hurth, and E. Lunghi, Nucl. Phys. **B802**, 40 (2008).

[29] LHCb Collaboration, Phys. Rev. Lett. **110**, 021801 (2013).

[30] LHCb Collaboration, Phys. Rev. Lett. **111**, 101805 (2013).

[31] CMS Collaboration, arXiv:1307.5025.

[32] S. Hansmann-Menzemer, "Experimental Results on Flavour Physics," EPS-HEP Conference, Stockholm, 2013 (unpublished).

[33] LHCb Collaboration, J. High Energy Phys. 07 (2012) 133.

[34] LHCb Collaboration, Phys. Rev. Lett. **111**, 112003 (2013).

[35] A. Lenz, U. Nierste, J. Charles, S. Descotes-Genon, A. Jantsch, C. Kaufhold, H. Lacker, S. Monteil, V. Niess, and S. T'Jampens, Phys. Rev. D **83**, 036004 (2011).

[36] S. Jäger and J. M. Camalich, J. High Energy Phys. 05 (2013) 043.

[37] A. Khodjamirian, T. Mannel, A. A. Pivovarov, and Y. M. Wang, J. High Energy Phys. 09 (2010) 089.

[38] D. Van Dyk, "Constraints on $|\Delta B| = |\Delta S| = 1$ Wilson Coefficients," EPS-HEP Conference, Stockholm, 2013 (unpublished).

[39] T. Huber, E. Lunghi, M. Misiak, and D. Wyler, Nucl. Phys. **B740**, 105 (2006).

[40] P. Langacker, Rev. Mod. Phys. **81**, 1199 (2009).

[41] A. J. Buras, F. De Fazio, and J. Girrbach, J. High Energy Phys. 02 (2013) 116.

[42] D. Becirevic, V. Lubicz, and F. Mescia, Nucl. Phys. **B769**, 31 (2007).

[43] Z. Liu, S. Meinel, A. Hart, R. R. Horgan, E. H. Muller, and M. Wingate, arXiv:1101.2726.

[44] A. Bharucha, T. Feldmann, and M. Wick, J. High Energy Phys. 09 (2010) 090.

[45] LHCb Collaboration, J. High Energy Phys. 02 (2013) 105.

[46] D. Becirevic, N. Kosnik, F. Mescia, and E. Schneider, Phys. Rev. D **86**, 034034 (2012).

[47] R. Zhou *et al.* (Fermilab Lattice and MILC Collaborations), Proc. Sci., LATTICE2011 (2011) 298 [arXiv:1111.0981].

[48] LHCb Collaboration, J. High Energy Phys. 07 (2013) 084.

[49] CDF Collaboration, Phys. Rev. Lett. **107**, 201802 (2011).

[50] LHCb Collaboration, arXiv:1306.2577 [Physics Letters B (to be published)].

[51] T. Gutsche, M. A. Ivanov, J. G. Korner, V. E. Lyubovitskij, and P. Santorelli, Phys. Rev. D **87**, 074031 (2013).

[52] T. Feldmann and M. W. Y. Yip, Phys. Rev. D **85**, 014035 (2012); **86**, 079901(E) (2012).

[53] Y.-M. Wang, Y. Li, and C.-D. Lu, Eur. Phys. J. C **59**, 861 (2009).

[54] Y.-M. Wang, Y.-L. Shen, and C.-D. Lu, Phys. Rev. D **80**, 074012 (2009).

[55] W. Detmold, C.-J. D. Lin, S. Meinel, and M. Wingate, Phys. Rev. D **87**, 074502 (2013).

[56] C.-D. Lu and W. Wang, Phys. Rev. D **85**, 034014 (2012).

[57] D. Becirevic and A. Tayduganov, Nucl. Phys. **B868**, 368 (2013).

[58] J. Matias, Phys. Rev. D **86**, 094024 (2012).

[59] M. Doring, U.-G. Meißner, and W. Wang, arXiv:1307.0947.

[60] T. Blake, U. Egede, and A. Shires, J. High Energy Phys. 03 (2013) 027.