PHYSICAL REVIEW D **93,** 054008 (2016)

Generalized helicity formalism, higher moments, and the $B \to K_{J_K}(\to K\pi)\overline{\ell}_1\ell_2$ angular distributions

James Gratrex,[1,*] Markus Hopfer,[1,2,†] and Roman Zwicky[1,‡]

[1]*Higgs Centre for Theoretical Physics, School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3JZ, Scotland*

[2]*Institut für Physik, Karl-Franzens-Universität Graz, Universitätsplatz 5, 8010 Graz, Austria*

(Received 15 October 2015; published 8 March 2016)

We generalize the Jacob-Wick helicity formalism, which applies to sequential decays, to effective field theories of rare decays of the type $B \to K_{J_K}(\to K\pi)\overline{\ell}_1\ell_2$. This is achieved by reinterpreting local interaction vertices $\overline{b}\Gamma'_{\mu_1\ldots\mu_n}s\overline{\ell}\Gamma^{\mu_1\ldots\mu_n}\ell$ as a coherent sum of $1 \to 2$ processes mediated by particles whose spin ranges between zero and n. We illustrate the framework by deriving the full angular distributions for $\overline{B} \to \overline{K}\ell_1\overline{\ell}_2$ and $\overline{B} \to \overline{K}^*(\to \overline{K}\pi)\ell_1\overline{\ell}_2$ for the complete dimension-six effective Hamiltonian for nonequal lepton masses. Amplitudes and decay rates are expressed in terms of Wigner rotation matrices, leading naturally to the method of moments in various forms. We discuss how higher-spin operators and QED corrections alter the standard angular distribution used throughout the literature, potentially leading to differences between the method of moments and the likelihood fits. We propose to diagnose these effects by assessing higher angular moments. These could be relevant in investigating the nature of the current LHCb anomalies in $R_K = \mathcal{B}(B \to K\mu^+\mu^-)/\mathcal{B}(B \to Ke^+e^-)$ as well as angular observables in $B \to K^*\mu^+\mu^-$.

DOI: 10.1103/PhysRevD.93.054008

I. INTRODUCTION

Helicity amplitudes (HAs), as defined by Jacob and Wick [1], describe $A \to BC$ $(1 \to 2)$ transitions and have definite transformation properties under rotation. The key idea is that the angular and helicity information are equivalent to each other. Angular decay distributions follow (e.g. [2–4]) from evaluating the HAs with B and C in the forward direction, with the angular information encoded in Wigner D matrix functions, reminiscent of the Wigner-Eckart theorem.

The intent of this paper is to generalize this method to decays of the type $A \to (B_1B_2)C$ which are schematically described by local interactions of the form

$$H^{\text{eff}} \sim (AC)_{\mu_1\ldots\mu_n}(B_1B_2)^{\mu_1\ldots\mu_n}. \tag{1}$$

We do so by rewriting the $1 \to 3$ decay as a sequence of $1 \to 2$ processes, by inserting multiple complete sets of polarization states between the Lorentz contractions of AC and B_1B_2 above. This leads to a reinterpretation of the decay in terms of a sum over intermediate particles of spin J, where J can range from 0 up to n depending on the specific structure of the operators. Symbolically we may write

$$\mathcal{A}(A \to (B_1B_2)C) = \sum_{J\geq 0}^{n} \mathcal{A}(A \to B_J(\to B_1B_2)C), \tag{2}$$

with $\mathcal{A}$ denoting the amplitude. We refer to this case as the B-particle factorization approximation. At the formal level, the main work is the decomposition of the Lorentz tensors into irreducible objects under the spatial rotation group (reminiscent of the $3+1$ decomposition of cosmological perturbation theory for example).

Important examples of such decays are given by the rare radiative decays $B \to K\ell^+\ell^-$ and $B \to K^*(\to K\pi)\ell^+\ell^-$. Besides evaluating nonperturbative matrix elements to these decays (e.g. [5–17]), it has become clear that it is beneficial to consider general properties of the amplitudes entering the angular distributions (e.g. [18–21]). Our work can be seen to be part of the latter category.

We evaluate the $\overline{B} \to \overline{K}^{(*)}\ell^+\ell^-$ angular decay distributions within the generalized helicity framework developed in this paper, providing an alternative method to traditional techniques using Dirac trace technology [22,23]. An important consequence of the manner in which we derive the distribution is that it lends itself to the methods of moments (MoM), which use the decomposition of the distribution into orthogonal functions to obtain observables independently of each other. This is a complementary method to the likelihood fit to extract the dynamical information from the decay, and was recently studied from an experimental viewpoint in [24]. We discuss the impact of including higher partial waves in both the $(K\pi)$- and especially the dilepton-system. The latter give rise to corrections, in the form of higher moments, to the standard form of the angular distribution used in the literature. The sources of higher dilepton partial waves are higher spin operators and electroweak corrections, both of which we discuss qualitatively. The two sources can be distinguished

*j.gratrex@ed.ac.uk
†markus.hopfer@emtensor.com,
‡roman.zwicky@ed.ac.uk.

by their different behavior in higher partial moments. We encourage experimental investigation of higher moments from various viewpoints. In particular, we discuss how higher moments can be used to diagnose the size of QED effects in $B \to K\ell^+\ell^-$ (with $\ell = e, \mu$) and test leakage of J/Ψ-contributions into the lower dilepton-spectrum. Both are of importance in view of R_K as well as the angular anomalies in the low dilepton-spectrum, which have recently been reported by the LHCb collaboration in [25] and [26,27] respectively.

The paper is organized as follows. In Sec. II the methodology is introduced ending with a formal expression for the fourfold decay distribution in terms of rotation matrices and HAs. Specific angular distributions for $\overline{B} \to \overline{K}^{(*)}\ell_1\overline{\ell}_2$,[1] with detailed results in appendices C (and a *Mathematica* notebook in the arXiv version [29]) and D, are given in Sec. III. The method of total and partial moments is presented in Sec. IV. Section V contains the discussion of including higher partial waves: a qualitative assessment of higher spin operators and QED corrections is presented in subsections V B and V C respectively. The relevance of testing for higher moments is emphasised in subsection V D. The paper ends with conclusions in Sec. VI. Additional material, such as the leptonic HAs and a few brief remarks on $\Lambda_b \to \Lambda(\to (p,n)\pi)\ell_1\overline{\ell}_2$, is presented in Appendices A 3 and E respectively. In appendix B we provide the kinematic conventions for computation of the angular distribution by the sole use of Dirac trace technology.

II. GENERALIZED HELICITY FORMALISM FOR EFFECTIVE THEORIES

We first review the standard helicity formalism in Sec. II A, and qualitatively apply it to sequential $1 \to 2$ decays in Sec. II A 1, specializing to the spin configuration relevant for our decays at the end. In Sec. II B the formalism is extended to include decays like $B \to K_{J_K}(\to K\pi)\overline{\ell}_1\ell_2$ described by effective field theories for $b \to s\overline{\ell}_1\ell_2$ transitions. The framework can be straightforwardly applied to the entire zoo of semileptonic and rare flavor decays such as $B_s \to K^*\ell\nu$, $B \to D^{(*)}\ell\nu$, $D \to (\pi,\rho)\mu\mu$, $D \to (\pi,\rho)\mu\nu$, $K \to \pi\mu\mu$ etc., and can also be extended to include initial particles with nonzero spin.

A. The basic idea of the helicity formalism and its extension

The discussion in this section is standard and we refer the reader to [2–4] for more extensive reviews, as well as the pioneering paper of Jacob and Wick [1]. In a $1 \to 2$ (say $A \to B_1B_2$) decay a particle of spin J_A and helicity M_A decays into two particles of momentum $\vec{p}_1$ and $\vec{p}_2$ with helicities λ_1 and λ_2 respectively. In the center-of-mass frame ($\vec{p}_1 = -\vec{p}_2$) the system can be characterized by the two helicities and the direction (i.e. the solid angles θ and ϕ). By inserting a complete set of two-particle angular momentum states the corresponding matrix element can be written

$$\begin{aligned}\mathcal{A}(A \to B_1B_2) &= \langle\theta,\phi,\lambda_1,\lambda_2|J_A,M_A\rangle \\ &= \sum_{j,m}\langle\theta,\phi,\lambda_1,\lambda_2|j,m,\lambda_1,\lambda_2\rangle\langle j,m,\lambda_1,\lambda_2|J_A,M_A\rangle \\ &= \underbrace{\langle\theta,\phi,\lambda_1,\lambda_2|J_A,M_A,\lambda_1,\lambda_2\rangle}_{\sqrt{\frac{2J_A+1}{4\pi}}D^{J_A}_{M_A,\lambda_1-\lambda_2}(\phi,\theta,-\phi)}\underbrace{\langle J_A,M_A,\lambda_1,\lambda_2|J_A,M_A\rangle}_{\mathcal{A}^{J_A}_{M_A,\lambda_1,\lambda_2}}\end{aligned} \tag{3}$$

as a product of Wigner D-functions and a HA $\mathcal{A}^{J_A}_{M_A,\lambda_1,\lambda_2}$. The Wigner matrix is a $(2J_A+1)$-dimensional $SO(3)$ representation in the helicity basis. The essence is that the distribution of the amplitude over the angles is then governed by the rotation matrix as a function of the helicities. In practice one only needs to compute the HA.

The process $B \to J/\Psi(\to \ell^+\ell^-)K^*(\to K\pi)$ constitutes a well-known example of a sequential $1 \to 2$ decay where the formalism can be applied [30]. The idea of this paper is to extend this formalism to the case where the $\overline{\ell}_1\ell_2$-pair emerges from a local interaction vertex $O_{ij} \sim \bar{s}\Gamma_i b\overline{\ell}\Gamma_j\ell$ with effective Hamiltonian $H^{\text{eff}} \sim \sum_{ij}C_{ij}O_{ij}$. This is achieved by reinterpreting the local interaction vertex as originating from a sum of particles whose spin depends on the number of Lorentz contractions between the $\Gamma_{i,j}$ structures. Elements of this program have appeared in the literature, e.g. [31] for $B \to K_J\ell^+\ell^-$, but we are unaware of a systematic presentation that allows the incorporation of a generic effective Hamiltonian as well as other decay types.

1. Helicity formalism for $B_{J_B} \to K_{J_K}(\to K_1K_2)\gamma_{J_\gamma}(\to \overline{\ell}_1\ell_2)$

Let us consider the following sequential decay

$$B_{J_B} \to K_{J_K}(\to K_1K_2)\gamma_{J_\gamma}(\to \overline{\ell}_1\ell_2) \tag{4}$$

where J_B, J_γ and J_K denote the spin of the particles B, γ_J and K_J. The notation is close to the main application of this paper but we emphasise that at this point the methodology is completely general. Assuming the decay to be a series of sequential $1 \to 2$ decays the amplitude can be written in terms of a product of $1 \to 2$ HAs times the corresponding Wigner functions

[1]Throughout this work, we use nonequal leptons $\ell_1 \neq \ell_2$, in order to accommodate semileptonic decays of the type $B \to \rho\ell\nu$ as well as potential lepton flavor violation [28], motivated by the R_K measurement.

$$\mathcal{A}(\Omega_B,\Omega_\ell,\Omega_K|\lambda_B,\lambda_{K_1},\lambda_{K_2},\lambda_1,\lambda_2) \sim \sum_{\lambda_\gamma\lambda_K} \overline{D}^{J_B}_{\lambda_B,\lambda_\gamma-\lambda_K}(\Omega_B)\mathcal{H}_{\lambda_\gamma\lambda_K}\overline{D}^{J_K}_{\lambda_K,\lambda_{K_2}-\lambda_{K_1}}(\Omega_K)\mathcal{K}_{\lambda_{K_1},\lambda_{K_2}} \times \overline{D}^{J_\gamma}_{\lambda_\gamma,\lambda_\ell}(\Omega_\ell)\mathcal{L}_{\lambda_{\ell_1}\lambda_{\ell_2}}, \tag{5}$$

where the λ_i are helicity indices, and

$$\lambda_\ell \equiv \lambda_1 - \lambda_2, \tag{6}$$

is a shorthand that we use frequently throughout the paper. The HAs $\mathcal{H}$, $\mathcal{K}$ and $\mathcal{L}$ correspond to the transitions $B_{J_B} \to K_{J_K}\gamma_{J_\gamma}$, $K_{J_K} \to K_1K_2$ and $\gamma_{J_\gamma} \to \overline{\ell}_1\ell_2$ respectively. The helicities of the internal particles γ_J and K_J have to be coherently summed over. The Wigner D-functions

$$D^j_{m',m}(\Omega=(\alpha,\beta,\gamma)) = \langle jm'|e^{-i\alpha J_z}e^{-i\beta J_y}e^{-i\gamma J_z}|jm\rangle \tag{7}$$

are irreducible $SO(3)$-representations of dimension $2j+1$. The J_i are the generators of angular momentum, and the states $|j,m\rangle$ carry angular momentum j and helicity m and are orthonormalized $\langle j,m|j',m'\rangle = \delta_{jj'}\delta_{mm'}$. To avoid proliferation of indices we denote complex conjugation by a bar instead of the more standard asterisk.

Adaptation to $J_B = 0$ and $K_1 = K$ and $K_2 = \pi$ In order to ease the notation slightly we move straight to the case $\overline{B} \to \overline{K}_J(\to K\pi)\gamma_{J_\gamma}(\to \ell_1\overline{\ell}_2)$.[2] The relation $D^{J_B=0}_{\lambda_B=0,\lambda_\gamma-\lambda_K}(\Omega) = \delta_{0,\lambda_\gamma-\lambda_K}$ implies equality of helicities

$$\lambda \equiv \lambda_\gamma = \lambda_K. \tag{8}$$

One may therefore reduce $\mathcal{H}_{\lambda_\gamma\lambda_K} \to H_\lambda$, which is the quantity known as the HA in the $\overline{B} \to \overline{K}^*\ell^+\ell^-$-literature and carries the nontrivial dynamic information. The HA $\mathcal{K}_{\lambda_{K_1},\lambda_{K_2}}$ reduces to a scalar constant (denoted by $g_{K_JK\pi}$) since $K_1 \to K$, $K_2 \to \pi$ are both scalar particles. The third HA $\mathcal{L}_{\lambda_1\lambda_2}$ depends on the interaction vertex of the leptons, but is trivial to calculate once the interaction is known. We may rewrite the amplitude (5) as

$$\mathcal{A}(\overline{B} \to \overline{K}_{J_K}(\to K\pi)\gamma_{J_\gamma}(\to \ell_1\overline{\ell}_2)) \sim \sum_{\lambda=-J_K}^{J_K} D^{J_K}_{\lambda,0}(\Omega_K)D^{J_\gamma}_{\lambda,\lambda_\ell}(\Omega_\ell)\mathcal{A}^{J_\gamma}_{\lambda,\lambda_1,\lambda_2}, \tag{9}$$

where the angles, depicted in Fig. 1, are $\Omega_K = (0,\theta_K,0)$ and $\Omega_\ell = (\phi_\ell,\theta_\ell,-\phi_\ell)$. Note, the passage from $\overline{D}$ to D-functions from (5) to (9) is related to passing from B to $\overline{B}$.

[2]We choose $\bar{B} \to \bar{K}_J\ell_1\bar{\ell}_2$ transitions as the main template for the results in this paper. Such transitions are governed by the $b \to s$ Hamiltonian, which is the standard in the theory literature and is used to define the Wilson coefficients. In the more conceptual sections we shall refer to $B \to K\bar{\ell}_1\ell_2$-transitions.

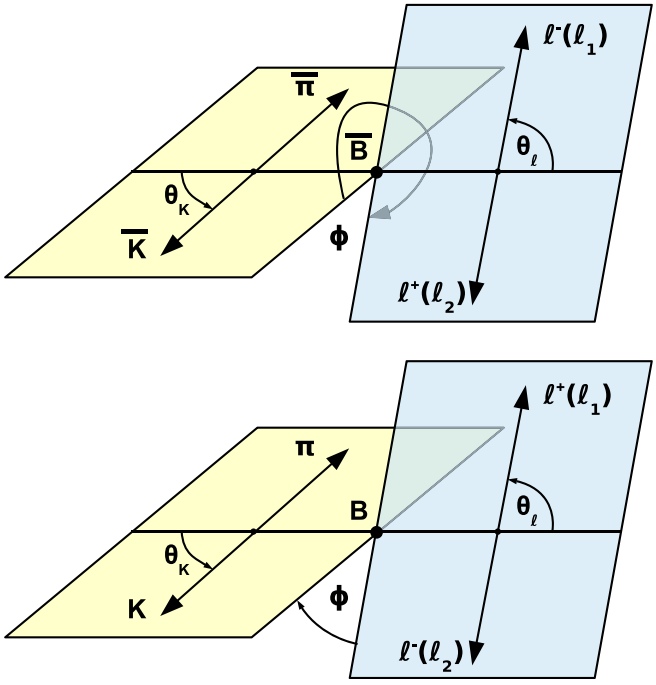

FIG. 1. Decay geometries for $\bar{B} \to \bar{K}^*\ell_1\bar{\ell}_2$ (above) and $B \to K^*\bar{\ell}_1\ell_2$ (below). In both cases $\ell_1 = \ell^-$, $\ell_2 = \ell^-$ denote the negatively charged lepton. The conventions are the same as used by the LHCb collaboration in [32] (cf. Appendix A therein). Comparison to the convention used by the theory community can be found in Appendix C 2. The pictures are slightly misleading in that the angles $\theta_{\ell,K}$ are drawn in the rest frame of the lepton-pair and the K^*-meson. For decays which are not self-tagging, such as $B_s, \bar{B}_s \to \phi(\to K^+K^-)\mu^+\mu^-$ at the LHCb, the angles $(\theta_\ell,\theta_K,\phi) \to (\pi-\theta_\ell,\pi-\theta_K,2\pi-\phi)$, and one can only measure the sum of both decay rates [second equation (C6)].

In the lepton-pair factorization approximation, defined more explicitly in the following section, the amplitude $\mathcal{A}^{J_\gamma}_{\lambda,\lambda_1,\lambda_2} \sim H_\lambda\mathcal{L}_{\lambda_1\lambda_2}|_{J_\gamma}$ is the product of the hadronic and leptonic matrix elements. The angle ϕ_ℓ is the helicity angle and is usually called simply ϕ. Before commenting on different conventions of the angles we quote the fourfold differential decay

$$\frac{d^4\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi} \sim \sum_{\lambda_1\lambda_2}|\mathcal{A}|^2 \sim \tag{10}$$

$$\sum_{\lambda_1,\lambda_2=-1/2}^{1/2}\sum_{\lambda,\lambda'=-J_\gamma}^{J_\gamma}\mathcal{A}^{J_\gamma}_{\lambda,\lambda_1,\lambda_2}\overline{\mathcal{A}}^{J_\gamma}_{\lambda',\lambda_1,\lambda_2}D^{J_{K_J}}_{\lambda,0}(\Omega_K)\overline{D}^{J_{K_J}}_{\lambda',0}(\Omega_K) \times D^{J_\gamma}_{\lambda,\lambda_\ell}(\Omega_\ell)\overline{D}^{J_\gamma}_{\lambda',\lambda_\ell}(\Omega_\ell), \tag{11}$$

in terms of amplitudes and Wigner D-functions. For the angles we use the $\overline{B} \to \overline{K}^*\ell\ell$ decay as a reference and use the same conventions as the LHCb collaboration [32] (Appendix A), which differ from those used by the theory

community. More precise statements, including a conversion diagram, can be found in Appendix C 2.

B. Effective theories rewritten as a coherent sum of sequential decays

In this section we give the formal steps to derive the expression of the angular distributions. The reader interested in the final result can directly proceed to Sec. III.

The amplitude (9) is of a completely general form for the decay where γ_{J_γ} is an actual particle of spin J_γ. In $\overline{B} \to \overline{K}^*(\to \overline{K}\pi)\ell_1\overline{\ell}_2$ a part of the amplitude is in this form where the photon corresponds to the intermediate state ($\gamma_1 = \gamma$). In general there are effective vertices, so-called contact terms, where the intermediate particles are not present. In the interest of clarity we quote the effective Hamiltonian for $b \to s\ell\overline{\ell}$[3]:

$$\begin{aligned} H^{\text{eff}} &= c_H \hat{H}^{\text{eff}}, \\ c_H &\equiv -\frac{4G_F}{\sqrt{2}}\frac{\alpha}{4\pi} V_{\text{ts}}^* V_{\text{tb}}, \\ \hat{H}^{\text{eff}} &= \sum_{i=V,A,S,P,\mathcal{T}} (C_i O_i + C_i' O_i'). \end{aligned} \tag{12}$$

Above G_F is Fermi's constant, α the fine structure constant, V_{tD} are Cabibbo-Kobayashi-Maskawa (CKM) elements and the operators are

$$\begin{aligned} O_{S(P)} &= \overline{s}_L b\overline{\ell}(\gamma_5)\ell, \\ O_{V(A)} &= \overline{s}_L\gamma^\mu b\overline{\ell}\gamma_\mu(\gamma_5)\ell, \\ O_{\mathcal{T}} &= \overline{s}_L\sigma^{\mu\nu} b\overline{\ell}\sigma_{\mu\nu}\ell, \end{aligned} \tag{13}$$

where $O' = O|_{s_L \to s_R}$, the labels refer to the lepton interaction vertex, $q_{L,R} \equiv 1/2(1\mp\gamma_5)q$, $\overline{\ell}$, $\ell \to \overline{\ell}_1$, ℓ_2 for different lepton flavors and a few additional relevant remarks deferred to Appendix A 2. In passing we add that the notation $O_{9(10)} = O_{V(A)}$ is more common throughout the literature. In the case where electroweak corrections are neglected at the matrix element level one may factorize the hadronic from the leptonic part. We refer to this as the lepton-pair factorization approximation (LFA) (B-particle factorization approximation in the introduction). Schematically (12) is written as a product of a hadronic part H and a leptonic part $\mathcal{L}$ with a certain number of Lorentz contractions between them:

$$H^{\text{eff}} \sim \sum_{a=1}^{N_0} H^a \mathcal{L}_a + \sum_{b=1}^{N_1} H_\mu^b \mathcal{L}_b^\mu + \sum_{c=1}^{N_2} H_{\mu_1\mu_2}^c \mathcal{L}_c^{\mu_1\mu_2}. \tag{14}$$

[3]The adaptation from $\ell\overline{\ell} \to \ell_1\overline{\ell}_2$ is trivial and will not be spelled out explicitly.

The sum over a, b and c extends over operators with 0, 1 and 2 Lorentz contractions between quark and lepton operators. In the example of $C_V O_V = H_\mu \mathcal{L}^\mu$ we would have $H_\mu = C_V \overline{s}_L\gamma_\mu b$ and $\mathcal{L}^\mu = \overline{\ell}\gamma^\mu\ell$. On a formal level we might think of $O_V(O_9)$ as originating from integrating out a vector and a scalar particle, in the sense that the Lorentz contraction over index μ can be written as the sum of products of a spin-one and a timelike spin-0 polarization vector. This is expressed by the well-known completeness relation (e.g. [18,23,31])

$$\begin{aligned} g^{\mu\nu} &= \sum_{\lambda,\lambda' \in \{t,\pm,0\}} \omega^\mu(\lambda)\overline{\omega}^\nu(\lambda') G_{\lambda\lambda'}, \\ G_{\lambda\lambda'} &= \text{diag}(1,-1,-1,-1), \end{aligned} \tag{15}$$

where the first entry in $G_{\lambda\lambda'}$ refers to $\lambda = \lambda' = t$ and an explicit parametrization is given by

$$\begin{aligned} \omega^\mu(\pm) &= (0,\pm 1, i, 0)/\sqrt{2}, \\ \omega^\mu(0) &= (q_z, 0, 0, q_0)/\sqrt{q^2}, \\ \omega^\mu(t) &= (q_0, 0, 0, q_z)/\sqrt{q^2}, \end{aligned} \tag{16}$$

which is consistent with the parametrization $q^\mu = (q_0, 0, 0, q_z)$. The polarization vectors $\omega(\pm, 0)$ are compatible with the Jacob-Wick phase convention [1] (cf. Appendix B and the corresponding footnote for further remarks). Let us pause a moment and emphasise that intermediate results do depend on the convention, which enters the definition of the HAs, and this dependence has to be taken into account when comparing to HAs appearing in the literature. We choose the convention in [18], since it is compatible with the Condon-Shortly convention that is standard for Clebsch-Gordon coefficients and Wigner matrices (e.g. [33]).

We may think of ω as being associated with the Lorentz group $SO(3,1)$. In the rest frame $q_z = 0$ the timelike polarization tensor transforms as a scalar under the restriction of $SO(3,1)$ to spatial rotations $SO(3)$.[4] For an effective operator with n Lorentz indices the relation (15) can be inserted n times to obtain a HA with n helicity indices. More precisely, the direct product of $SO(3,1)$ polarization tensors decomposes into irreducible representations of $SO(3)$ polarization tensors $\epsilon^{j,\lambda}_{\mu_1\ldots\mu_n}$ of spin $j = 0,\ldots,n$ and helicities $\lambda = -j,\ldots,j$. Using the expressions in Eqs. (A8) and (A9) the analogue of $\mathcal{A}^{J_\gamma}_{\lambda,\lambda_1,\lambda_2}$ in (9) on each spin component can be written as[5]

[4]Formally the branching rule for the Lorentz four vector $(1/2,1/2)$ is $(1/2,1/2)_{\text{SO}(3,1)}|_{\text{SO}(3)} \to (\mathbf{1}+\mathbf{3})_{\text{SO}(3)}$.

[5]In the notation used throughout the literature $H^t = \langle H_b^\mu\rangle\epsilon_\mu^{0,0} = \langle H_b^\mu\rangle\omega_\mu(t)$ is known as the timelike HA [23,31]. By virtue of the equation of motion the timelike HAs can be absorbed into the scalar and pseudoscalar HAs, cf. Appendix C 5.

$$\mathcal{A}^{J_\gamma}_{\lambda,\lambda_1,\lambda_2} = \begin{cases} \langle H_a\rangle\langle \mathcal{L}_a\rangle + \langle H_b^\mu\rangle\langle \mathcal{L}_b^\alpha\rangle \epsilon_\mu^{0,0}\overline{\epsilon}_\alpha^{0,0} + \langle H_c^{\mu\nu}\rangle\langle \mathcal{L}_c^{\alpha\beta}\rangle \epsilon_{\mu\nu}^{0,0}\cdot\overline{\epsilon}_{\alpha\beta}^{0,0} & J_\gamma = 0 \\ \langle H_b^\mu\rangle\langle \mathcal{L}_b^\alpha\rangle \epsilon_\mu^{1,\lambda}\overline{\epsilon}_\alpha^{1,\lambda_\ell} + \langle H_c^{\mu\nu}\rangle\langle \mathcal{L}_c^{\alpha\beta}\rangle \epsilon_{\mu\nu}^{1,\lambda}\cdot\overline{\epsilon}_{\alpha\beta}^{1,\lambda_\ell} & J_\gamma = 1 \\ \langle H_c^{\mu\nu}\rangle\langle \mathcal{L}_c^{\alpha\beta}\rangle \epsilon_{\mu\nu}^{2,\lambda}\cdot\overline{\epsilon}_{\alpha\beta}^{2,\lambda_\ell} & J_\gamma = 2 \end{cases} \tag{17}$$

where summation over Lorentz indices and the number of operators in (14) are both implied, the scalar product "·" is detailed in (A8) and

$$\begin{aligned} \langle H_a^{\mu_1\cdots\mu_n}\rangle &\equiv \langle \overline{K}_J(\lambda)|H_a^{\mu_1\cdots\mu_n}|\overline{B}\rangle, \\ \langle \mathcal{L}_a^{\mu_1\cdots\mu_n}\rangle &\equiv \langle \ell_1(\lambda_1)\overline{\ell}_2(\lambda_2)|\mathcal{L}_a^{\mu_1\cdots\mu_n}|0\rangle, \end{aligned} \tag{18}$$

are the leptonic and hadronic matrix elements. The helicities in (17) are the helicities of the outgoing particles of the HAs, with λ for $K_J(\lambda)$ in $H^{B\to K_{J_K}}$ and $\lambda_\ell = \lambda_1 - \lambda_2$ for $\ell_1(\lambda_1)\overline{\ell}_2(\lambda_2)$ in $\mathcal{L}^{\gamma_{J_\gamma}\to\ell_1\overline{\ell}_2}$. This is the main idea of the formalism: the angular dependence from the ingoing to outgoing particle is governed by the Wigner D-function, e.g. $\overline{\epsilon}^{J_{\ell,\lambda}} = \overline{D}^{J_\gamma}_{\lambda,\lambda_\ell}(\Omega_\ell)\overline{\epsilon}^{J_{\ell,\lambda_\ell}}$ for $\mathcal{L}^{\gamma_{J_\gamma}(\lambda)\to\ell_1(\lambda_1)\overline{\ell}_2(\lambda_2)}$, which is inherent in (3). The generalized HA then becomes essentially a sum over all spin components J_γ necessary to saturate the Lorentz indices in the effective Hamiltonian,

$$\begin{aligned} &\mathcal{A}(\overline{B}\to\overline{K}_{J_K}(\to K\pi)\ell_1\overline{\ell}_2) \\ &= \frac{\sqrt{2J_K+1}}{4\pi}\sum_{J_\gamma=0}^{n}\sum_{\lambda=-\min(J_\gamma,J_K)}^{\min(J_\gamma,J_K)} D^{J_K}_{\lambda,0}(\Omega_K)D^{J_\gamma}_{\lambda,\lambda_\ell}(\Omega_\ell)\mathcal{A}^{J_\gamma}_{\lambda,\lambda_1,\lambda_2}, \end{aligned} \tag{19}$$

where the overall factor follows from (3). A schematic representation of Eq. (19) is given in Fig. 2. The differential decay distribution (10) is replaced by a similar expression

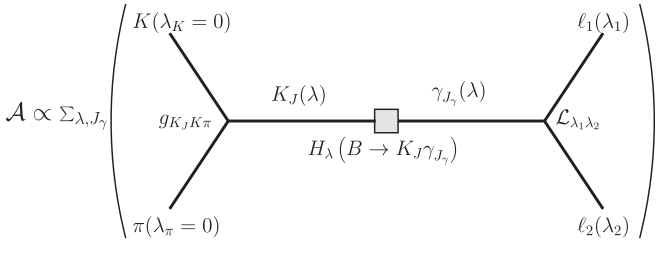

FIG. 2. A diagrammatic interpretation of the process, Eq. (19), used to set up the formalism. The decay to two leptons is treated as being mediated by an effective particle γ_{J_γ} of spin J_γ. The factor $g_{K_JK\pi}$ has no dependence on helicities and depends only on the dynamics of the K^* decay.

$$\begin{aligned} \frac{d^4\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi} &\sim \sum_{\lambda_1\lambda_2}|\mathcal{A}|^2 = \frac{2J_K+1}{4\pi}\sum_{\lambda_1\lambda_2}\sum_{J_\gamma\lambda}\sum_{J'_\gamma\lambda'} \\ &\times \mathcal{A}^{J_\gamma}_{\lambda,\lambda_1,\lambda_2}\overline{\mathcal{A}}^{J'_\gamma}_{\lambda',\lambda_1,\lambda_2}D^{J_K}_{\lambda,0}(\Omega_K)\overline{D}^{J_K}_{\lambda',0}(\Omega_K)D^{J_\gamma}_{\lambda,\lambda_\ell}(\Omega_\ell)\overline{D}^{J_\gamma}_{\lambda',\lambda_\ell}(\Omega_\ell), \end{aligned} \tag{20}$$

with additional coherent sums over the spins J_γ

$$\sum_{\lambda_1\lambda_2} = \sum_{\lambda_1,\lambda_2=-1/2}^{1/2}, \qquad \sum_{J_\gamma\lambda} = \sum_{J_\gamma=0}^{2}\sum_{\lambda=-\min(J_\gamma,J_{K_J})}^{\min(J_\gamma,J_{K_J})} \tag{21}$$

and likewise for the sum over J'_γ, λ'.

III. ANGULAR DISTRIBUTION AND WIGNER D-FUNCTIONS

We now apply the method introduced in the previous section to decays governed by the $b\to s\ell_1\overline{\ell}_2$ effective Hamiltonian (12). First we consider the decay $\overline{B}\to\overline{K}^*(\to\overline{K}\pi)\ell_1\overline{\ell}_2$, and then in Sec. III C we present similar results for $\overline{B}\to\overline{K}\ell_1\overline{\ell}_2$. The related decay $\Lambda_b\to\Lambda(\to N\pi)\ell_1\overline{\ell}_2$, where $N = (p,n)$, can also be treated within this formalism, and will be briefly considered in Appendix E.

TABLE I. The definitions of the Γ^X and their associated spin $J_\gamma(X)$. The contributions $J_\gamma(X) = 0, 1$ give rise to the S_ℓ- and P_ℓ-wave amplitudes respectively. The basic polarisation vector ω_μ is given in (16) and the composed ones can be found in Eq. (A5). The precise value of the helicity index λ_X is specified when the leptonic and hadronic HAs are defined in Eqs. (A13), (C14), (D4). Note that the additional structure $\Gamma^{T_5} = \sigma_{\mu\nu}\gamma_5$ can be absorbed into the other tensor structures due to the identity $\sigma^{\alpha\beta}\gamma_5 = -\frac{i}{2}\epsilon^{\alpha\beta\mu\nu}\sigma_{\mu\nu}$ (with the $\epsilon_{0123} = +1$ convention for the Levi-Civita tensor). Timelike contributions $\gamma^\mu[\gamma_5]\omega_\mu(t)$ can be absorbed into $\Gamma^{S,P}$ respectively, as detailed in Appendix C 5. Above $\sigma_{\mu\nu} = i/2[\gamma_\mu,\gamma_\nu]$.

	$\Gamma^{S[P]}$	$\Gamma^{V[A]}$	$\Gamma^{T[T_t]}$
Γ^X	$\mathbf{1}_4[\gamma_5]$	$\gamma^\mu[\gamma_5]\omega_\mu(\lambda_X)$	$\sigma^{\mu\nu}\omega^{1,\lambda_X}_{\mu\nu}[\omega^{t,\lambda_X}_{\mu\nu}]$
$J_\gamma(X)$	0	1	1

A. $\overline{B} \to \overline{K}^*(\to \overline{K}\pi)\ell_1\overline{\ell}_2$

The use of the effective Hamiltonian (12) in the LFA restricts the partial waves to $J_\gamma = 0, 1$ terms in equation (17). The discussion of higher partial waves ($J_\gamma \geq 2$) is deferred to Sec. V. The matrix element for (12) is then given by the sum of an S_ℓ- and P_ℓ-wave amplitude (with the subscript ℓ referring to the partial wave in the angle θ_ℓ):

$$\begin{aligned}\hat{\mathcal{M}}_{\lambda_1,\lambda_2} &= \langle \overline{K}^*(\to \overline{K}\pi)\ell_1(\lambda_1)\overline{\ell}_2(\lambda_2)|\hat{H}^{\rm eff}|\overline{B}\rangle \\ &= \frac{\sqrt{3}}{4\pi}\Big[\mathcal{A}^0_{0,\lambda_1,\lambda_2} D^1_{0,0}(\Omega_K)\delta_{\lambda_1\lambda_2} \\ &\quad + \sum_{\lambda=\pm,0} \mathcal{A}^1_{\lambda,\lambda_1,\lambda_2} D^1_{\lambda,0}(\Omega_K) D^1_{\lambda,\lambda_\ell}(\Omega_\ell)\Big], \end{aligned} \tag{22}$$

where the hat denotes the effective Hamiltonian without the c_H prefactor (12). There is no D-wave since the two-indices in the tensor operator (12) are antisymmetric and therefore in a spin 1 representation (cf. discussion in Sec. V B on higher spin operators). The K^* has spin 1 and is therefore always in a P_K-wave in the θ_K-angle, with analogous meaning for the K subscript as before. Above we have used $D^0_{0,\lambda_\ell}(\Omega) = \delta_{0\lambda_\ell}$ to impose $\delta_{\lambda_1\lambda_2}$ on the scalar part of the matrix element. The principal objects to be calculated are the amplitudes $\mathcal{A}^{J_\gamma}_{\lambda,\lambda_1,\lambda_2}$. For $H^{\rm eff}$ (12) the S_ℓ- and P_ℓ-wave amplitudes (that is to say $\mathcal{A}^0$ and $\mathcal{A}^1$ respectively) are written as

$$\begin{aligned}\mathcal{A}^0_{0,\lambda_1,\lambda_2} &= H^S\mathcal{L}^S_{\lambda_1,\lambda_2} + H^P\mathcal{L}^P_{\lambda_1,\lambda_2}, \\ \mathcal{A}^1_{\lambda,\lambda_1,\lambda_2} &= -H^V_\lambda\mathcal{L}^V_{\lambda_1,\lambda_2} - H^A_\lambda\mathcal{L}^A_{\lambda_1,\lambda_2} + H^T_\lambda\mathcal{L}^T_{\lambda_1,\lambda_2} - 2H^{T_t}_\lambda\mathcal{L}^{T_t}_{\lambda_1,\lambda_2}, \end{aligned} \tag{23}$$

with the relative signs and factor of 2 emerging from the (double) completeness relation (A3), and the leptonic and the hadronic HAs are

$$\begin{aligned} H^X_\lambda &= \langle \overline{K}^*(\lambda)|\bar{s}\overline{\Gamma}^X b|\overline{B}\rangle, \\ \mathcal{L}^X_{\lambda_1\lambda_2} &\equiv \langle \ell_1(\lambda_1)\overline{\ell}_2(\lambda_2)|\overline{\ell}\Gamma^X\ell|0\rangle, \end{aligned} \tag{24}$$

the expressions in (18) contracted with the corresponding polarization vectors; explicit expressions for the Γ^X are given in Table 1. Explicit results, as well as a more precise prescription concerning Γ^X, are given in Appendices A 3 and C 5 in Eqs. (A13) and (C14), respectively. Squaring the matrix element in (22), summing over external helicities and averaging over final-state spins, one obtains an angular distribution

$$\begin{aligned} I_{K^*}(q^2,\Omega_K,\Omega_\ell) &\equiv \frac{32\pi}{3}\frac{d^4\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi} \\ &= \frac{32\pi}{3}\mathcal{N}\sum_{\lambda_1,\lambda_2}|\hat{\mathcal{M}}_{\lambda_1,\lambda_2}|^2, \end{aligned} \tag{25}$$

with I_{K^*} being a shorthand and $32\pi/3$ is a convenient normalization factor. The factor $\mathcal{N}$,

$$\mathcal{N} \equiv |c_H|^2\kappa_{\rm kin}, \qquad \kappa_{\rm kin} \equiv \frac{\sqrt{\lambda_B}\sqrt{\lambda_{\gamma^*}}}{2^6\pi^3 m_B^3 q^2}, \tag{26}$$

is the product of the prefactor resulting from the effective Hamiltonian c_H (12) and the kinematic phase space factor. The matrix element is defined in (22). Above $\lambda_B \equiv \lambda(m_B^2, m_{K^*}^2, q^2)$ and $\lambda_{\gamma^*} \equiv \lambda(q^2, m_{\ell_1}^2, m_{\ell_2}^2)$ where $\lambda(a,b,c)$ is the Källén-function defined in (B1) and related to the absolute value of the three-momentum of the K^* and the lepton pair by (B2).

1. Angular distribution

The squared matrix element initially contains a plethora of different products of four Wigner functions. However, these correspond to pairs of direct products that can be reduced to single Wigner functions by the Clebsch-Gordan series

$$D^j_{m,n}(\Omega)D^l_{p,q}(\Omega) = \sum_{J=|j-l|}^{j+l}\sum_{M=-J}^{J}\sum_{N=-J}^{J} C^{Jjl}_{Mmp}C^{Jjl}_{Nnq}D^J_{M,N}(\Omega). \tag{27}$$

Applied separately over the angles $\Omega_K = (0,\theta_K,0)$ and $\Omega_\ell = (\phi,\theta_\ell,-\phi)$, along with the identity $\overline{D}^l_{m,m'}(\Omega) = (-1)^{m'-m}D^l_{-m,-m'}(\Omega)$, this allows the angular distribution to be written in the compact form

$$\begin{aligned} &I^{(0)}_{K^*}(q^2,\Omega_K,\Omega_\ell) \\ &= {\rm Re}[G^{0,0}_0(q^2)\Omega^{0,0}_0 + G^{0,1}_0(q^2)\Omega^{0,1}_0 + G^{0,2}_0(q^2)\Omega^{0,2}_0 \\ &\quad + G^{2,0}_0(q^2)\Omega^{2,0}_0 + G^{2,1}_0(q^2)\Omega^{2,1}_0 + G^{2,1}_1(q^2)\Omega^{2,1}_1 \\ &\quad + G^{2,2}_0(q^2)\Omega^{2,2}_0 + G^{2,2}_1(q^2)\Omega^{2,2}_1 + G^{2,2}_2(q^2)\Omega^{2,2}_2], \end{aligned} \tag{28}$$

where the superscript (0) is a reminder that only S_ℓ- and P_ℓ-wave contributions were used to describe the amplitude (22). The angular functions Ω are given in terms of Wigner D functions

$$\begin{aligned} \Omega^{l_K,l_\ell}_m &\equiv \Omega^{l_K,l_\ell}_m(\Omega_K,\Omega_\ell) \\ &\equiv D^{l_K}_{m,0}(\Omega_K)D^{l_\ell}_{m,0}(\Omega_\ell) \\ &= D^{l_K}_{m,0}(\Omega'_K)D^{l_\ell}_{m,0}(\Omega'_\ell). \end{aligned} \tag{29}$$

The variables $\Omega'_K = (\phi,\theta_K,-\phi)$ and $\Omega'_\ell = (0,\theta_\ell,0)$ form an angular reparametrization that will prove convenient when we discuss partial moments. The label l_K corresponds to the

$(K\pi)$-system, l_ℓ to the dilepton system, and the common index m is the azimuthal component ϕ of either partial wave. The observables $G_m^{l_K,l_\ell}$ are functions of q^2 and the relation to the standard observables in the literature is given in Sec. III B. The explicit Wigner D-functions used above are given by

$$\begin{aligned}
D^0_{0,0}(\Omega) &= 1,\\
D^2_{0,0}(\Omega) &= \frac{1}{2}(3\cos^2\theta - 1),\\
D^2_{2,0}(\Omega) &= \sqrt{\frac{3}{8}}e^{-2i\phi}\sin^2\theta,\\
D^1_{0,0}(\Omega) &= \cos\theta,\\
D^1_{1,0}(\Omega) &= -\frac{1}{\sqrt{2}}e^{-i\phi}\sin\theta,\\
D^2_{1,0}(\Omega) &= -\sqrt{\frac{3}{8}}e^{-i\phi}\sin 2\theta,
\end{aligned} \tag{30}$$

and can be related to spherical harmonics $Y_{lm}(\theta,\phi)$ or associated Legendre polynomials $P_{lm}(x)$ as

$$\begin{aligned}
D^l_{m,0}(\phi,\theta,-\phi) &= \sqrt{\frac{4\pi}{2l+1}}\overline{Y}_{lm}(\theta,\phi)\\
&= \sqrt{\frac{(l-m)!}{(l+m)!}}P_{lm}(\cos\theta)e^{-im\phi}.
\end{aligned} \tag{31}$$

We comment briefly on four features of the angular distribution (28), all of which are encoded by the double Clebsch-Gordan series (27) but which can also be seen to emerge from the underlying physics:

(i) The second helicity index of all Wigner D-functions in the angular distribution is zero. The latter is the difference of the helicities of the final-state particles, which is zero since these helicities are summed incoherently, $(\lambda_1 - \lambda_2) - (\lambda_1 - \lambda_2) = 0$.

(ii) The first helicity index m is identical in all pairs of Wigner D-functions appearing in the angular distribution. This index contains the helicities of the internal particles, summed coherently. One can also see this as a property of the freedom of defining the reference plane for the angle ϕ.

(iii) The range of the indices l_K and l_ℓ is fixed between the range $0,\ldots,2\max[J_{K,\ell}]$. Including only $J_\gamma \le 1$ contributions emerging from the dimension-six effective Hamiltonian (12) hence imposes $0 \le l_\ell \le 2$, and likewise $J_K = 1$ imposing $0 \le l_K \le 2$.

(iv) The absence of angular structures with $l_K = 1$ is specific to this decay, due to the final state consisting of (pseudo)scalar mesons.

The first three features are universal to such decay chains and apply even if some of the particles involved are fermions, for example in the decay $\Lambda_b \to \Lambda(\to (p,n)\pi)\ell_1\overline{\ell}_2$, see Appendix E.

B. Relation of the $G_m^{l_K,l_\ell}$ to standard literature observables

The functions $G_m^{l_K,l_\ell}$, omitting the explicit q^2-dependence hereafter, are defined in terms of the standard basis of observables $g_i(q^2)$ parametrized in (C1) by

$$\begin{aligned}
G_0^{0,0} &= \frac{4}{9}(3(g_{1c}+2g_{1s}) - (g_{2c}+2g_{2s})),\\
G_0^{0,1} &= \frac{4}{3}(g_{6c}+2g_{6s}), \qquad G_0^{0,2} = \frac{16}{9}(g_{2c}+2g_{2s}),\\
G_0^{2,0} &= \frac{4}{9}(6(g_{1c}-g_{1s}) - 2(g_{2c}-g_{2s})),\\
G_0^{2,1} &= \frac{8}{3}(g_{6c}-g_{6s}), \qquad G_0^{2,2} = \frac{32}{9}(g_{2c}-g_{2s}),\\
G_1^{2,1} &= \frac{16}{\sqrt{3}}\mathcal{G}_5, \qquad G_1^{2,2} = \frac{32}{3}\mathcal{G}_4, \qquad G_2^{2,2} = \frac{32}{3}\mathcal{G}_3,
\end{aligned} \tag{32}$$

where we have defined $\mathcal{G}_{3,4,5} \equiv (g_{3,4,5} + ig_{9,8,7})$.

The twelve quantities (32), keeping in mind that the last three are complex, have been rewritten in several ways in the literature. A frequently used form is the set of observables given in [34], constructed to be insensitive to form factors. In the notation of LHCb [26], which includes their, and therefore our, angular conventions, the observables are given in terms of $G_m^{l_K,l_\ell}$ by[6]:

$$\begin{aligned}
\langle P_1\rangle_{\rm bin}|_{\rm LHCb} &= \frac{\langle {\rm Re}[G_2^{2,2}]\rangle_{\rm bin}}{\mathcal{N}_{\rm bin}},\\
\langle P_2\rangle_{\rm bin}|_{\rm LHCb} &= \frac{\langle 2G_0^{0,1} - G_0^{2,1}\rangle_{\rm bin}}{3\mathcal{N}_{\rm bin}},\\
\langle P_3\rangle_{\rm bin}|_{\rm LHCb} &= \frac{\langle {\rm Im}[G_2^{2,2}]\rangle_{\rm bin}}{2\mathcal{N}_{\rm bin}},\\
\langle P_4'\rangle_{\rm bin}|_{\rm LHCb} &= \frac{\langle {\rm Re}[G_1^{2,2}]\rangle_{\rm bin}}{4\mathcal{N}'_{\rm bin}},\\
\langle P_8'\rangle_{\rm bin}|_{\rm LHCb} &= \frac{\langle {\rm Im}[G_1^{2,2}]\rangle_{\rm bin}}{4\mathcal{N}'_{\rm bin}},\\
\langle P_5'\rangle_{\rm bin}|_{\rm LHCb} &= \frac{\langle {\rm Re}[G_1^{2,1}]\rangle_{\rm bin}}{2\sqrt{3}\mathcal{N}'_{\rm bin}},\\
\langle P_6'\rangle_{\rm bin}|_{\rm LHCb} &= \frac{\langle {\rm Im}[G_1^{2,1}]\rangle_{\rm bin}}{2\sqrt{3}\mathcal{N}'_{\rm bin}},
\end{aligned} \tag{33}$$

where we defined

[6]The extension of these relations to CP-odd and CP-even combinations, in the spirit of [35], is straightforward, see Sec. IV of [34].

$$\langle f(q^2)\rangle_{\rm bin} = \int_{\rm bin} dq^2 f(q^2),$$

as the integral over q^2 bins of the observable of interest, and[7]

$$\mathcal{N}_{\rm bin} \equiv 4\left\langle G_0^{0,2} - \frac{1}{2}G_0^{2,2}\right\rangle_{\rm bin},$$
$$\mathcal{N}'_{\rm bin} = \sqrt{-\left\langle G_0^{0,2} - \frac{1}{2}G_0^{2,2}\right\rangle_{\rm bin}\langle G_0^{0,2} + G_0^{2,2}\rangle_{\rm bin}}. \quad (34)$$

Three other combinations of the $G_m^{l_K,l_\ell}$ can be related to the branching fraction $\frac{d\Gamma}{dq^2}$, the forward-backward asymmetry $A_{\rm FB}$ and the longitudinal polarization fraction F_L [36]:

$$\left\langle\frac{d\Gamma}{dq^2}\right\rangle_{\rm bin} = \frac{3}{4}\langle G_0^{0,0}\rangle_{\rm bin},$$
$$\langle A_{\rm FB}\rangle_{\rm bin}|_{\rm LHCb} = \frac{1}{2}\frac{\langle G_0^{0,1}\rangle_{\rm bin}}{\langle G_0^{0,0}\rangle_{\rm bin}},$$
$$\langle F_L\rangle_{\rm bin} = \frac{\langle G_0^{0,0}\rangle_{\rm bin} + \langle G_0^{2,0}\rangle_{\rm bin}}{3\langle G_0^{0,0}\rangle_{\rm bin}}. \quad (35)$$

The observables in Eqs. (33), (34), (35) correspond to the twelve g_i. The definitions of the P_i' above correspond to those used by LHCb [26]; we give the correspondence to the observables defined in [34] in Appendix C 2.

C. $\overline{B} \to \overline{K}\ell_1\overline{\ell}_2$

Having shown the $\overline{B} \to \overline{K}^*\ell_1\overline{\ell}_2$ HA analysis in detail we are going to be rather brief on $\overline{B} \to \overline{K}\ell_1\overline{\ell}_2$. Skipping the step in (5) we directly write down the S_ℓ- and P_ℓ-wave amplitudes [analogue of Eq. (23)]:

$$\mathcal{A}^0_{0,\lambda_1,\lambda_2} = h^S\mathcal{L}^S_{\lambda_1,\lambda_2} + h^P\mathcal{L}^P_{\lambda_1,\lambda_2},$$
$$\mathcal{A}^1_{0,\lambda_1,\lambda_2} = -h^V\mathcal{L}^V_{\lambda_1,\lambda_2} - h^A\mathcal{L}^A_{\lambda_1,\lambda_2} + h^T\mathcal{L}^T_{\lambda_1,\lambda_2} - 2h^{T_t}\mathcal{L}^{T_t}_{\lambda_1,\lambda_2}, \quad (36)$$

where the $\mathcal{L}^X_{\lambda_1,\lambda_2}$ are the same as in the $\overline{B} \to \overline{K}^*\ell_1\overline{\ell}_2$ decay, and the hadronic HAs are taken over the same set of operators, but defined instead for $\overline{B} \to \overline{K}$ transitions. We again refer the reader to Appendix A 1 for a clarification of the signs and factor of 2 that emerge from the (double) completeness relation.

The reduced matrix element is then the sum of the S_ℓ- and P_ℓ-wave amplitude

$$\hat{\mathcal{M}}_{\lambda_1,\lambda_2} = \frac{1}{\sqrt{4\pi}}(\mathcal{A}^0_{0,\lambda_1,\lambda_2}\delta_{\lambda_1\lambda_2} + \mathcal{A}^1_{0,\lambda_1,\lambda_2}D^1_{0,\lambda_\ell}(\Omega_\ell)), \quad (37)$$

where $\Omega_\ell = (0,\theta_\ell,0)$ in this case. The angular distribution (with $0 \le \theta_\ell \le \pi$) is given by squaring the matrix element

$$I_K(q^2,\theta_\ell) \equiv \frac{d^2\Gamma}{dq^2 d\cos\theta_\ell} = \mathcal{N}\sum_{\lambda_1,\lambda_2}|\hat{\mathcal{M}}_{\lambda_1,\lambda_2}|^2. \quad (38)$$

Using (37) one obtains

$$\begin{aligned} I_K^{(0)} &= G^{(0)}(q^2) + G^{(1)}(q^2)D^1_{0,0}(\Omega_\ell) + G^{(2)}(q^2)D^2_{0,0}(\Omega_\ell) \\ &= G^{(0)}(q^2) + G^{(1)}(q^2)P_1(\cos\theta_\ell) + G^{(2)}(q^2)P_2(\cos\theta_\ell) \\ &= G^{(0)}(q^2) + G^{(1)}(q^2)\cos\theta_\ell + G^{(2)}(q^2)\frac{1}{2}(3\cos^2\theta_\ell - 1), \end{aligned} \quad (39)$$

where we used $P_l(\cos\theta_\ell) = D^l_{0,0}(\Omega_\ell)$ and $D^0_{0,0}(\Omega_\ell) = 1$. For convenience, we have given results in terms of the explicit angle θ_ℓ using equation (30). The superscript (0) is again a reminder that the restriction to $l_\ell \le 2$ is a consequence of only including S_ℓ- and P_ℓ-waves in (37). The explicit functions $G^{(0,1,2)}$, whose q^2-dependence we omit hereafter, are given in appendix D in equation (D2) in terms of HAs.[8]

With respect to the parametrization of the angular distribution used in the experimental community, [38]

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_\ell} = \frac{3}{4}(1-F_H)(1-\cos^2\theta_\ell) + \frac{1}{2}F_H + A_{\rm FB}\cos\theta_\ell, \quad (40)$$

the relationship to the $G^{(i)}$ in (39) is given by

$$\Gamma = 2\langle G^{(0)}\rangle,$$
$$A_{\rm FB} = \sigma_X\frac{\langle G^{(1)}\rangle}{2\langle G^{(0)}\rangle},$$
$$F_H = \frac{\langle G^{(0)}\rangle + \langle G^{(2)}\rangle}{\langle G^{(0)}\rangle}, \quad (41)$$

where $\langle X\rangle = \int dq^2 X$ denotes the integration or appropriate binning over q^2 and $\sigma = \pm 1$ depending on the conventions.[9]

[7] In terms of the $g_i(q^2)$ basis, $\mathcal{N}_{\rm bin} = \frac{64}{3}\langle g_{2s}\rangle_{\rm bin}$ and $\mathcal{N}'_{\rm bin} = \frac{16}{3}\sqrt{-\langle g_{2c}\rangle_{\rm bin}\langle g_{2s}\rangle_{\rm bin}}$.

[8] The observables $G^{(l_\ell)}$ and the angular coefficients used in the literature [37] are related by $a(q^2) = G^{(0)} - \frac{1}{2}G^{(2)}$, $b(q^2) = G^{(1)}$ and $c(q^2) = \frac{3}{2}G^{(2)}$ where $I_K^{(0)} = a + b\cos\theta_\ell + c\cos^2\theta_\ell$.

[9] In our conventions by definition $\sigma_{\rm GHZ} = 1$ and the translation to the LHCb conventions [38] are as follows $\sigma_{\rm GHZ} = \sigma(B^\pm)$ and $\sigma_{\rm GHZ} = -\sigma(B^0,\bar{B}^0)$. The charged and neutral decays are different because the neutral mode, being observed in K_S, is not self-tagging. Comparing with the theory paper [37] we find $\sigma_{\rm GHZ} = -\sigma_{\rm BHP}$ for both charged and neutral modes.

TABLE II. Moments $M_m^{l_K,l_\ell}$ in terms of $G_m^{l_K,l_\ell}$ as defined by Eq. (45) with factor of proportionality $c_m^{l_K,l_\ell}$ evaluated with (46).

$M_m^{l_K,l_\ell}$	$G_0^{0,0}$	$\frac{1}{3}G_0^{0,1}$	$\frac{1}{5}G_0^{0,2}$	$\frac{1}{5}G_0^{2,0}$	$\frac{1}{15}G_0^{2,1}$	$\frac{1}{25}G_0^{2,2}$	$\frac{1}{30}G_1^{2,1}$	$\frac{1}{50}G_1^{2,2}$	$\frac{1}{50}G_2^{2,2}$

IV. METHOD OF TOTAL AND PARTIAL MOMENTS

The MoM is a powerful tool to extract the angular observables $G_m^{l_K,l_\ell}$ by the use of orthogonality relations. In B physics, for example, the method has been applied to $B \to J/\Psi(\to \bar{\ell}\ell)K^*(\to K\pi)$ type decays [30] during the first B-factory era.

In experiment the angular information on $B \to K^*\ell\ell$ has been extracted through the likelihood fit method, at the level of $I_{K^*}^{(0)}$ [27], and it has also been suggested for analysis at the amplitude level [39]. A possible advantage of the MoM over the likelihood fit is that it is less sensitive to theoretical assumptions. More precisely, one can test each angular term independent of the rest of the distribution. Generically the fourfold angular distribution can be expanded over the complete set of functions $\Omega_m^{l_K,l_\ell}$ (29)

$$I_{K^*}(q^2,\Omega_K,\Omega_\ell) = \sum_{l_K,l_\ell\geq 0}\sum_{m=0}^{\min(l_K,l_\ell)} \mathrm{Re}[G_m^{l_K,l_\ell}\Omega_m^{l_K,l_\ell}(\theta_K,\theta_\ell,\phi)], \tag{42}$$

of which the distribution $I_{K^*}^{(0)}$ (28) is a subset. Note that the sum over m does not need to be continued for negative values since I_{K^*} is real-valued. By using the orthogonality properties of the Wigner D-functions (e.g. [40]) with $\Omega = (\alpha,\beta,\gamma)$

$$\int_{-1}^{1} d\cos\beta \int_0^{2\pi} \mathrm{d}\alpha \int_0^{2\pi} \mathrm{d}\gamma D^j_{m,n}(\Omega)\overline{D}^l_{p,q}(\Omega) = \frac{8\pi^2}{2j+1}\delta_{jl}\delta_{mp}\delta_{nq}, \tag{43}$$

the MoM allows us to extract the observables $G_m^{l_K,l_\ell}$ from the angular distribution. In particular one can test for the absence of all higher moments and therefore test very specifically the assumptions made when deriving the distribution $I_{K^*}^{(0)}$ (28). We refer to this method as the method of (total) moments or simply MoM with results given in Sec. IV A. Integrating over a subset of angles, referred to as partial moments, is discussed in Sec. IV B. In the latter case orthogonality does not hold in the generic case and different $G_m^{l_K,l_\ell}$ enter the same moment.

Elements of the MoM have previously been applied to $\Lambda_b \to \Lambda(\to (p,n)\pi)\bar{\ell}_1\ell_2$ [41] and more systematically to the other channels discussed in this paper, crucially including a study of how to account for detector-resolution acceptance effects, in [24]. Our study differs from the latter in that we start at the level of the HAs, and obtain the distribution (42) through a direct computation, whereas the other studies proceed backwards and directly expand the decay distribution in the orthogonal basis of associated Legendre polynomials. Our approach is therefore advantageous in that it provides additional insight, by clarifying the structure of the decay distribution (28) and what type of physics goes beyond it. This is an aspect we return to in Sec. V.

A. Method of total moments

In order to condense the notation slightly we define the scalar product

$$\langle f(\Omega)|g(\Omega)\rangle_{\theta_K\theta_\ell\phi} \equiv \frac{1}{8\pi}\int_{-1}^{1} d\cos\theta_K \int_{-1}^{1} d\cos\theta_\ell \times \int_0^{2\pi} d\phi \overline{f}(\Omega)g(\Omega), \tag{44}$$

normalized such that $\langle 1|1\rangle = 1$. Using $\langle f(\Omega)|g(\Omega)\rangle_{\theta_K\theta_\ell\phi}$ we can thus extract all observables $G_m^{l_K,l_\ell}$ separately from each other, by taking moments[10]

$$M_m^{l_K,l_\ell} \equiv \langle \Omega_m^{l_K,l_\ell}|I_{K^*}(q^2,\Omega_K,\Omega_\ell)\rangle_{\theta_K\theta_\ell\phi} = c_m^{l_K,l_\ell}G_m^{l_K,l_\ell}, \tag{45}$$

where

$$c_m^{l_K,l_\ell} = \frac{1+\delta_{m0}}{2(2l_K+1)(2l_\ell+1)}. \tag{46}$$

Using the equation above the terms in (28) are given in Table II. Furthermore, the orthogonality condition also implies that

$$\begin{aligned} M_m^{j,j'} &= 0, \quad \forall\ m \quad \text{and} \quad j \geq 3 \quad \text{or} \quad j' \geq 3, \\ M_m^{1,j'} &= 0, \quad \forall\ j', m. \end{aligned} \tag{47}$$

Hence the higher and $l_K = 1$ moments vanish, providing a very specific test of the theoretical assumptions behind $I_{K^*}^{(0)}$.

B. Partial moments

The results given previously show how to extract the individual $G_m^{l_K,l_\ell}$. We propose the method of partial moments whereby one integrates only over a subset of angles. The distributions might be regarded as generalizations of uni- and double-angular distributions as these in

[10]The moments $M_m^{l_K,l_\ell}$ and the quantities $S_{l_\ell,l_K,m}$ introduced in [24] are related as follows: $8\pi G_0^{0,0}S_{l_\ell,l_K,m} = G_m^{l_K,l_\ell} = M_m^{l_K,l_\ell}/c_m^{l_K,l_\ell}$.

turn can be viewed as partial moments with respect to unity. The method is effectively a hybrid between the likelihood fit and the total MoM. To this end we define the further scalar products

$$\langle f(\Omega)|g(\Omega)\rangle_{\theta\phi} \equiv \frac{1}{4\pi}\int_{-1}^{1} d\cos\theta \int_0^{2\pi} d\phi \overline{f}(\Omega) g(\Omega),$$
$$\langle f(\Omega)|g(\Omega)\rangle_{\theta_K\theta_\ell} \equiv \frac{1}{4}\int_{-1}^{1} d\cos\theta_K \int_{-1}^{1} d\cos\theta_\ell \overline{f}(\Omega) g(\Omega), \tag{48}$$

again normalized such that $\langle 1|1\rangle = 1$. The orthogonality relation (43) can then be rewritten as

$$\langle D^l_{p,0}(\Omega_\ell)|D^j_{m,0}(\Omega_\ell)\rangle_{\theta_\ell\phi} = \frac{1}{2l+1}\delta_{jl}\delta_{mp}. \tag{49}$$

1. Integrating over θ_ℓ,ϕ: $k^{l_\ell}_m(\theta_K)$-moments

The partial moment over θ_ℓ and ϕ is defined and given by

$$k^{l_\ell}_m(\theta_K) \equiv \langle D^{l_\ell}_{m,0}(\Omega_\ell)|I_{K^*}(q^2,\Omega_K,\Omega_\ell)\rangle_{\theta_\ell\phi} = \frac{1+\delta_{m0}}{2(2l_\ell+1)}\sum_{l_K\geq 0} D^{l_K}_{m,0}(\Omega_K) G^{l_K,l_\ell}_m \tag{50}$$

Assuming the distribution (28) ($l_K = 0, 2$) there are six nonvanishing moments

$$\begin{aligned}
k^0_0(\theta_K) &= G^{0,0}_0 + G^{2,0}_0 D^2_{0,0}(\Omega_K) \\
&= G^{0,0}_0 + \frac{1}{2}(3\cos^2\theta_K - 1)G^{2,0}_0, \\
k^1_0(\theta_K) &= \frac{1}{3}(G^{0,1}_0 + G^{2,1}_0 D^2_{0,0}(\Omega_K)) \\
&= \frac{1}{3}\left(G^{0,1}_0 + \frac{1}{2}(3\cos^2\theta_K - 1)G^{2,1}_0\right), \\
k^2_0(\theta_K) &= \frac{1}{5}(G^{0,2}_0 + G^{2,2}_0 D^2_{0,0}(\Omega_K)) \\
&= \frac{1}{5}\left(G^{0,2}_0 + \frac{1}{2}(3\cos^2\theta_K - 1)G^{2,2}_0\right), \\
k^1_1(\theta_K) &= \frac{1}{6}G^{2,1}_1 D^2_{1,0}(\Omega_K) = \frac{-1}{6}\sqrt{\frac{3}{8}}\sin 2\theta_K G^{2,1}_1, \\
k^2_1(\theta_K) &= \frac{1}{10}G^{2,2}_1 D^2_{1,0}(\Omega_K) = \frac{-1}{10}\sqrt{\frac{3}{8}}\sin 2\theta_K G^{2,2}_1, \\
k^2_2(\theta_K) &= \frac{1}{10}G^{2,2}_2 D^2_{2,0}(\Omega_K) = \frac{1}{10}\sqrt{\frac{3}{8}}\sin^2\theta_K G^{2,2}_2,
\end{aligned} \tag{51}$$

where we used $D^0_{0,0}(\Omega_K) = 1$. As was the case in the MoM, with respect to the distribution $I^{(0)}_{K^*}$ higher partial moments vanish

$$k^{l_\ell}_m(\theta_K) = 0 \quad \forall\ l_\ell \geq 3, \quad \forall\ m. \tag{52}$$

2. Integrating over θ_K, ϕ: $l^{l_\ell}_m(\theta_\ell)$-moments

The partial moment over θ_K and ϕ is defined in complete analogy with the previous partial moment (50) by,

$$l^{l_K}_m(\theta_\ell) \equiv \langle D^{l_K}_{m,0}(\Omega'_K)|I_{K^*}(q^2,\Omega'_K,\Omega'_\ell)\rangle_{\theta_K\phi} = \frac{1+\delta_{m0}}{2(2l_K+1)}\sum_{l_\ell\geq 0} D^{l_\ell}_{m,0}(\Omega_K) G^{l_K,l_\ell}_m, \tag{53}$$

where we make use of the reparametrization of angles given in (29). Again assuming the distribution (28) ($l_\ell = 0, 1, 2$) there are four nonvanishing moments

$$\begin{aligned}
l^0_0(\theta_\ell) &= G^{0,0}_0 + G^{0,1}_0 D^1_{0,0}(\Omega'_\ell) + G^{0,2}_0 D^2_{0,0}(\Omega'_\ell) \\
&= G^{0,0}_0 + \cos\theta_\ell G^{0,1}_0 + \frac{1}{2}(3\cos^2\theta_\ell - 1)G^{0,2}_0, \\
l^2_0(\theta_\ell) &= \frac{1}{5}(G^{2,0}_0 + G^{2,1}_0 D^1_{0,0}(\Omega'_\ell) + G^{2,2}_0 D^2_{0,0}(\Omega'_\ell)) \\
&= \frac{1}{5}\left(G^{2,0}_0 + \cos\theta_\ell G^{2,1}_0 + \frac{1}{2}(3\cos^2\theta_\ell - 1)G^{2,2}_0\right), \\
l^2_1(\theta_\ell) &= \frac{1}{10}(G^{2,1}_1 D^1_{1,0}(\Omega'_\ell) + G^{2,2}_1 D^2_{1,0}(\Omega'_\ell)) \\
&= \frac{-1}{10\sqrt{2}}\left(\sin\theta_\ell G^{2,1}_1 + \sqrt{\frac{3}{4}}\sin 2\theta_\ell G^{2,2}_1\right), \\
l^2_2(\theta_\ell) &= \frac{1}{10}G^{2,2}_2 D^2_{2,0}(\Omega'_\ell) = \frac{1}{10}\sqrt{\frac{3}{8}}\sin^2\theta_\ell G^{2,2}_2,
\end{aligned} \tag{54}$$

where we used $D^0_{0,0}(\Omega'_\ell) = 1$. With respect to the distribution $I^{(0)}_{K^*}$ higher partial moments vanish

$$l^{l_\ell}_m(\theta_\ell) = 0, \quad \forall\ l_K \geq 3, \quad \forall\ m \quad \text{and} \quad l_K = 1, \quad \forall\ m. \tag{55}$$

3. Integrating over θ_K, θ_ℓ: $p^{l_K,l_\ell}_{m,m'}(\phi)$-moments

Finally, we can consider projecting on to moments of the form $D^l_{m,0}(\Omega_K)D^{l'}_{m',0}(\Omega'_\ell)$ with respect to θ_K, θ_ℓ. In this case the full orthogonality relation (43) no longer holds, but due to (27) there exist selection rules as to which of the G^{l_K,l_ℓ}_m can contribute to the partial moments

$$p^{l_K,l_\ell}_{m,m'}(\phi) \equiv \langle D^{l_K}_{m,0}(0,\theta_K,0)D^{l\ell}_{m',0}(0,\theta_\ell,0)|I_{K^*}(q^2,\Omega_K,\Omega_\ell)\rangle_{\theta_K\theta_\ell}. \tag{56}$$

Assuming $I^{(0)}_{K^*}$ a few nonvanishing moments are

$$p^{0,0}_{0,0}(\phi) = \frac{1}{6}(6G^{0,0}_0 + \mathrm{Re}[e^{-2i\phi}G^{2,2}_2]), \qquad p^{0,1}_{0,0}(\phi) = \frac{1}{3}G^{0,1}_0,$$
$$p^{0,2}_{0,0}(\phi) = \frac{1}{30}(6G^{0,2}_0 - \mathrm{Re}[e^{-2i\phi}G^{2,2}_2]), \qquad p^{2,0}_{0,0}(\phi) = \frac{1}{30}(6G^{2,0}_0 - \mathrm{Re}[e^{-2i\phi}G^{2,2}_2]),$$
$$p^{2,1}_{0,0}(\phi) = \frac{1}{15}G^{2,1}_0, \qquad p^{2,2}_{0,0}(\phi) = \frac{1}{150}(6G^{2,2}_0 + \mathrm{Re}[e^{-2i\phi}G^{2,2}_2]),$$
$$p^{2,1}_{1,1}(\phi) = \frac{1}{15}\mathrm{Re}[e^{-i\phi}G^{2,1}_1], \qquad p^{2,2}_{1,1}(\phi) = \frac{1}{25}\mathrm{Re}[e^{-i\phi}G^{2,2}_1]. \tag{57}$$

A consequence of the fact that the full orthogonality of the Wigner functions has been lost is that higher moments contain lower G-functions. As an interesting example we quote

$$p^{4,1}_{2,0}(\phi) = \frac{1}{9\sqrt{10}}(G^{0,1}_0 + G^{2,1}_0) = \frac{4}{9\sqrt{10}}g_{6c}. \tag{58}$$

This quantity is of some interest since $g_{6c} = 0$ in the Standard Model (SM), as it involves scalar and tensor operators at the level of the dimension-six effective Hamiltonian (12).

V. INCLUDING HIGHER PARTIAL WAVES

The compact form of the angular distribution $I^{(0)}_{K^*}$ (28) is a consequence of the LFA and the restriction to the P_K-wave in the $(K\pi)$-channel. In this section we elaborate on the consequences of going beyond these approximations. The double partial wave expansion is outlined in Sec. V A followed by a qualitative discussion of the effect of higher spin operators and the inclusion of electroweak effects in Secs. V B and V C respectively. In Sec. V D we emphasise how testing for higher moments can be used to diagnose the size of QED corrections. Throughout this section we change the notation from $\ell_1\overline{\ell}_2 \to \ell^+\ell^-$ for the sake of familiarity and simplicity.

A. Double partial wave expansion

In order to discuss the origin of generic terms in the full distribution (42), it is advantageous to return to the amplitude level. Somewhat symbolically we may rewrite the amplitude (19), omitting the sum over J_γ, as

$$\mathcal{A}(B \to K_J(\lambda)(\to K\pi)\ell^+(\lambda_1)\ell^-(\lambda_2)) = \mathcal{A}^{J_\gamma,J_K}_{\lambda,\lambda_\ell}\overline{D}^{J_K}_{\lambda,0}(\Omega_K)\overline{D}^{J_\gamma}_{\lambda,\lambda_\ell}(\Omega_\ell) \tag{59}$$

with $\lambda_\ell = \lambda_1 - \lambda_2$ as defined in (6). The two opening angles θ_K and θ_ℓ allow for two separate partial wave expansions. The partial waves in the θ_K- and θ_ℓ-angles are denoted by $S_K, P_K, \ldots$ and $S_\ell, P_\ell, \ldots$ respectively.

Throughout this work we mostly restricted ourselves to $K_J = K^*$ thereby imposing $J_K = 1$ i.e. a P_K-wave. The signal of K^* is part of the $(K\pi)$ P_K-wave. The importance of considering the S_K-wave interference through $K^*_0(800)$ [also known as $\kappa(800)$] was emphasized a few years ago in [42]. The separation of the various partial waves in the $(K\pi)$-channel is a problem that can be solved experimentally e.g. [43]. We refer the reader to Ref. [19] for a generic study of the lowest partial waves at high q^2.

The second partial wave expansion originates from the lepton angle θ_ℓ, which will be the main focus hereafter. By restricting ourselves to the dimension-six effective Hamiltonian equation (12) as well as the lepton-pair factorization approximation (LFA)[11] only S_ℓ- and P_ℓ-waves were allowed [cf. Eq. (22)], bounding $l_\ell \leq 2$ in (42). This pattern is broken by the inclusion of higher spin operators and nonfactorizable corrections between the lepton pair and the quarks. It is therefore important to be able to distinguish these two effects from each other.

B. Qualitative discussion of effects of higher spin operators in H^{eff}

Operators of higher dimension are suppressed and neglected in the standard analysis. Operators of higher spin in the lepton and quark parts are necessarily of higher dimension and bring in new features. An operator of (lepton- and quark-pair) spin j is given by

$$O^{(j)} = \overline{s}_L\Gamma^{(j-)}_{\mu_1\ldots\mu_j}b\overline{\ell}\Gamma^{(j+)\mu_1\ldots\mu_j}\ell \tag{60}$$

with $\Gamma^{(j\pm)}_{\mu_1\ldots\mu_j} \equiv \gamma_{\{\mu_1}D^\pm_{\mu_2}\ldots D^\pm_{\mu_j\}}$, $D^\pm \equiv \overleftarrow{D} \pm \vec{D}$, with $\vec{D}$ the directional covariant derivative and curly brackets denoting symmetrization in the Lorentz indices. In passing let us note that in this notation $O^{(1)} = O_V \equiv O_9$ with O_V defined in (12). The operator (60) is of dimension $d_{O^{(j)}} = 4 + 2j$ and the corresponding Wilson coefficients are suppressed by powers of m_W. Neglecting electroweak corrections and including the dimensional estimate of the matrix elements the leading relative contributions are given by $(m_b/m_W)^{2(j-1)}$ where $(m_b/m_W) \sim 6 \times 10^{-3}$.

Operators of the form (60) present new opportunities to test physics beyond the SM provided that their contribution

[11]We remind the reader that in the LFA no electroweak gauge bosons are exchanged between the lepton pair and other particles when calculating the matrix element. This is the same approximation that is relevant for the endpoint relations [18,44].

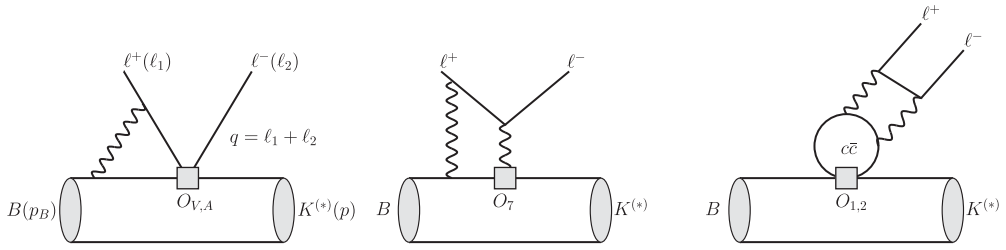

FIG. 3. Examples of virtual QED corrections to $B \to K\ell^+\ell^-$, where either a photon is exchanged between the decaying b-quark and a final state lepton, with effective operators $O_{V,A}$ (left) and O_7 (middle), or a second photon is emitted by the charm loop (right). Other topologies relevant for higher moments include the interaction of the leptons with the spectator as well as the B- and $K^{(*)}$-meson.

is larger than that of the breaking of lepton factorization through electroweak corrections. The operator $O^{(2)}$, for example, gives rise to nonvanishing moments of the type $G_2^{2,4}$ in $B \to K^*(\to K\pi)\overline{\ell}^+\ell^-$ and $G_4^{4,4}$ in $B \to K_2(\to K\pi)\overline{\ell}^+\ell^-$, [45] both of which are absent in the LFA.

C. Qualitative discussion of QED corrections

The $B \to K\ell^+\ell^-$ channel allows the discussion of the consequences of going beyond the LFA in a simplified setup, and is of particular relevance because of a recent LHCb measurement [25].

In the LFA (38) the single opening angle θ_ℓ of the decay is restricted to $l_\ell \leq 2$ moments in $I_K^{(0)}$ (42). More precisely, $l_\ell \leq 2j$ with $O^{(j)}$ as in (60) [see also the discussion following Eq. (28)]. From the viewpoint of a generic $1 \to 3$ decay there is no reason for this restriction, as it is only the sum of the total (orbital and spin) angular momentum that is conserved. However, in the LFA the $B \to K[\ell^+\ell^-]$ decay mimics a $1 \to 2$ process, imposing this constraint. This pattern is broken by exchanges of photons and W- and Z-bosons, as depicted in Fig. 3 for a few operators relevant to the decay. The W and Z are too heavy to impact on the matrix elements, but their effect is included in the Wilson coefficient.

As stated above QED corrections turn the decay into a true $1 \to 3$ process, and this necessitates a reassessment of the kinematics. By crossing the process can be written as a $2 \to 2$ process

$$B(p_B) + \ell^-(-\ell_1) \to K(p) + \ell^-(\ell_2), \tag{61}$$

with Mandelstam variables $s = (p + \ell_2)^2$, $t = (\ell_1 + \ell_2)^2 = q^2$ and $u = (p + \ell_1)^2$,

$$s[u] = \frac{1}{2}[(m_B^2 + m_K^2 + 2m_\ell^2 - q^2) \pm \beta_\ell \sqrt{\lambda(m_B^2, m_K^2, q^2)} \cos\theta_\ell], \tag{62}$$

obeying the Mandelstam constraint $s + t + u = m_B^2 + m_K^2 + 2m_\ell^2$. Crucially, the kinematic variables s and u become *explicit* functions of the angle θ_ℓ. In a generic computation these variables enter (poly)logarithms, which when expanded give contributions to any order l_ℓ in the Legendre polynomials. This statement applies at the amplitude level and therefore also to the decay distribution (39)

$$\frac{d^2\Gamma(B \to K\ell^+\ell^-)}{dq^2 d\cos\theta_\ell} = \sum_{l_\ell \geq 0} G^{(l_\ell)} P_{l_\ell}(\cos\theta_\ell). \tag{63}$$

The $B \to K\ell\ell$ moments are simply given by

$$M_{\overline{\ell}\ell}^{(l_\ell)} = \int_{-1}^{1} d\cos\theta_\ell P_{l_\ell}(\cos\theta_\ell) \frac{d^2\Gamma(B \to K\ell^+\ell^-)}{dq^2 d\cos\theta_\ell} = \frac{1}{2l_\ell + 1} G_{\overline{\ell}\ell}^{(l_\ell)} \tag{64}$$

where we have introduced a lepton-subscript for further reference. In the SM the effects are dependent on the lepton mass, for example through logarithms of the form $\ln(m_\ell/m_b)$ times the fine structure constant. There are new qualitative features of which we would like to highlight the following two:

(i) Both vector and axial couplings $O_{V(A)} = O_{9(10)}$ (12) contribute to any moment $l_\ell \geq 0$. In the LFA l_ℓ-odd terms (forward-backward asymmetric) arise from broken parity through interference of O_V and O_A (12). The physical interpretation is that there is a preferred direction for charged leptons in the presence of the charged quarks of the decay. In the specific diagram Fig. 3 (left) it is the charge of the b-quark which attracts or repels the charged lepton(s) with definite preference. It is possible that one can establish a higher degree of symmetry by using charge-averaged forward-backward asymmetries.

(ii) A key question is how the moments vary in l_ℓ. In the absence of a computation a precise answer is not possible. Nevertheless we can assess the question semiquantitatively by considering for example the triangle graph between the photon, a lepton and the b-quark in Fig. 3 (left) and the corresponding one with the s-quark. Neglecting the Dirac

structures the triangle graph is given by $C_0(m_\ell^2, p_B^2, s[\cos\theta_\ell], m_\ell^2, 0, m_b^2)$.[12] Expanding this function in partial waves $C_0 = \sum_{l_\ell \geq 0} C_0^{(l_\ell)} P_{l_\ell}(\cos\theta_\ell)$ we find that $|C_0^{(l_\ell)}|$ does fall off in l_ℓ. Averaging over several configurations (cf. footnote[13]) we conclude that the $l_\ell = 2$ (D-wave) contribution is suppressed by approximately a factor of 2 with respect to $l_\ell = 0$, with a slightly steeper falloff with increasing l_ℓ for the b-quark versus s-quark vertex correction. Note the graph where the photon couples to the other lepton comes with a different Dirac structure and is not obtainable through a straightforward symmetry prescription. We therefore think that it is sensible to consider those graphs separately. We stress that this semiquantitative analysis does not replace a complete QED computation, which would include corrections to Wilson coefficients, all virtual corrections and importantly also the real photon emission.

We now turn to the most important consideration, the relative size of the QED corrections versus higher spin operators. For effective field theories of the type $\langle H^{\rm eff}\rangle \sim C^{(j)}(\mu_F)\langle O^{(j)}(\mu_F)\rangle$ (12), the precise separation scale μ_F is arbitrary to a certain degree and effects are therefore contained in the Wilson coefficients as well as the matrix elements. We find it convenient to discuss the effect at the level of the Wilson coefficients. For the latter QED corrections arising from modes from m_W to $\mu_F \simeq m_b$ can be absorbed into a tower of the higher spin operators $O^{(j)}$ (60). The leading contribution to the corresponding Wilson coefficients from the initial matching procedure and the mixing due to QED behaves parametrically as

$$C^{(j)} = \frac{\mathcal{O}(1)}{(m_W^2)^j} + \alpha f_j \cdot \left(\frac{m_W^2}{m_b^2}\right)^{(j-1)} \frac{\mathcal{O}(1)}{(m_W^2)^j}, \quad \text{for } j \geq 1, \tag{65}$$

where we have implicitly used $\mu_F = m_b$ in $\langle H^{\rm eff}\rangle \sim C^{(j)}(\mu_F)\langle O^{(j)}(\mu_F)\rangle$. Above α is the fine structure constant and f_j parametrizes the comparatively moderate falloff of the higher moments due to QED. In the SM one therefore expects QED effects to dominate over those due to higher spin operators, except for $j = 2$ where they could be comparable [45]. At the level of matrix elements this hierarchy could even shift further toward QED as a result of infrared enhancements through $\ln(m_\ell/m_b)$-contributions.

The discussion of $B \to K^*(\to K\pi)\ell^+\ell^-$ is similar, but involves the kinematics of a $1 \to 4$ decay. The decay distribution becomes a generic function of all three angles θ_ℓ, θ_K and ϕ. It should be added that the selection of the $K^* \to K\pi$ signal (P_K-wave) restricts $l_K = 0, 2$.

D. On the importance of testing for higher moments for $B \to K^{(*)}\bar{\ell}^+\ell^-$

We have stressed throughout the text that it is of importance to probe for moments that are vanishing in the decay distributions $I_{K^*}^{(0)}$ (28) of $\overline{B} \to \overline{K}^*(\to K\pi)\ell\ell$) and $I_K^{(0)}$ (38) of $\overline{B} \to \overline{K}\ell\ell$ respectively. In this section we highlight specific cases of current experimental anomalies in exclusive decay modes where their nature might be clarified using an analysis of (higher) moments.

1. Diagnosing QED background to R_K

In the SM the decays $B^+ \to K^+e^+e^-$ and $B^+ \to K^+\mu^+\mu^-$ are identical up to phase-space lepton mass effects and electroweak corrections. The observable

$$R_K|_{[q^2_{\rm min}, q^2_{\rm max}]} \equiv \frac{\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathcal{B}(B^+ \to K^+e^+e^-)}\Big|_{[q^2_{\rm min}, q^2_{\rm max}]} \tag{66}$$

has been put forward in Ref. [46] as an interesting test of lepton flavour universality (LFU). Above $q^2_{\rm min\,/\,max}$ stands for the bin boundaries. Neglecting electroweak corrections the SM prediction is $R_K|_{[1,6]\ \rm GeV^2} \simeq 1.0003(1)$ [47], which is at 2.6σ-tension with the LHCb measurement at 3 fb^{-1} [25]

$$R_K = 0.745^{+0.090}_{-0.074}(\text{stat}) \pm 0.036(\text{syst}). \tag{67}$$

Previous measurements [48,49], with much larger uncertainties, were found to be consistent with the SM as well as (67). This led to investigations of physics beyond the SM with $C_9^{\bar{e}e} \neq C_9^{\bar{\mu}\mu}$ (where $O_9^{\bar{\ell}\ell} \equiv \bar{b}\gamma_\alpha s\bar{\ell}\gamma^\alpha\ell$) amongst other variants for which we quote a few recent works [50–59] as well as the general review [60] for further references.

Let us summarize the aspects of QED corrections which are of relevance for the discussion below: (i) they break lepton factorization and therefore give rise to higher moments, and (ii) they depend on the lepton mass, for example through logarithmic terms of $\ln(m_\ell/m_b)$. In view of the lack of a full QED computation[14] we suggest diagnosing the size of QED corrections, as well as their lepton dependence, by experimentally assessing higher

[12] We use conventions for the Passarino-Veltman function $C_0(p_1^2, p_2^2, p_3^2, m_1^2, m_2^2, m_3^2)$ such that the two-particle cuts begin at $p_1^2 \geq (m_1 + m_2)^2$, $p_2^2 \geq (m_2 + m_3)^2$ and $p_3^2 \geq (m_3 + m_1)^2$.

[13] We have refined this analysis by taking into account that the b- and s-quark only carry a fraction of the momentum of the corresponding mesons. This amounts to the substitution $p_B^2 \to (p_B - xp)^2$ and $s[u] \to (\ell_2[\ell_1] + xp)^2$ with x being the momentum fraction carried by the s-quark. For the vertex diagrams one expects the Feynman mechanism (i.e. $x \simeq 0$) to dominate. This changes when spectator corrections are taken into account.

[14] A partial result, photon emission from initial and final state, was reported in [61].

moments.[15] The latter is directly relevant for R_K. Let us be slightly more concrete and define the normalized angular functions as follows $\hat{G}^{(l_\ell)}_{\bar{\ell}\ell} \equiv G^{(l_\ell)}_{\bar{\ell}\ell}/(2G^{(0)}_{\bar{\ell}\ell})$ (63) (in this convention $2G^{(0)}_{\bar{\ell}\ell} = d\Gamma(B \to K\ell\ell)/dq^2$, $\hat{G}^{(1)}_{\bar{\ell}\ell} = A_{\rm FB}$ and $\hat{G}^{(2)}_{\bar{\ell}\ell} = (F_H - 1)/2$ in the notation of [37]). We would like to stress the following points:

(i) *How to distinguish QED corrections from higher dimensional operators:* both contributions give rise to higher moments but crucially the QED corrections dominate for moments of increasing l_ℓ, cf. the discussion at the end of Sec. V C and specifically Eq. (65). A J_γ-wave at the amplitude level contributes to a $l_\ell = J_\gamma + 1$ moment through interference with the SM P_ℓ-wave. We conclude that QED and higher spin operators could be comparable for $\hat{G}^{(3)}_{\bar{\ell}\ell}$ but for $\hat{G}^{(l_\ell>3)}_{\bar{\ell}\ell}$ one would expect the former to dominate.[16]

(ii) *Lepton-flavor dependence of QED corrections:* differences between $\hat{G}^{(l_\ell\geq 3)}_{\bar{\mu}\mu}$ and $\hat{G}^{(l_\ell\geq 3)}_{\bar{e}e}$ in the range above $q^2 > 1$ GeV^2 indicate the importance of the flavor dependence. This gives an indication on how much the branching fractions (zeroth moments) and therefore R_K is affected by QED through lepton mass effects. Note that due to $\ln(m_\ell/m_b)$-effects it is conceivable that $\hat{G}^{(l_\ell\geq 3)}_{\bar{\mu}\mu}$ is small, say $O(1\%)$, but that $\hat{G}^{(l_\ell\geq 3)}_{\bar{e}e}$ is larger. Note for example that $\hat{G}^{(1)}_{\mu\mu} = A_{\rm FB}$ is consistent with the SM prediction excluding QED, which is $O(m_\mu)$, within errors in the few percent range [38].

2. Combinatorial background in $B \to K^*\mu^+\mu^-$ below the narrow charmonium resonance region

A characteristic feature of $B \to K^{(*)}\ell^+\ell^-$ transitions is the large contribution to the branching fraction through the intermediate narrow charmonium states J/Ψ and $\Psi(2S)$. For example $\mathcal{B}(B \to KJ/\Psi)\mathcal{B}(J/\Psi \to \mu^+\mu^-) \simeq (8 \times 10^{-4})(6 \times 10^{-2}) \simeq 5 \times 10^{-5}$ is three orders of magnitude larger than the measured differential branching fraction, $d\mathcal{B}(B \to K^*\mu^+\mu^-)/dq^2 \simeq 2 \times 10^{-8}/\text{GeV}^2$ [64], well below the narrow charmonium resonances region. It is therefore legitimate to be concerned with possible combinatorial backgrounds in this region.

Assuming that such backgrounds are relevant this raises the question as to how they can be distinguished from the signal event. In the case where they can be absorbed into the background fit-function they would not impact on the analysis. Whether or not this is the case is a nontrivial question. Pragmatically, however, background events can be expected to perturb the hierarchy of the moments as compared to the true signal event. One would expect the background events to fall off only slowly for higher moments in the lepton partial wave.[17] Hence the size of these effects can be diagnosed through the measurement of higher moments as a function of q^2, independent of model assumptions. By the latter we mean that higher moments peaking below the charmonium resonances will be indicative of the type of combinatorial background mentioned above.

A possible example of such backgrounds is the process $B \to K\mu^+\mu^-\gamma$ where the photon is not detected but energetic enough to cause a significant downward shift in $q^2 = (\ell_1 + \ell_2)^2$. Such an event would be rejected as a $B \to K\mu^+\mu^-$ signal because the reconstructed B-mass $m_{K\mu\mu}$ would fall outside the signal window [i.e. $m_{K\mu\mu} < m_B - \Delta$ and $\Delta \simeq O(100\ \text{MeV})$]. If additionally a π-meson from the underlying event is detected, the event could conspire to enter the signal window of $B \to K^*(\to K\pi)\mu^+\mu^-$ (i.e. $m_{K\pi\mu\mu} \simeq m_B$ and $m_{K\pi} \simeq m_{K^*}$). It is therefore conceivable that the small chance of the events described above is overcome by the enhancement by three orders of magnitude of the J/Ψ transition. If such events are present and not rejected then this leads to a bias in $B \to K^*(\to K\pi)\mu^+\mu^-$ transitions below the narrow charmonium resonances. More precisely, denoting the momentum of the undetected photon by r, the shift in q^2 is as follows $q^2 \simeq m^2_{J/\Psi} = (\ell_1 + \ell_2 + r)^2 \to q^2_{\rm signal} \equiv (\ell_1 + \ell_2)^2 < m^2_{J/\Psi}$.

This is particularly relevant as some of the anomalies from the LHCb measurements, in particular the angular observable P_5', are most pronounced in bins just below the J/Ψ-resonance [26,27]. To what extent such operators correspond to new physics in $O_9 \equiv O_V$ [65,66] or effects from charm resonances [67] is a difficult question since they contribute to the same helicity amplitude. They can be distinguished from each other by analyzing the q^2-spectrum of the observables and by the determination of the strong phases which can originate from the charm resonances [67]. This could be through the determination of the complex-valued residues of the resonance poles [67], or

[15] Collinear photon emission in the inclusive case was studied recently in [62]. The additional photon of course leads to terms which go beyond the $I^{(0)}_K$ angular distribution. Note, in view of the presence of these terms through virtual corrections they also have to be present in real emission by virtue of the Bloch-Nordsieck QED infrared cancellation theorem [63]. The authors [62] find within their approximation that the third and fourth moment are two orders of magnitude smaller than the leading contributions. This is in the expected parametric range but one cannot draw precise conclusions on the size of this effect for the exclusive channels discussed in this paper.

[16] Another criterion could be that corrections from higher spin operators are uniform in the lepton mass provided that lepton flavor universality is unbroken. This is though delicate since the measurement of R_K questions this aspect.

[17] Similar things can be said about the hadronic partial wave, but as the detection of the P_K-wave is part of the signal selection the presence of such higher waves would have less influence. However, the remaining background might impact on the S_K-wave, which does matter since the S_K-wave enters the analysis.

simply the strong phase in the region below the q^2-resonance through $\mathrm{Im}[G_1^{2,1}] \sim P_6'$, which corresponds to the imaginary part of $\mathrm{Re}[G_1^{2,1}] \sim P_5'$ (33).

VI. CONCLUSIONS

In this work we have generalized the standard helicity formalism to effective field theories of the $b \to s\ell\ell$-type. The framework applies to any semileptonic and radiative decay. The formalism has been used to derive the angular distributions $I_{K^*}^{(0)}$ (28) and $I_K^{(0)}$ (39) for nonequal lepton masses with the full dimension-six effective Hamiltonian, including in particular scalar and tensor operators. Explicit results for $\overline{B} \to \overline{K}^*\ell_1\overline{\ell}_2$ and for $\overline{B} \to \overline{K}\ell_1\overline{\ell}_2$ can be found in Appendices C and D respectively as well as a *Mathematica* notebook (notebookGHZ.nb) provided in the arXiv version [29]. Comments on differences conversion of observables between theory and experiment with the literature are reported in Appendix C 2 a. Minor discrepancies in tensor contributions with respect to previous results are discussed in Appendix C 1 b.

The approach clarifies how the lepton factorization approximation determines the specific form of the angular distributions $I_{K^*}^{(0)}$ and $I_K^{(0)}$, and how these distributions are extended by the inclusion of virtual and real QED corrections, as well as higher-spin operators in the effective Hamiltonian. Higher-dimensional spin operators provide new opportunities to search for physics beyond the SM. We have argued that, within the SM, QED effects and higher-spin operators can be distinguished from each other by their differing falloff behavior in increasingly higher moments in the θ_ℓ-angle.[18]

Assessing higher moments can shed light on current anomalies with respect to the SM. We have argued (cf. Sec. V D 1) that higher moments in $B \to K\ell^+\ell^-$ ($\ell = e, \mu$) are a window into QED corrections and therefore of importance with regard to the R_K measurement [25]. In view of tensions of angular predictions in $B \to K^*\mu^+\mu^-$ with experiment [26,27], the higher moments can be of help in assessing their origin, such as the possible leakage of J/Ψ events into the lower nearby q^2-bins (cf. Sec. V D 2). As another example we mention the $R(D^{(*)}) = \mathcal{B}(B \to D^{(*)}\tau\nu)/\mathcal{B}(B \to D^{(*)}\mu\nu)$ ratio measurement [68–70], suggestive of some tension with the SM. A higher moment analysis could again be useful in assessing the impact of QED, lepton mass or cross channel backgrounds on these results.

To measure and bound higher moments is relevant as their contributions can bias likelihood fits. We therefore encourage the investigation of higher moments in several experimental channels from the various perspectives discussed above.

[18] In addition higher-spin operators can be distinguished from QED corrections by universality in the lepton flavor. However, it should be kept in mind that lepton-universality is questioned by the R_K measurement.

ACKNOWLEDGMENTS

We are grateful to Simon Badger, Damir Becirevic, Tom Blake, Christoph Bobeth, Marcin Chrzaszcz, Peter Clarke, Sebastian Descotes-Genon, Martin Gorbahn, Enrico Lunghi, Joaquim Matias, Matthias Neubert, Kostas Petridis, Alexey Petrov, Maurizio Piai, Steve Playfer, Nico Serra, David Straub, Olcyr Sumensari, Javier Virto, Renata Zukanov and in particular to Greig Cowan for many useful discussions. J. G. acknowledges the support of an STFC studentship (Grant No. ST/K501980/1). M. H. acknowledges support from the Doktoratskolleg "Hadrons in Vacuum, Nuclei and Stars" of the Austrian Science Fund, FWF DK W1203-N16.

Note added.—Recently a paper using the helicity formalism for $B \to K^*\ell^+\ell^-$ appeared [71]. The paper uses the standard Jacob-Wick formalism and therefore includes HAs of definite spin. This is an approximation that holds up to lepton mass corrections in the SM and does not allow the inclusion of scalar operators for example.

APPENDIX A: RESULTS RELEVANT FOR ALL DECAY MODES

1. Decomposition of $SO(3,1)$ into $SO(3)$ up to spin 2

The aim of this Appendix is to give some more detail about the decomposition (15) and in particular extend it to the two-index case, which includes the discussion of spin 2, 1, 0.

In Sec. II B it was shown that insertion of the completeness relation (15) corresponds to the decomposition, or branching rule,

$$(1/2,1/2)_{SO(3,1)}|_{SO(3)} \to (\mathbf{1}+\mathbf{3})_{SO(3)}, \tag{A1}$$

where $(1/2,1/2)$ is the irreducible vector Lorentz representation. We remind the reader that the irreducible Lorentz representations, denoted by (j_1, j_2), are characterized by the eigenvalues of the two Casimir operators of $SO(3,1)$. Inserting the completeness relation twice therefore corresponds to taking the tensor product $(1/2,1/2) \otimes (1/2,1/2)$ which decomposes as

$$\begin{aligned}
&((1/2,1/2) \otimes (1/2,1/2))_{SO(3,1)} \\
&\quad = ([(1,1)] \oplus [(1,0) \oplus (0,1)] \oplus (0,0))_{SO(3,1)}|_{SO(3)} \\
&\quad \to ([1\cdot\mathbf{5} \oplus 1\cdot\mathbf{3} \oplus 1\cdot\mathbf{1}] \oplus [2\cdot\mathbf{3}] \oplus 1\cdot\mathbf{1})_{SO(3)} \\
&\quad = (1\cdot\mathbf{5} \oplus 3\cdot\mathbf{3} \oplus 2\cdot\mathbf{1})_{SO(3)}.
\end{aligned} \tag{A2}$$

The double completeness relation

$$g_{\alpha\beta}g_{\gamma\delta} = \delta_{\alpha\beta\gamma\delta} + \delta^t_{\alpha\beta\gamma\delta} + \delta^{tt}_{\alpha\beta\gamma\delta} \tag{A3}$$

can be decomposed

$$\delta_{\alpha\beta\gamma\delta} = \sum_{J=0}^{2}\sum_{\lambda=-J}^{J} \omega_{\alpha\gamma}^{J,\lambda}\overline{\omega}_{\beta\delta}^{J,\lambda},$$
$$\delta_{\alpha\beta\gamma\delta}^{t} = -\sum_{\lambda=-1}^{1} \omega_{\alpha\gamma}^{t,\lambda}\overline{\omega}_{\beta\delta}^{t,\lambda} - \sum_{\lambda=-1}^{1} \omega_{\gamma\alpha}^{t,\lambda}\overline{\omega}_{\delta\beta}^{t,\lambda},$$
$$\delta_{\alpha\beta\gamma\delta}^{tt} = \omega_{\alpha\gamma}^{tt}\overline{\omega}_{\beta\delta}^{tt}, \tag{A4}$$

into parts containing zero, one and two timelike polarization vectors

$$\omega_{\alpha\gamma}^{t,\lambda} = \omega_\alpha(t)\omega_\gamma(\lambda), \qquad \omega_{\alpha\gamma}^{tt} = \omega_\alpha(t)\omega_\gamma(t),$$
$$\omega_{\alpha\gamma}^{J,\lambda} = \sum_{\lambda_1,\lambda_2=-1}^{1} C_{\lambda\lambda_1\lambda_2}^{J11}\omega_\alpha(\lambda_1)\omega_\gamma(\lambda_2), \tag{A5}$$

with $\lambda = \lambda_1 + \lambda_2$ in the first term and the polarization vectors $\omega_\alpha(\lambda)$ are parametrized as

$$\omega^\mu(\pm) = (0, \pm 1, i, 0)/\sqrt{2},$$
$$\omega^\mu(0) = (q_z, 0, 0, q_0)/\sqrt{q^2},$$
$$\omega^\mu(t) = (q_0, 0, 0, q_z)/\sqrt{q^2}, \tag{A6}$$

which we reproduce from (16) for the reader's convenience. A few explanations seem in order. The minus sign in front of $\delta_{\alpha\beta\gamma\delta}^{t}$ in (A4) is due to there being an odd number of timelike polarization vectors. The first, second and third term in (A3) correspond respectively to the (1,1)-, $[(1,0)\oplus(0,1)]$- and (0,0)-terms in (A2). It is convenient to rewrite the double completeness relation (A3) in a form that makes the decomposition into the different spins j explicit

$$g_{\alpha\beta}g_{\gamma\delta} = \sum_{J=0}^{2}\sum_{\lambda=-J}^{J} \epsilon_{\alpha\gamma}^{J,\lambda}\cdot\overline{\epsilon}_{\beta\delta}^{J,\lambda}. \tag{A7}$$

Above the scalar product "$\cdot$" stands for

$$\epsilon_{\alpha\gamma}^{\lambda}\cdot\overline{\epsilon}_{\beta\delta}^{J,\lambda'} = \delta_{J0}[\omega_{\alpha\gamma}^{0,0}\overline{\omega}_{\beta\delta}^{0,0} + \omega_{\alpha\gamma}^{tt}\overline{\omega}_{\beta\delta}^{tt}]$$
$$+ \delta_{J1}[\omega_{\alpha\gamma}^{1,\lambda}\overline{\omega}_{\beta\delta}^{1,\lambda'} - \omega_{\alpha\gamma}^{t,\lambda}\overline{\omega}_{\beta\delta}^{t,\lambda'} - \omega_{\gamma\alpha}^{t,\lambda}\overline{\omega}_{\delta\beta}^{t,\lambda'}]$$
$$+ \delta_{J2}[\omega_{\alpha\gamma}^{2,\lambda}\overline{\omega}_{\beta\delta}^{2,\lambda'}]. \tag{A8}$$

The single completeness relation (15) in the analogous notation of (A7) reads

$$g_{\alpha\beta} = \sum_{J=0}^{1}\sum_{\lambda=-J}^{J} \epsilon_{\alpha}^{J,\lambda}\overline{\epsilon}_{\beta}^{J,\lambda}, \tag{A9}$$

with $\epsilon_\alpha^{J,\lambda} = \delta_{J1}\omega_\alpha(\lambda) + \delta_{J0}\omega_\alpha(t)$.

When applying the double completeness relation to generic decay structures, it can be seen from (A4) that in general one expects two distinct contributions to the amplitude from $\delta_{\alpha\beta\gamma\delta}^{t}$,

$$H_{\mu\nu}L^{\mu\nu} \to -(\mathcal{H}_{t\lambda}\mathcal{L}_{t\lambda} + \mathcal{H}_{\lambda t}\mathcal{L}_{\lambda t}) + \cdots,$$

where $\mathcal{H}_{t\lambda} = H_{\mu\nu}\overline{\omega}_{t,\lambda}^{\mu\nu}$, and analogous notation for $\mathcal{H}_{\lambda t}$, $\mathcal{L}_{t\lambda}$ and $\mathcal{L}_{\lambda t}$. If, however, the objects $H_{\mu\nu}$ and $L^{\mu\nu}$ are both symmetric or antisymmetric in the Lorentz indices, then $\mathcal{H}_{\lambda t}\mathcal{L}_{\lambda t} = \mathcal{H}_{t\lambda}\mathcal{L}_{t\lambda}$ and the two contributions can be combined. We have used this simplification in defining the generalized HAs for the $\overline{B}\to\overline{K}^*\ell_1\overline{\ell}_2$ (23) and $\overline{B}\to\overline{K}\ell_1\overline{\ell}_2$ (36) decays respectively, resulting in the extra factor of 2 associated with the terms $H_\lambda^{T_t}\mathcal{L}_{\lambda_1,\lambda_2}^{T_t}$, $h^{T_t}\mathcal{L}_{\lambda_1,\lambda_2}^{T_t}$ relative to other contributions in the generalized HAs.

2. Additional remarks on effective Hamiltonian

Here we collect a few additional remarks to the effective $b\to s\ell\ell$ Hamiltonian quoted in Eqs. (12), (13). Contributions proportional to $V_{us}V_{ub}^*$ have been neglected. The chromoelectric and chromomagnetic operators O_7 and O_8, along with the contributions of the four-quark operators $O_{1,\ldots,6}$, can be absorbed into O_V through defining an effective Wilson coefficient $C_V^{\rm eff} = C_9^{\rm eff}$. We can rewrite $O_{\mathcal{T}}^{(\prime)} = 1/2(O_T \pm O_{T_5})$, with the latter defined as

$$O_T = \overline{s}\sigma_{\mu\nu}b\overline{\ell}\sigma_{\mu\nu}\ell, \qquad O_{T5} = \overline{s}\sigma_{\mu\nu}\gamma_5 b\overline{\ell}\sigma_{\mu\nu}\ell, \tag{A10}$$

(note: $O_{T5} = \overline{s}\sigma_{\mu\nu}b\overline{\ell}\sigma_{\mu\nu}\gamma_5\ell = -\frac{i}{2}\epsilon^{\alpha\beta\mu\nu}\overline{s}\sigma_{\alpha\beta}b\overline{\ell}\sigma_{\mu\nu}\ell$ with the last equality depending on conventions) and the relation between the Wilson coefficients is therefore

$$C_{\mathcal{T}}^{(\prime)} = C_T \pm C_{T5}, \qquad C_{T(5)} = \frac{1}{2}(C_{\mathcal{T}} \pm C_{\mathcal{T}}') \tag{A11}$$

in the sense that $C_T O_T + C_{T5}O_{T5} = C_{\mathcal{T}}O_{\mathcal{T}} + C_{\mathcal{T}}'O_{\mathcal{T}}'$.

3. Definitions and results of leptonic helicity amplitudes

The calculation of the leptonic helicity amplitudes is an important part of the generalized helicity formalism described in this paper, and the method for their calculation is outlined in [3]. In the Dirac basis of the Clifford algebra, with σ^i as the usual 2×2 Pauli matrices,

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$
$$\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix},$$
$$\gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{A12}$$

the particle u and antiparticle v spinor are given by

$$u\left(\frac{1}{2}\right) = \left(\sqrt{E_1 + m_{\ell_1}}, 0, \quad \sqrt{E_1 - m_{\ell_1}}, 0\right)^T = (\beta_1^+, 0, \beta_1^-, 0)^T,$$
$$u\left(-\frac{1}{2}\right) = \left(0, \sqrt{E_1 + m_{\ell_1}}, 0, \quad -\sqrt{E_1 - m_{\ell_1}}, \right)^T = (0, \beta_1^+, 0, -\beta_1^-)^T,$$
$$v\left(\frac{1}{2}\right) = \left(\sqrt{E_2 - m_{\ell_2}}, 0, \quad -\sqrt{E_2 + m_{\ell_2}}, 0\right)^T = (\beta_2^-, 0, -\beta_2^+, 0)^T,$$
$$v\left(-\frac{1}{2}\right) = \left(0, \sqrt{E_2 - m_{\ell_2}}, 0, \quad \sqrt{E_2 + m_{\ell_2}}\right)^T = (0, \beta_2^-, 0, \beta_2^+)^T,$$

with implicit definition of $\beta_i^\pm \equiv \sqrt{E_i \pm m_{\ell_i}}$. The spinors are normalized as $\overline{u}(\lambda_1)u(\lambda_2) = \delta_{\lambda_1\lambda_2} 2m_{\ell_1}$ and $\overline{v}(\lambda_1)v(\lambda_2) = -\delta_{\lambda_1\lambda_2} 2m_{\ell_2}$. The leptonic HAs (18) contracted with polarization vectors give rise to the HAs $\mathcal{L}_{\lambda_1\lambda_2}$

$$\mathcal{L}^X_{\lambda_1\lambda_2} \equiv \langle \ell_1(\lambda_1)\overline{\ell}_2(\lambda_2)|\overline{\ell}\Gamma^X\ell|0\rangle = \overline{u}(\lambda_1)\Gamma^X v(\lambda_2), \tag{A13}$$

(where $\ell_1 = e^-$ for example) and the $\Gamma^X|_{\lambda_X\to\lambda_\ell}$ ($\lambda_\ell = \lambda_1 - \lambda_2$) as defined in Table 3.1. Using all the equations above the evaluation of the lepton HAs is then straightforward and the results are presented below, for lepton masses $m_{\ell_1} \neq m_{\ell_2}$ in the first set of matrices and $m_{\ell_1} = m_{\ell_2} \equiv m_\ell$ in the second set.[19] The first row (column) corresponds to $\lambda_1(\lambda_2) = -\frac{1}{2}$ and the second row (column) corresponds to $\lambda_1(\lambda_2) = +\frac{1}{2}$. For the $\overline{B} \to \overline{K}^*\ell_1\overline{\ell}_2$ decay mode, ie $\ell_1 = \ell^-$, the lepton HAs are given by

$$\mathcal{L}^V(\lambda_1,\lambda_2) = \begin{pmatrix} \beta_1^+\beta_2^+ - \beta_1^-\beta_2^- & -\sqrt{2}(\beta_1^+\beta_2^+ + \beta_1^-\beta_2^-) \\ -\sqrt{2}(\beta_1^+\beta_2^+ + \beta_1^-\beta_2^-) & \beta_1^+\beta_2^+ - \beta_1^-\beta_2^- \end{pmatrix} \to \begin{pmatrix} 2m_\ell & -\sqrt{2q^2} \\ -\sqrt{2q^2} & 2m_\ell \end{pmatrix},$$
$$\mathcal{L}^A(\lambda_1,\lambda_2) = \begin{pmatrix} \beta_1^+\beta_2^- - \beta_1^-\beta_2^+ & \sqrt{2}(\beta_1^+\beta_2^- + \beta_1^-\beta_2^+) \\ -\sqrt{2}(\beta_1^+\beta_2^- + \beta_1^-\beta_2^+) & \beta_1^-\beta_2^+ - \beta_1^+\beta_2^- \end{pmatrix} \to \begin{pmatrix} 0 & \sqrt{2q^2}\beta_\ell \\ -\sqrt{2q^2}\beta_\ell & 0 \end{pmatrix},$$
$$\mathcal{L}^S(\lambda_1,\lambda_2) = \begin{pmatrix} \beta_1^+\beta_2^- + \beta_1^-\beta_2^+ & 0 \\ 0 & \beta_1^+\beta_2^- + \beta_1^-\beta_2^+ \end{pmatrix} \to \begin{pmatrix} \sqrt{q^2}\beta_\ell & 0 \\ 0 & \sqrt{q^2}\beta_\ell \end{pmatrix},$$
$$\mathcal{L}^P(\lambda_1,\lambda_2) = \begin{pmatrix} \beta_1^+\beta_2^+ + \beta_1^-\beta_2^- & 0 \\ 0 & -\beta_1^+\beta_2^+ - \beta_1^-\beta_2^- \end{pmatrix} \to \begin{pmatrix} \sqrt{q^2} & 0 \\ 0 & -\sqrt{q^2} \end{pmatrix},$$
$$\mathcal{L}^T(\lambda_1,\lambda_2) = \begin{pmatrix} -i\sqrt{2}(\beta_1^+\beta_2^- + \beta_1^-\beta_2^+) & -2i(\beta_1^+\beta_2^- - \beta_1^-\beta_2^+) \\ 2i(\beta_1^+\beta_2^- - \beta_1^-\beta_2^+) & i\sqrt{2}(\beta_1^+\beta_2^- + \beta_1^-\beta_2^+) \end{pmatrix} \to \begin{pmatrix} -i\sqrt{2q^2}\beta_\ell & 0 \\ 0 & i\sqrt{2q^2}\beta_\ell \end{pmatrix},$$
$$\mathcal{L}^{T_t}(\lambda_1,\lambda_2) = \begin{pmatrix} i(\beta_1^+\beta_2^+ + \beta_1^-\beta_2^-) & -i\sqrt{2}(\beta_1^+\beta_2^+ - \beta_1^-\beta_2^-) \\ -i\sqrt{2}(\beta_1^+\beta_2^+ - \beta_1^-\beta_2^-) & i(\beta_1^+\beta_2^+ + \beta_1^-\beta_2^-) \end{pmatrix} \to \begin{pmatrix} i\sqrt{q^2} & -2i\sqrt{2}m_\ell \\ -2i\sqrt{2}m_\ell & i\sqrt{q^2} \end{pmatrix}, \tag{A14}$$

where $\beta_{1,2}^\pm = \sqrt{E_{1,2} \pm m_{\ell_{1,2}}}$ as before. Above we have used $\beta_i^+\beta_i^- \to E\beta_\ell$ for $m_{\ell_{1,2}} \to m_\ell$ since $E^2 = q^2/4$, where E is the energy of either lepton in the rest frame of the lepton pair. Note that the scalar transitions S and P are necessarily diagonal since $\lambda_\ell = \lambda_1 - \lambda_2 = 0$. Timelike vector and axial lepton HAs are integrated into the hadron HAs (C15).

APPENDIX B: DETAILS ON KINEMATICS FOR DECAY MODES

While within the formalism described in this paper it is not essential to consider the full kinematics of the decay, as the evaluation of the hadronic and leptonic HAs can be performed within their respective rest frames, we collect here the kinematics used in calculating the angular distribution using the Dirac trace technology approach [22,23] in order to facilitate comparison. The Källén function that often appears in our results is given by

$$\lambda(a,b,c) \equiv a^2 + b^2 + c^2 - 2(ab + ac + bc). \tag{B1}$$

[19]The expressions for $m_{\ell_1} \neq m_{\ell_2}$ can be applied to studies of lepton flavor-violating processes in all the decay modes considered in this paper within the lepton factorization approximation, and are also applicable to decays involving an $l\bar{\nu}$ in the final state e.g. $B \to D^*\ell\bar{\nu}$.

For a decay $A \to B + C$, in the rest-frame of A, it is related to the absolute value the spatial momentum of the B and C particles as

$$|\vec{p}_B| = |\vec{p}_C| = \frac{\sqrt{\lambda(m_A^2, m_B^2, m_C^2)}}{2m_A}. \tag{B2}$$

1. Basis-dependent kinematics for $\overline{B} \to \overline{K}^* \ell_1 \overline{\ell}_2$

We parametrize the kinematics of the ($\ell_1 = \ell^-$ and $\ell_2 = \ell^-$)

$$\overline{B} \to \overline{K}^*(\to \overline{K}(p_K)\pi(p_\pi))\ell_1(\ell_1)\overline{\ell}_2(\ell_2) \tag{B3}$$

decay mode. To do so we need all four momenta p_π, p_K ($p = p_\pi + p_K$), ℓ_1 and ℓ_2 ($q = \ell_1 + \ell_2$) in a specific frame for which we choose the $\overline{B}$-restframe. It is simplest to first obtain $\ell_{1,2}$ and $p_{\pi,K}$ in the rest frame of the lepton pair and the $\overline{K}^*$-meson respectively:

$$\begin{aligned} &\ell_{1,2}\text{-rest frame}: \ell_1^\mu = (E_1, |\vec{p}_\ell|\hat{\ell}), \quad \ell_2^\mu = (E_2, -|\vec{p}_\ell|\hat{\ell}), \\ &p_{\pi,K}\text{-rest frame}: p_K^\mu = (E_K, |\vec{p}_K|\hat{k}), \quad p_\pi^\mu = (E_\pi, -|\vec{p}_K|\hat{k}), \end{aligned} \tag{B4}$$

and the definitions

$$\begin{aligned} \hat{\ell} &= (\cos\phi\sin\theta_\ell, -\sin\phi\sin\theta_\ell, \cos\theta_\ell), \quad |\vec{p}_\ell| = \frac{\sqrt{\lambda_{\gamma^*}}}{2\sqrt{q^2}}, \\ \hat{k} &= (-\sin\theta_K, 0, -\cos\theta_K), \quad |\vec{p}_K| = \frac{\sqrt{\lambda_{K^*}}}{2m_{K^*}}, \end{aligned} \tag{B5}$$

where

$$\begin{aligned} \lambda_{\gamma^*} &\equiv \lambda(q^2, m_1^2, m_2^2), \\ \lambda_{K^*} &\equiv \lambda(m_{K^*}^2, m_K^2, m_\pi^2), \\ \lambda_B &\equiv \lambda(m_B^2, m_{K^*}^2, q^2), \end{aligned} \tag{B6}$$

are the explicit Källén function used throughout. The lepton and hadron energies are then given by $E_{1,2} = \sqrt{m_{\ell_{1,2}}^2 + |\vec{p}_\ell|^2}$, $E_{\pi,K} = \sqrt{m_{\pi,K}^2 + |\vec{p}_K|^2}$ and obey $E_1 + E_2 = \sqrt{q^2}$ and $E_\pi + E_K = m_{K^*}$.

The polarization vectors $\eta^\mu(\lambda)$ of the $\overline{K}^*$-meson in its rest frame, using the convention in [18], are[20]

$$\eta^\mu(0) = (0, 0, 0, 1), \qquad \eta^\mu(\pm) = (0, \mp 1, i, 0)/\sqrt{2}. \tag{B7}$$

In the $\overline{B}$-restframe, $p_B = (m_B, 0, 0, 0)$, the momenta take the following form

$$\begin{aligned} (\ell_1)^\mu &= (f_\ell(E_1, q_0, q_z), |\vec{p}_\ell|\sin\theta_\ell\cos\phi, -|\vec{p}_\ell|\sin\theta_\ell\sin\phi, f_\ell(E_1, q_z, q_0)), \\ (\ell_2)^\mu &= (f_\ell(E_2, q_0, -q_z), -|\vec{p}_\ell|\sin\theta_\ell\cos\phi, +|\vec{p}_\ell|\sin\theta_\ell\sin\phi, f_\ell(E_2, q_z, -q_0)), \\ (p_K)^\mu &= (f_{K^*}(E_K, p_0, q_z), -|\vec{p}_K|\sin\theta_K, 0, -f_{K^*}(E_K, q_z, p_0)), \\ (p_\pi)^\mu &= (f_{K^*}(E_\pi, p_0, -q_z), |\vec{p}_K|\sin\theta_K, 0, -f_{K^*}(E_\pi, q_z, -p_0)), \end{aligned} \tag{B8}$$

with $f_\ell(a,b,c) = (ab + c|\vec{p}_\ell|\cos\theta_\ell)/\sqrt{q^2}$ and $f_{K^*}(a,b,c) = (ab + c|\vec{p}_K|\cos\theta_K)/m_K^*$, and it is easily verified that

$$\begin{aligned} q^\mu &= (\ell_1 + \ell_2)^\mu = (q_0, 0, 0, q_z), \\ p^\mu &= (p_K + p_\pi)^\mu = (p_0, 0, 0, -q_z), \end{aligned} \tag{B9}$$

($p_0 = E_{K^*}$) while the polarization vectors of the $\overline{K}^*$ in the $\overline{B}$-rest frame are

$$\begin{aligned} \eta^\mu(0) &= (-q_z, 0, 0, p_0)/m_{K^*}, \\ \eta^\mu(\pm) &= (0, \mp 1, i, 0)/\sqrt{2}, \end{aligned} \tag{B10}$$

where $p_0 + q_0 = m_B$ and $q_z = \sqrt{\lambda_B}/(2m_B)$, in accordance with (B2), is the three-momentum of the lepton pair.

For completeness let us add that in the case of:

(i) $B \to K^* \overline{\ell}_1(\ell_1)\ell_2(\ell_2)$ the replacement rule $\hat{\ell} \to \hat{\ell}_{\phi\to-\phi} = (\cos\phi\sin\theta_\ell, +\sin\phi\sin\theta_\ell, \cos\theta_\ell)$ applies. Note this is coherent with Fig. 4 in the next section;

(ii) identical lepton masses the following replacements are in order

$$E_{1,2} \to \sqrt{q^2}/2, \qquad \sqrt{\lambda_{\gamma^*}} \to (q^2)\beta_\ell \tag{B11}$$

where we recall that $\beta_\ell \equiv \sqrt{1 - \frac{4m_\ell^2}{q^2}}$.

2. Basis-independent kinematics for $\overline{B} \to \overline{K}^* \ell_1 \overline{\ell}_2$

Introducing the notation

$$Q^\mu = (\ell_1 - \ell_2)^\mu, \qquad P^\mu = (p_K - p_\pi)^\mu, \tag{B12}$$

in addition to (B9). the invariants that can be formed out of p, P, q and Q are given by

[20]The polarisation vector η corresponds to γ in [18] (c.f. appendix A therein). The exact correspondence between the convention used in [18], and also in this paper, and the Jacob-Wick convention [1,3] is $\eta(\pm)_\mu|_{[18]} = -\eta(\pm)_\mu|_{[1]}$, $\eta(0)_\mu|_{[18]} = \eta(0)_\mu|_{[1]}$. The final distributions remain the same but the off-diagonal elements of the lepton HAs (or matrices) change sign (A14). Note in particular that the hadron HAs (C15) remain unchanged.

$$
\begin{aligned}
q\cdot Q &= m_{\ell_1}^2 - m_{\ell_2}^2,\\
Q^2 &= 2(m_{\ell_1}^2 + m_{\ell_2}^2) - q^2,\\
q\cdot p &= \frac{1}{2}(m_B^2 - m_{K^*}^2 - q^2),\\
p\cdot P &= m_K^2 - m_\pi^2,\\
P^2 &= 2(m_K^2 + m_\pi^2) - m_{K^*}^2,\\
q\cdot P &= \frac{2p\cdot Pq\cdot p + \cos\theta_K\sqrt{\lambda_B\lambda_{K^*}}}{2m_{K^*}^2},\\
Q\cdot P &= \frac{p\cdot P\sqrt{\lambda_B\lambda_{\gamma^*}}\cos\theta_\ell + 2q\cdot p\sqrt{\lambda_{K^*}\lambda_{\gamma^*}}\cos\theta_K\cos\theta_\ell}{2m_{K^*}^2q^2}\\
&\quad + \frac{\sqrt{\lambda_{K^*}\lambda_{\gamma^*}}\sin\theta_K\sin\theta_\ell\cos\phi}{m_{K^*}\sqrt{q^2}} + \frac{q\cdot Qq\cdot P}{q^2},\\
Q\cdot p &= \frac{2q\cdot Qq\cdot p + \cos\theta_\ell\sqrt{\lambda_B\lambda_{\gamma^*}}}{2q^2},\\
\epsilon(P,p,Q,q) &= -\frac{\sin\theta_K\sin\theta_\ell\sin\phi\sqrt{\lambda_B\lambda_{K^*}\lambda_{\gamma^*}}}{2m_{K^*}\sqrt{q^2}}
\end{aligned}
\tag{B13}
$$

with $p^2 = m_{K^*}^2$, the λ's defined in Eq. (B6), $\epsilon(P,p,Q,q) = \epsilon_{\alpha\beta\gamma\delta}P^\alpha p^\beta Q^\gamma q^\delta$ and the $\epsilon_{0123} = 1$ convention for the Levi-Civita tensor. Note that the kinematic invariants for $B\to K^*\overline{\ell}_1(\ell_1)\ell_2(\ell_2)$ are the same up to $\epsilon(P,p,Q,q)\to -\epsilon(P,p,Q,q)$ which originates from the only change in angles $\phi\to-\phi$.

APPENDIX C: SPECIFIC RESULTS FOR $\overline{B}\to\overline{K}^*(\to\overline{K}\pi)\ell_1\overline{\ell}_2$

1. Fourfold differential decay rate

The angular distribution for $\overline{B}\to\overline{K}^*(\to\overline{K}\pi)\ell_1\overline{\ell}_2$ is usually presented in the form (e.g. [36])

$$
\begin{aligned}
&\frac{8\pi}{3}\frac{d^4\Gamma}{dq^2d\cos\theta_\ell d\cos\theta_K d\phi}\\
&\quad = \frac{I_{K^*}^{(0)}}{4} = (g_{1s} + g_{2s}\cos2\theta_\ell + g_{6s}\cos\theta_\ell)\sin^2\theta_K\\
&\qquad + (g_{1c} + g_{2c}\cos2\theta_\ell + g_{6c}\cos\theta_\ell)\cos^2\theta_K\\
&\qquad + (g_3\cos2\phi + g_9\sin2\phi)\sin^2\theta_K\sin^2\theta_\ell\\
&\qquad + (g_4\cos\phi + g_8\sin\phi)\sin2\theta_K\sin2\theta_\ell\\
&\qquad + (g_5\cos\phi + g_7\sin\phi)\sin2\theta_K\sin\theta_\ell,
\end{aligned}
\tag{C1}
$$

which can be condensed as

$$
\begin{aligned}
&\frac{8\pi}{3}\frac{d^4\Gamma}{dq^2d\cos\theta_\ell d\cos\theta_K d\phi}\\
&\quad = \mathrm{Re}[(g_{1s} + g_{2s}\cos2\theta_\ell + g_{6s}\cos\theta_\ell)\sin^2\theta_K\\
&\qquad + (g_{1c} + g_{2c}\cos2\theta_\ell + g_{6c}\cos\theta_\ell)\cos^2\theta_K\\
&\qquad + e^{-2i\phi}\mathcal{G}_3\sin^2\theta_K\sin^2\theta_\ell\\
&\qquad + e^{-i\phi}\sin2\theta_K(\mathcal{G}_4\sin2\theta_\ell + \mathcal{G}_5\sin\theta_\ell)],
\end{aligned}
\tag{C2}
$$

where we have defined

$$
\mathcal{G}_{3,4,5} = (g_{3,4,5} + ig_{9,8,7}). \tag{C3}
$$

We have introduced the notation g_i rather than J_i in order to minimize the potential of confusion due to the angular conventions discussed in Appendix C 2. The relationship between the $g_i(q^2)$ and the $G_m^{l_K,l_\ell}(q^2)$ was given in (32) but is repeated here for convenience:

$$
\begin{aligned}
G_0^{0,0} &= \frac{4}{9}(3(g_{1c} + 2g_{1s}) - (g_{2c} + 2g_{2s})),\\
G_0^{0,1} &= \frac{4}{3}(g_{6c} + 2g_{6s}), \qquad G_0^{0,2} = \frac{16}{9}(g_{2c} + 2g_{2s}),\\
G_0^{2,0} &= \frac{4}{9}(6(g_{1c} - g_{1s}) - 2(g_{2c} - g_{2s})),\\
G_0^{2,1} &= \frac{8}{3}(g_{6c} - g_{6s}), \qquad G_0^{2,2} = \frac{32}{9}(g_{2c} - g_{2s}),\\
G_1^{2,1} &= \frac{16}{\sqrt{3}}\underbrace{(g_5 + ig_7)}_{=\mathcal{G}_5}, \qquad G_1^{2,2} = \frac{32}{3}\underbrace{(g_4 + ig_8)}_{=\mathcal{G}_4},\\
G_2^{2,2} &= \frac{32}{3}\underbrace{(g_3 + ig_9)}_{=\mathcal{G}_3}.
\end{aligned}
\tag{C4}
$$

Explicit results for the $G_m^{l_K,l_\ell}$ are presented in Sec. C 3 for the case of identical final-state leptons $m_{\ell_1} = m_{\ell_2}$ and Sec. C 4 for the more general case $m_{\ell_1}\neq m_{\ell_2}$.

a. Kinematic endpoint relations in terms of $G_m^{l_K,l_\ell}$

In Ref. [18] it was shown that the HAs obey symmetry relations at the kinematic endpoint due to symmetry enhancement. This is due to the K^* being at rest resulting in symmetry in all space directions i.e. helicity directions. The relations for the HAs in Eq. (13) in [18] lead to the following equivalent of Eq. (21) in [18]

$$
\begin{aligned}
&G_0^{0,0}\neq 0, \qquad G_0^{2,2}\to\mathrm{Re}[G_0^{2,2}],\\
&G_1^{2,2}\to -2\mathrm{Re}[G_0^{2,2}], \qquad G_2^{2,2}\to 2\mathrm{Re}[G_0^{2,2}],
\end{aligned}
\tag{C5}
$$

with all other five $G_m^{l_K,l_\ell}$ vanishing. Recall that $G_0^{0,0}$ is proportional to the total decay rate. The relations between the $G_m^{2,2}$ are not accidental but have to do

with the symmetries of a multiplet. The factor of two between $G_0^{2,2}$ and $G_{1(2)}^{2,2}$, once more, originates from absorbing $G_{-1(2)}^{2,2}$ into $G_{1(2)}^{2,2}$. The results of the threshold expansion, linear in the K^* momentum $\kappa \sim \lambda(m_B^2, m_{K^*}^2, q^2)$, can be inferred from Eq. (30) in [18] taking into account the different angular conventions detailed in Fig. 4.

b. Comparison of angular distribution with the literature

The angular distribution (C1) was first computed in the SM for massless leptons in [22], extended to include equal lepton masses in [23,72]. A full dimension-six operator basis was considered in [73]. The basis was extended to lepton mass corrections for (pseudo)scalar operators in [35], enforcing the g_{6c}-structure, and tensor operators by the authors in [36,74]. We compare our results with regard to [36], which is the latest reference.

Taking into account the change $g_{4,6,7,9} \to -J_{4,6,7,9}$ (cf. Fig. 4) and comparing at the level of form factors (naive factorization) only we find agreement except for tensor interference terms. Agreement is established when $C_{T5} \to -C_{T5}$ in [36]. The latter might be related to the fact that the relations $\mathrm{tr}[\gamma^\alpha\gamma^\beta\gamma^\gamma\gamma^\delta\gamma_5] = 4\lambda i\epsilon^{\alpha\beta\gamma\delta}$ and $\sigma^{\alpha\beta}\gamma_5 = -\lambda\frac{i}{2}\epsilon^{\alpha\beta\gamma\delta}\sigma_{\gamma\delta}$ (with $\lambda = \pm 1$ depending on conventions —$\lambda = 1$ in this paper) are not consistent with Eq. C.16 [36] (v3).

A minor difference is that the authors of [36] have chosen not to present the tensor contribution in $J_{8,9}(g_{8,9})$, since such contributions vanish in the narrow-width approximation.[21] In addition, we find that a few of the HAs in [36] Eq. (C13) do not agree with their definitions. These disagreements do, however, drop out in the final expression.

2. Angular conventions

In this section we discuss and compare the LHCb and theory angular conventions. The main result is shown in form of a commutative diagram in Fig. 4. We proceed by first discussing the CP-conjugate modes in each case and then link the conventions with each other.

The LHCb conventions [32], which are the same as adapted in this paper, are shown in Fig. 1. The rationale behind the definition of the conjugate mode is as follows. First, particles are mapped into antiparticles, corresponding to a C-transformation. Then the momenta of all particles are reversed, changing the angle $\phi \to 2\pi - \phi$. This leads to sign changes in $g_{7,8,9}$. Hence the conjugate mode corresponds to a full CP-transformation

$$\left.\frac{d^4\overline{\Gamma}}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi}\right|_{\mathrm{LHCb}} = \left.\frac{d^4\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi}\right|_{CP},$$

and the quantity

$$\left.\frac{d^4(\Gamma \pm \overline{\Gamma})}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi}\right|_{\mathrm{LHCb}},$$

is therefore CP-even (-odd). Above $\Gamma = \Gamma(\overline{B} \to \overline{K}^*\ell_1\overline{\ell}_2)$ and $\overline{\Gamma} = \Gamma(B \to K^*\overline{\ell}_1\ell_2)$.

The theory conventions for CP conjugates are such that they facilitate the implementation of decays which are not self-tagging [such as $\overline{B}_s, B_s \to \phi(\to K^+K^-)\ell^+\ell^-$ at hadron colliders]. When going between conjugate modes the conventions are that the angles transform as $(\theta_\ell, \theta_K, \phi) \to (\pi - \theta_\ell, \pi - \theta_K, 2\pi - \phi)$,[22] which leads to sign changes in $g_{5,6,8,9}$. This transformation rule corresponds to a full CP-conjugation, but with the angles θ_ℓ, θ_K associated to the same particle rather than the antiparticle.

To find the transformation between the theory and LHCb conventions is not straightforward because it is difficult to find a theory paper that resolves the fourfold ambiguity of defining the angle ϕ and/or shows a figure consistent with the definitions used in the corresponding work. We have taken a pragmatic route in verifying that the results in [35,36,72] agree with each other for common contributions, and crucially that our results are in agreement with these contributions for $\overline{B} \to \overline{K}^*\ell_1\overline{\ell}_2$ if $J_{4,6,7,9} = -g_{4,6,7,9}$ and $J_{1,2,3,5,8} = g_{1,2,3,5,8}$. This completes the diagram in Fig. 4.[23]

[21] In v3 of [36] it is stated that agreement with v4 of [75] is found up to a sign of an interference term between a scalar and a tensor HA. This suggests that we agree with [36] but disagree with [75] on that sign, as well as the sign of C_{T5}.

[22] Equivalently one can use the angular redefinitions $(\pi - \theta_\ell, \theta_K, \pi - \phi)$ and $(\theta_\ell - \pi, \theta_K, 2\pi - \phi)$, which are sometimes stated in the literature.

[23] One can come to the same conclusion in another way [76]. Let us again consider $\bar{B} \to \bar{K}^*\ell_1\bar{\ell}_2$. In general $\theta_K|_{\mathrm{LHCb}} = \theta_K|_{\mathrm{theory}}$ is chosen to be the same angle and the theoretical community chooses $\theta_\ell|_{\mathrm{LHCb}} = \pi - \theta_\ell|_{\mathrm{theory}}$. The only unknown remains the angle ϕ, for which one may use the scalar- and cross-product definitions. Using Appendix A in [32] and likewise in [72], we infer that $\cos\phi|_{\mathrm{LHCb}} = \cos\phi|_{\mathrm{theory}}$ and $\sin\phi|_{\mathrm{LHCb}} = -\sin\phi|_{\mathrm{theory}}$. Taking all angular changes into account this results in sign changes in $g_{4,6,7,9}$ which is consistent with our explicit computations mentioned above.

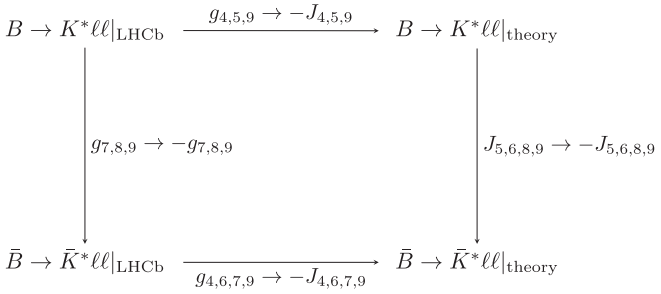

FIG. 4. Changes of angular functions when going from one mode to the other. For CP conjugates the conjugation of the CP-odd (weak) phases are suppressed. Angular functions whose signs do not change are not indicated.

In summary,

$$\left.\frac{d^4(\Gamma \pm \overline{\Gamma})}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi}\right|_{\rm LHCb} \Leftrightarrow S[A]_{1,2,3,4,5,6,7,8,9},$$
$$\left.\frac{d^4(\Gamma \pm \overline{\Gamma})}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi}\right|_{\rm theory} \Leftrightarrow S[A]_{1,2,3,4,7}, A[S]_{5,6,8,9}, \tag{C6}$$

where the CP-even (-odd) quantities are

$$S_i[A_i] = \frac{g_i \pm g_i^{CP}}{\Gamma + \Gamma^{CP}} \tag{C7}$$

with adapted notation from [35]. Written in yet another way (C6) is equivalent to

$$(g, A, S)_{1,2,3,5,8}|_{\rm LHCb} = +(J, A, S)_{1,2,3,5,8}|_{\rm theory},$$
$$(g, A, S)_{4,6,7,9}|_{\rm LHCb} = -(J, A, S)_{4,6,7,9}|_{\rm theory}. \tag{C8}$$

In order to understand (C6) and (C8) one has to keep in mind that $\overline{B} \to \overline{K}^*(\ell_1\overline{\ell}_2)$ rather than its conjugate is the reference decay. Note that at the LHCb (hadron collider) $B_s \to \phi\mu^+\mu^-$ is untagged and therefore, setting aside the issue of production asymmetry, only $S_{1,2,3,4,7}, A_{5,6,8,9}$ are experimentally accessible.

a. Angular observables in the literature and conventions

We aim to find the relation between angular $P_i^{(\prime)}$ observables as defined by the theorists [34] and their adaptation by LHCb [26]. In matching the results and creating the dictionary one needs to pay attention to the fact that [26] and [34] define the $P_i{}'$ in terms of g_i and J_i differently, as well as the different angular conventions for g_i and J_i per se (shown in Fig. 4).

Amongst the twelve observables discussed in Sec. III B, eight of them, $P_{1,2,3}$, $P'_{4,5,6,8}$ and $A_{\rm FB}$, depend on angles and definitions.

The $P_i{}'$ and $A_{\rm FB}$ are defined by LHCb [26] as

$$P'_{4,5,6,8}|_{\rm LHCb} = \frac{S_{4,5,7,8}|_{\rm LHCb}}{\sqrt{F_L(1-F_L)}},$$
$$A_{\rm FB}|_{\rm LHCb} = \frac{3(S_{6s}|_{\rm LHCb})}{4(\Gamma + \overline{\Gamma})}, \tag{C9}$$

where S_i is defined in Eq. (C7) and $2\mathcal{N}'_{\rm bin} = \sqrt{F_L(1-F_L)}$ in our notation used in Sec. III B. LHCb has not defined $P_{1,2,3}$ and we shall assume the same functional form as in the theory paper [34].

In [34] the eight equivalent angular observables are defined as follows[24]

$$P_1 = \frac{1}{2S_{2s}} S_3 = \frac{1}{2S_{2s}}(S_3|_{\rm LHCb}) = +P_1|_{\rm LHCb},$$
$$P_2 = \frac{1}{8S_{2s}} S_{6s} = \frac{1}{8S_{2s}}(-S_{6s}|_{\rm LHCb}) = -P_2|_{\rm LHCb},$$
$$P_3 = \frac{-1}{4S_{2s}} S_9 = \frac{-1}{4S_{2s}}(-S_9|_{\rm LHCb}) = -P_3|_{\rm LHCb},$$
$$P'_4 = \frac{1}{\mathcal{N}'_{\rm bin}} S_4 = \frac{1}{\mathcal{N}'_{\rm bin}}(-S_4|_{\rm LHCb}) = -2P'_4|_{\rm LHCb},$$
$$P'_5 = \frac{1}{2\mathcal{N}'_{\rm bin}} S_5 = \frac{1}{2\mathcal{N}'_{\rm bin}}(S_5|_{\rm LHCb}) = +P'_5|_{\rm LHCb},$$
$$A_{\rm FB} = -\frac{3S_{6s}}{4(\Gamma + \overline{\Gamma})} = -\frac{3(-S_{6s}|_{\rm LHCb})}{4(\Gamma + \overline{\Gamma})} = +A_{\rm FB}|_{\rm LHCb},$$
$$P'_6 = \frac{-1}{2\mathcal{N}'_{\rm bin}} S_7 = \frac{-1}{2\mathcal{N}'_{\rm bin}}(-S_7|_{\rm LHCb}) = +P'_6|_{\rm LHCb},$$
$$P'_8 = \frac{-1}{\mathcal{N}'_{\rm bin}} S_8 = \frac{-1}{\mathcal{N}'_{\rm bin}}(S_8|_{\rm LHCb}) = -2P'_8|_{\rm LHCb},$$

where we have directly translated into the LHCb conventions. It seems that we differ from the theory community in the sign of the observables $S[A]_{7,8,9}$. For example, both $P_6{}' = P_6{}'|_{\rm LHCb}$ and $P_8{}' = -2P_8{}'|_{\rm LHCb}$ differ from the relations given in the caption of Table 1 in [65] by the aforementioned sign. Our relation $S[A]_9 = -S[A]_9|_{\rm LHCb}$ also differs from the one given by [66] in Table 1 by a sign.

3. $G_m^{l_K,l_\ell}$ for $\overline{B} \to \overline{K}^*\ell_1\overline{\ell}_2$ in terms of helicity amplitudes for $m_{\ell_i} \equiv m_\ell$

When the masses of the leptons are identical, we obtain for $G_m^{l_K,l_\ell} = \mathcal{N}q^2\mathbb{G}_m^{l_K,l_\ell}$ [with $\mathcal{N}$ defined in (26)],

[24]Note that [35,36] define $A_{\rm FB} = \frac{3S_{6s}}{4(\Gamma+\overline{\Gamma})}$ which results in $A_{\rm FB}|_{[34,35]} = -A_{\rm FB}|_{\rm LHCb}$.

$$
\begin{aligned}
\mathbb{G}_0^{0,0} &= \frac{4}{9}(1-\hat{m}_\ell{}^2)(|H_+^V|^2+|H_-^V|^2+|H_0^V|^2+(V\to A))+\frac{4\hat{m}_\ell{}^2}{3}(|H_+^V|^2+|H_-^V|^2+|H_0^V|^2-(V\to A))+\frac{2}{3}\beta_\ell^2|H^S|^2+\frac{2}{3}|H^P|^2\\
&\quad+\frac{8}{9}(1+8\hat{m}_\ell{}^2)(|H_+^{T_t}|^2+|H_-^{T_t}|^2+|H_0^{T_t}|^2)+\frac{4}{9}\beta_\ell^2(|H_+^T|^2+|H_-^T|^2+|H_0^T|^2)+\frac{16}{3}\hat{m}_\ell\mathrm{Im}[H_+^V\overline{H}_+^{T_t}+H_-^V\overline{H}_-^{T_t}+H_0^V\overline{H}_0^{T_t}],\\
\mathbb{G}_0^{0,1} &= \frac{4\beta_\ell}{3}(\mathrm{Re}[H_+^V\overline{H}_+^A-H_-^V\overline{H}_-^A]+\mathrm{Im}[\sqrt{2}H_0^T\overline{H}^P+2H_0^{T_t}\overline{H}^S]-2\hat{m}_\ell\mathrm{Re}[H_0^V\overline{H}^S]+4\hat{m}_\ell\mathrm{Im}[H_+^A\overline{H}_+^{T_t}-H_-^A\overline{H}_-^{T_t}]),\\
\mathbb{G}_0^{0,2} &= -\frac{2}{9}\beta_\ell^2(2|H_0^V|^2-|H_+^V|^2-|H_-^V|^2+(V\to A)-2(2|H_0^T|^2-|H_+^T|^2-|H_-^T|^2)-4(2|H_0^{T_t}|^2-|H_+^{T_t}|^2-|H_-^{T_t}|^2)),\\
\mathbb{G}_0^{2,0} &= -\frac{4}{9}(1-\hat{m}_\ell{}^2)(|H_+^V|^2+|H_-^V|^2-2|H_0^V|^2+(V\to A))-\frac{4\hat{m}_\ell{}^2}{3}(|H_+^V|^2+|H_-^V|^2-2|H_0^V|^2-(V\to A))\\
&\quad+\frac{4}{3}\beta_\ell^2|H^S|^2+\frac{4}{3}|H^P|^2-\frac{8}{9}(1+8\hat{m}_\ell{}^2)(|H_+^{T_t}|^2+|H_-^{T_t}|^2-2|H_0^{T_t}|^2)-\frac{4}{9}\beta_\ell^2(|H_+^T|^2+|H_-^T|^2-2|H_0^T|^2)\\
&\quad-\frac{16}{3}\hat{m}_\ell\mathrm{Im}[H_+^V\overline{H}_+^{T_t}+H_-^V\overline{H}_-^{T_t}-2H_0^V\overline{H}_0^{T_t}],\\
\mathbb{G}_0^{2,1} &= -\frac{4\beta_\ell}{3}(\mathrm{Re}[H_+^V\overline{H}_+^A-H_-^V\overline{H}_-^A]-2\mathrm{Im}[\sqrt{2}H_0^T\overline{H}^P+2H_0^{T_t}\overline{H}^S]+4\hat{m}_\ell(\mathrm{Re}[H_0^V\overline{H}^S]+\mathrm{Im}[H_+^A\overline{H}_+^{T_t}-H_-^A\overline{H}_-^{T_t}])),\\
\mathbb{G}_0^{2,2} &= -\frac{2}{9}\beta_\ell^2(4|H_0^V|^2+|H_+^V|^2+|H_-^V|^2+(V\to A)-2(4|H_0^T|^2+|H_+^T|^2+|H_-^T|^2)-4(4|H_0^{T_t}|^2+|H_+^{T_t}|^2+|H_-^{T_t}|^2)),\\
\mathbb{G}_1^{2,1} &= \frac{4\beta_\ell}{\sqrt{3}}(H_+^VH_0^A+H_+^AH_0^V-H_0^V\overline{H}_-^A-H_0^A\overline{H}_-^V+2\hat{m}_\ell(H_+^V\overline{H}^S+H^S\overline{H}_-^V)\\
&\quad-\sqrt{2}i(H^P\overline{H}_-^T-H_+^T\overline{H}^P+\sqrt{2}(H^S\overline{H}_-^{T_t}-H_+^{T_t}\overline{H}^S))-4i\hat{m}_\ell(H_+^A\overline{H}_0^{T_t}+H_0^{T_t}\overline{H}_-^A-H_+^{T_t}\overline{H}_0^A-H_0^A\overline{H}_-^{T_t})),\\
\mathbb{G}_1^{2,2} &= \frac{4}{3}\beta_\ell^2(H_+^VH_0^V+H_0^V\overline{H}_-^V+(V\to A)-2(H_+^T\overline{H}_0^T+H_0^T\overline{H}_-^T+2(H_+^{T_t}\overline{H}_0^{T_t}+H_0^{T_t}\overline{H}_-^{T_t}))),\\
\mathbb{G}_2^{2,2} &= -\frac{8}{3}\beta_\ell^2(H_+^V\overline{H}_-^V+H_+^A\overline{H}_-^A-2(H_+^T\overline{H}_-^T+2H_+^{T_t}\overline{H}_-^{T_t})),
\end{aligned}
\tag{C10}
$$

where $\hat{m}_\ell = m_\ell/\sqrt{q^2}$ and we recall that $\beta_\ell \equiv \sqrt{1-4\hat{m}_\ell^2}$. The number m in $G_m^{l_K,l_\ell}$ corresponds to the units of plus helicities. The common factor of q^2 in all observables as compared with standard literature results is a consequence of our choice of normalisation, whereby all global factors are placed outside the HAs. The factors of i where they appear (explicitly and implicitly) in $\mathbb{G}_1^{2,1}$, $\mathbb{G}_1^{2,2}$ and $\mathbb{G}_2^{2,2}$ are not accidental, as the results given above are complex and one must take the real and imaginary parts of these results to recover the observables $g_{3,4,5,7,8,9}$.

Note that it is sometimes convenient to express results in terms of the transversity amplitudes, which possess a definite parity. The relations to the HAs used throughout this paper are

$$
\begin{aligned}
&H_{\parallel(\perp)}^{L/R} \equiv \frac{1}{\sqrt{2}}(H_+^{L/R} \pm H_-^{L/R}), \qquad H_S \equiv H^S, \qquad H_t \equiv H^t, \qquad H_{\parallel(\perp)}^{\mathcal{T}_t} \equiv \frac{1}{\sqrt{2}}(H_+^{T_t} \pm H_-^{T_t}),\\
&H_0^{\mathcal{T}_t} \equiv H_0^{T_t}, \qquad H_{\parallel(\perp)}^{\mathcal{T}} \equiv \frac{1}{\sqrt{2}}(H_+^T \pm H_-^T), \qquad H_{\parallel\perp}^{\mathcal{T}} \equiv H_0^T.
\end{aligned}
\tag{C11}
$$

In [36] the notation A_{ij}, with $i,j = \parallel,\perp,0$, is used for the transversity amplitudes. Note that when comparing to this paper the difference in the convention of the polarization vectors has to be taken into account.

4. $G_m^{l_K,l_\ell}$ for $\overline{B} \to \overline{K}^*\ell_1\overline{\ell}_2$ in terms of helicity amplitudes for $m_{\ell_1} \neq m_{\ell_2}$

The formalism discussed in this paper allows a simple extension to the case $m_{\ell_1} \neq m_{\ell_2}$, so that the results presented in (C10) can be adapted to test for possible lepton-flavor violating processes. Using the notation

$$
\frac{\lambda_{\gamma^*}}{4q^2} = \beta_1^+\beta_1^-\beta_2^+\beta_2^-,
\tag{C12}
$$

[λ_{γ^*} given in (B6) and $\beta_{1,2}^\pm \equiv \sqrt{E_{1,2} \pm m_{\ell_{1,2}}}$], we obtain the following expressions for $G_m^{l_K,l_\ell} = \mathcal{N}\tilde{G}_m^{l_K,l_\ell}$ [with $\mathcal{N}$ defined in (26)]:

$$
\begin{aligned}
\tilde{G}_0^{0,0} =& \frac{4}{9}\left(3E_1E_2+\frac{\lambda_{\gamma^*}}{4q^2}\right)(|H_+^V|^2+|H_-^V|^2+|H_0^V|^2+(V\to A))+\frac{4m_{\ell_1}m_{\ell_2}}{3}(|H_+^V|^2+|H_-^V|^2+|H_0^V|^2-(V\to A))\\
&+\frac{4}{3}\left(E_1E_2-m_{\ell_1}m_{\ell_2}+\frac{\lambda_{\gamma^*}}{4q^2}\right)|H^S|^2+\frac{4}{3}\left(E_1E_2+m_{\ell_1}m_{\ell_2}+\frac{\lambda_{\gamma^*}}{4q^2}\right)|H^P|^2\\
&+\frac{16}{9}\left(3(E_1E_2+m_{\ell_1}m_{\ell_2})-\frac{\lambda_{\gamma^*}}{4q^2}\right)(|H_+^{T_t}|^2+|H_-^{T_t}|^2+|H_0^{T_t}|^2)+\frac{8}{9}\left(3(E_1E_2-m_{\ell_1}m_{\ell_2})-\frac{\lambda_{\gamma^*}}{4q^2}\right)\\
&\times(|H_+^T|^2+|H_-^T|^2+|H_0^T|^2)+\frac{16}{3}(m_{\ell_1}E_2+m_{\ell_2}E_1)\mathrm{Im}[H_+^V\overline{H}_+^{T_t}+H_-^V\overline{H}_-^{T_t}+H_0^V\overline{H}_0^{T_t}]\\
&+\frac{8\sqrt{2}}{3}(m_{\ell_1}E_2-m_{\ell_2}E_1)\mathrm{Im}[H_+^A\overline{H}_+^T+H_-^A\overline{H}_-^T+H_0^A\overline{H}_0^T],\\
\tilde{G}_0^{0,1} =& \frac{4\sqrt{\lambda_{\gamma^*}}}{3}\left(\mathrm{Re}[H_+^V\overline{H}_+^A-H_-^V\overline{H}_-^A]+2\sqrt{2}\frac{m_{\ell_1}^2-m_{\ell_2}^2}{q^2}\mathrm{Re}[H_+^T\overline{H}_+^{T_t}-H_-^T\overline{H}_-^{T_t}]+2\frac{m_{\ell_1}+m_{\ell_2}}{\sqrt{q^2}}\mathrm{Im}[H_+^A\overline{H}_+^{T_t}-H_-^A\overline{H}_-^{T_t}]\right)\\
&+\sqrt{2}\frac{m_{\ell_1}-m_{\ell_2}}{\sqrt{q^2}}\mathrm{Im}[H_+^V\overline{H}_+^T-H_-^V\overline{H}_-^T]-\frac{m_{\ell_1}-m_{\ell_2}}{\sqrt{q^2}}\mathrm{Re}[H_0^A\overline{H}^P]-\frac{m_{\ell_1}+m_{\ell_2}}{\sqrt{q^2}}\mathrm{Re}[H_0^V\overline{H}^S]+\mathrm{Im}[\sqrt{2}H_0^T\overline{H}^P+2H_0^{T_t}\overline{H}^S]\Big),\\
\tilde{G}_0^{0,2} =& -\frac{2\lambda_{\gamma^*}}{9q^2}(2|H_0^V|^2-|H_+^V|^2-|H_-^V|^2+(V\to A)-2(2|H_0^T|^2-|H_+^T|^2-|H_-^T|^2)-4(2|H_0^{T_t}|^2-|H_+^{T_t}|^2-|H_-^{T_t}|^2)),\\
\tilde{G}_0^{2,0} =& -\frac{4}{9}\left(3E_1E_2+\frac{\lambda_{\gamma^*}}{4q^2}\right)(|H_+^V|^2+|H_-^V|^2-2|H_0^V|^2+(V\to A))-\frac{4m_{\ell_1}m_{\ell_2}}{3}(|H_+^V|^2+|H_-^V|^2-2|H_0^V|^2-(V\to A))\\
&+\frac{8}{3}\left(E_1E_2-m_{\ell_1}m_{\ell_2}+\frac{\lambda_{\gamma^*}}{4q^2}\right)|H^S|^2+\frac{8}{3}\left(E_1E_2+m_{\ell_1}m_{\ell_2}+\frac{\lambda_{\gamma^*}}{4q^2}\right)|H^P|^2\\
&-\frac{16}{9}\left(3(E_1E_2+m_{\ell_1}m_{\ell_2})-\frac{\lambda_{\gamma^*}}{4q^2}\right)(|H_+^{T_t}|^2+|H_-^{T_t}|^2-2|H_0^{T_t}|^2)-\frac{8}{9}\left(3(E_1E_2-m_{\ell_1}m_{\ell_2})-\frac{\lambda_{\gamma^*}}{4q^2}\right)(|H_+^T|^2+|H_-^T|^2-2|H_0^T|^2)\\
&-\frac{16}{3}(m_{\ell_1}E_2+m_{\ell_2}E_1)\mathrm{Im}[H_+^V\overline{H}_+^{T_t}+H_-^V\overline{H}_-^{T_t}-2H_0^V\overline{H}_0^{T_t}]-\frac{8\sqrt{2}}{3}(m_{\ell_1}E_2-m_{\ell_2}E_1)\mathrm{Im}[H_+^A\overline{H}_+^T+H_-^A\overline{H}_-^T-2H_0^A\overline{H}_0^T],\\
\tilde{G}_0^{2,1} =& -\frac{4\sqrt{\lambda_{\gamma^*}}}{3}\left(\mathrm{Re}[H_+^V\overline{H}_+^A-H_-^V\overline{H}_-^A]+2\sqrt{2}\frac{m_{\ell_1}^2-m_{\ell_2}^2}{q^2}\mathrm{Re}[H_+^T\overline{H}_+^{T_t}-H_-^T\overline{H}_-^{T_t}]+\frac{2(m_{\ell_1}+m_{\ell_2})}{\sqrt{q^2}}\mathrm{Im}[H_+^A\overline{H}_+^{T_t}-H_-^A\overline{H}_-^{T_t}]\right.\\
&\left.+\frac{\sqrt{2}(m_{\ell_1}-m_{\ell_2})}{\sqrt{q^2}}\mathrm{Im}[H_+^V\overline{H}_+^T-H_-^V\overline{H}_-^T]+2\frac{m_{\ell_1}-m_{\ell_2}}{\sqrt{q^2}}\mathrm{Re}[H_0^A\overline{H}^P]+2\frac{m_{\ell_1}+m_{\ell_2}}{\sqrt{q^2}}\mathrm{Re}[H_0^V\overline{H}^S]-2\mathrm{Im}[\sqrt{2}H_0^T\overline{H}^P+2H_0^{T_t}\overline{H}^S]\right),\\
\tilde{G}_0^{2,2} =& -\frac{2\lambda_{\gamma^*}}{9q^2}(4|H_0^V|^2+|H_+^V|^2+|H_-^V|^2+(V\to A)-2(4|H_0^T|^2+|H_+^T|^2+|H_-^T|^2)-4(4|H_0^{T_t}|^2+|H_+^{T_t}|^2+|H_-^{T_t}|^2)),\\
\tilde{G}_1^{2,1} =& \frac{4\sqrt{\lambda_{\gamma^*}}}{\sqrt{3}}\left((H_+^V\overline{H}_0^A+H_+^A\overline{H}_0^V-H_0^V\overline{H}_-^A-H_0^A\overline{H}_-^V)+\frac{m_{\ell_1}+m_{\ell_2}}{\sqrt{q^2}}(H_+^V\overline{H}^S+H^S\overline{H}_-^V)\right.\\
&-\sqrt{2}i(H^P\overline{H}_-^T-H_+^T\overline{H}^P+\sqrt{2}(H^S\overline{H}_-^{T_t}-H_+^{T_t}\overline{H}^S))\\
&+\frac{m_{\ell_1}-m_{\ell_2}}{\sqrt{q^2}}(H_+^A\overline{H}^P+H^P\overline{H}_-^A)-2i\frac{m_{\ell_1}+m_{\ell_2}}{\sqrt{q^2}}(H_+^A\overline{H}_0^{T_t}+H_0^{T_t}\overline{H}_-^A-H_+^{T_t}\overline{H}_0^A-H_0^A\overline{H}_-^{T_t})\\
&\left.-\sqrt{2}i\frac{m_{\ell_1}-m_{\ell_2}}{\sqrt{q^2}}(H_+^V\overline{H}_0^T+H_0^T\overline{H}_-^V-H_+^T\overline{H}_0^V-H_0^V\overline{H}_-^T)+2\sqrt{2}\frac{m_{\ell_1}^2-m_{\ell_2}^2}{q^2}(H_+^T\overline{H}_0^{T_t}+H_+^{T_t}\overline{H}_0^T-H_0^T\overline{H}_-^{T_t}-H_0^{T_t}\overline{H}_-^T)\right),\\
\tilde{G}_1^{2,2} =& \frac{4\lambda_{\gamma^*}}{3q^2}(H_+^V\overline{H}_0^V+H_0^V\overline{H}_-^V+(V\to A)-2(H_+^T\overline{H}_0^T+H_0^T\overline{H}_-^T+2(H_+^{T_t}\overline{H}_0^{T_t}+H_0^{T_t}\overline{H}_-^{T_t}))),\\
\tilde{G}_2^{2,2} =& -\frac{8\lambda_{\gamma^*}}{3q^2}(H_+^V\overline{H}_-^V+H_+^A\overline{H}_-^A-2(H_+^T\overline{H}_-^T+2H_+^{T_t}\overline{H}_-^{T_t})).
\end{aligned}
\tag{C13}
$$

5. Explicit helicity amplitudes in terms of form factors

We collect here the definitions of the helicity amplitudes in terms of which our results are expressed. The hadronic HA is defined by

$$H_\lambda^X = \langle \overline{K}^*(\lambda)|\overline{s}\overline{\Gamma}^X b|\overline{B}\rangle, \tag{C14}$$

with $\overline{\Gamma}^X|_{\lambda_X\to\lambda}$ as defined in Table 3.1 and the further replacement $\omega \to \overline{\omega}$. The definitions of the hadronic matrix elements used in the calculations are standard (e.g. [77]). Below we evaluate the HAs using form factors to make clear the relative signs between the various contributions, allowing for definite comparison with the literature.

Results for form factors for low q^2 can be found from light-cone sum rules (LCSR) with vector distribution amplitudes (DA) in [5,77] and B-meson DA in [6], and for high q^2 from lattice QCD [78]. Long-distance effects contribute to H_λ^V only, and include quark loops (QL), the chromomagnetic operator O_8, quark loop spectator scattering (QLSS) and weak annihilation (WA). At low q^2, effects have been evaluated in QCD factorization (QCDF) in the leading $1/m_b$-limit and in LCSR. Results for O_8, WA and QLSS in QCDF are given in [7], and additional contributions for O_8 in [8]. In Ref. [7] it was shown that quark loops can be integrated into the $1/m_b$ framework using the results from inclusive matrix element computations [9]. Results for O_8 and WA, as well as a prescription for dealing with endpoint divergences of QLSS, can be found in [10] and [11]. Results for charm loops beyond the $1/m_b$ approximation can be found in [12] for LCSR with B-meson DA, and [13,14] for LCSR (at $q^2 = 0$ only) for vector-meson DA. At high q^2 many of the long-distance contributions are suppressed in the formulation in terms of an OPE in $1/q^2$ (with $q^2 \simeq m_b^2$) [79,80]. It should be added that the large contribution of broad charm resonances in $B \to K\mu^+\mu^-$ observed by the LHCb collaboration [81] demands a reassessment of duality violations [67]. Long-distance contributions can be found elsewhere.

Explicit results for the $\overline{B} \to \overline{K}^*(\ell_1\overline{\ell}_2)$-mode are given by

$$\begin{aligned}
H_0^V &= \frac{4im_Bm_{K^*}[(C_V-C_V')(m_B+m_{K^*})A_{12}+m_b(C_7-C_7')T_{23}]}{\sqrt{q^2}(m_B+m_{K^*})},\\
H_0^A &= \frac{4im_Bm_{K^*}}{\sqrt{q^2}}(C_A-C_A')A_{12},\\
H_\pm^V &= \frac{i}{2(m_B+m_{K^*})}(\pm(C_V+C_V')\sqrt{\lambda_B}V-(m_B+m_{K^*})^2(C_V-C_V')A_1)+\frac{im_b}{q^2}(\pm(C_7+C_7')\sqrt{\lambda_B}T_1-(C_7-C_7')(m_B^2-m_{K^*}^2)T_2),\\
H_\pm^A &= \frac{i}{2(m_B+m_{K^*})}(\pm(C_A+C_A')\sqrt{\lambda_B}V-(m_B+m_{K^*})^2(C_A-C_A')A_1),\\
H^P &= \frac{i\sqrt{\lambda_B}}{2}\left(\frac{C_P-C_P'}{m_b+m_s}+\frac{m_{\ell_1}+m_{\ell_2}}{q^2}(C_A-C_A')\right)A_0,\\
H^S &= \frac{i\sqrt{\lambda_B}}{2}\left(\frac{C_S-C_S'}{m_b+m_s}+\frac{m_{\ell_1}-m_{\ell_2}}{q^2}(C_V-C_V')\right)A_0,\\
H_0^T &= \frac{2\sqrt{2}m_Bm_{K^*}}{m_B+m_{K^*}}(C_T+C_T')T_{23},\\
H_0^{T_t} &= \frac{2m_Bm_{K^*}}{m_B+m_{K^*}}(C_T-C_T')T_{23},\\
H_\pm^T &= \frac{1}{\sqrt{2q^2}}(\pm(C_T-C_T')\sqrt{\lambda_B}T_1-(C_T+C_T')(m_B^2-m_{K^*}^2)T_2),\\
H_\pm^{T_t} &= \frac{1}{2\sqrt{q^2}}(\pm(C_T+C_T')\sqrt{\lambda_B}T_1-(C_T-C_T')(m_B^2-m_{K^*}^2)T_2),
\end{aligned} \tag{C15}$$

where $C_{V(A)} = C_{9(10)}$ in the standard notation used in the literature and the q^2-dependence of the form factors is suppressed. Furthermore we have used

$$A_{12} = \frac{(m_B+m_{K^*})^2(m_B^2-m_{K^*}^2-q^2)A_1-\lambda_BA_2}{16m_Bm_{K^*}^2(m_B+m_{K^*})}, \qquad T_{23} = \frac{(m_B^2-m_{K^*}^2)(m_B^2+3m_{K^*}^2-q^2)T_2-\lambda_BT_3}{8m_Bm_{K^*}^2(m_B-m_{K^*})},$$

the same shorthand for zero-helicity form factor combinations as in [77,78].

The so-called timelike HAs, often denoted by H_t in the literature, have been absorbed into H^S and H^P. This is exceptional and follows from the vector and axial Ward identities $q^\mu \overline{u}(\ell_1)\gamma_\mu[\gamma_5]v(\ell_2) = (m_{\ell_1} \mp m_{\ell_2})\overline{u}(\ell_1)[\gamma_5]v(\ell_2)$. A similar simplification procedure could be repeated by use of the equation of motion $i\partial^\nu(\bar{s}i\sigma_{\mu\nu}b) = -(m_s + m_b)\bar{s}\gamma_\mu b + i\partial_\mu(\bar{s}b) - 2\bar{s}i\overleftarrow{D}_\mu b$ (as u sed in [82]) for $H_\lambda^{T_t}$ if all of the operators present in the equation were used in the effective Hamiltonian. Since the higher derivative operators are not present in the effective Hamiltonian used in this paper, such a simplification does not occur.

APPENDIX D: SPECIFIC RESULTS FOR $\overline{B} \to \overline{K}\ell_1\overline{\ell}_2$

The angular distribution for this decay is

$$\frac{d^2\Gamma}{dq^2 d\cos\theta_\ell} = G^{(0)}D^0_{0,0}(\Omega_\ell) + G^{(1)}D^1_{0,0}(\Omega_\ell) + G^{(2)}D^2_{0,0}(\Omega_\ell),$$

where, using the general leptonic HAs in Appendix A 3 and taking $m_{\ell_1} \neq m_{\ell_2}$, the functions $G^{(l_\ell)} = \mathcal{N}\tilde{G}^{(l_\ell)}$ [with $\mathcal{N}$ defined in (26)], are given in terms of $\overline{B} \to \overline{K}$ HAs by

$$\begin{aligned}
\tilde{G}^{(0)} &= \left(4(E_1E_2 + m_{\ell_1}m_{\ell_2}) + \frac{\lambda_{\gamma^*}}{3q^2}\right)|h^V|^2 + \left(4(E_1E_2 - m_{\ell_1}m_{\ell_2}) + \frac{\lambda_{\gamma^*}}{3q^2}\right)|h^A|^2 \\
&+ \left(4(E_1E_2 - m_{\ell_1}m_{\ell_2}) + \frac{\lambda_{\gamma^*}}{q^2}\right)|h^S|^2 + \left(4(E_1E_2 + m_{\ell_1}m_{\ell_2}) + \frac{\lambda_{\gamma^*}}{q^2}\right)|h^P|^2 \\
&+ 16\left(E_1E_2 + m_{\ell_1}m_{\ell_2} - \frac{\lambda_{\gamma^*}}{12q^2}\right)|h^{T_t}|^2 + 8\left(E_1E_2 - m_{\ell_1}m_{\ell_2} - \frac{\lambda_{\gamma^*}}{12q^2}\right)|h^T|^2 \\
&+ 16(m_{\ell_1}E_2 + m_{\ell_2}E_1)\mathrm{Im}[h^V\overline{h}^{T_t}] + 8\sqrt{2}(m_{\ell_1}E_2 - m_{\ell_2}E_1)\mathrm{Im}[h^A\overline{h}^T], \\
\tilde{G}^{(1)} &= -4\sqrt{\lambda_{\gamma^*}}\left(\mathrm{Re}\left[\frac{m_{\ell_1} + m_{\ell_2}}{\sqrt{q^2}}h^V\overline{h}^S + \frac{m_{\ell_1} - m_{\ell_2}}{\sqrt{q^2}}h^A\overline{h}^P\right] - \mathrm{Im}[2h^{T_t}\overline{h}^S + \sqrt{2}h^T\overline{h}^P]\right), \\
\tilde{G}^{(2)} &= -\frac{4\lambda_{\gamma^*}}{3q^2}(|h^V|^2 + |h^A|^2 - 2|h^T|^2 - 4|h^{T_t}|^2).
\end{aligned} \tag{D1}$$

The equivalent expressions for equal lepton masses are, using the notation $G^{(l_\ell)} = \mathcal{N}q^2\mathbb{G}^{(l_\ell)}$

$$\begin{aligned}
\mathbb{G}^{(0)} &= \frac{4}{3}(1 + 2\hat{m}_\ell{}^2)|h^V|^2 + \frac{4}{3}\beta_\ell^2|h^A|^2 + 2\beta_\ell^2|h^S|^2 + 2|h^P|^2 \\
&+ \frac{8}{3}(1 + 8\hat{m}_\ell{}^2)|h^{T_t}|^2 + \frac{4}{3}\beta_\ell^2|h^T|^2 + 16\hat{m}_\ell\mathrm{Im}[h^V\overline{h}^{T_t}], \\
\mathbb{G}^{(1)} &= -4\beta_\ell(2\hat{m}_\ell\mathrm{Re}[h^V\overline{h}^S] - \mathrm{Im}[2h^{T_t}\overline{h}^S + \sqrt{2}h^T\overline{h}^P]), \\
\mathbb{G}^{(2)} &= -\frac{4\beta_\ell^2}{3}(|h^V|^2 + |h^A|^2 - 2|h^T|^2 - 4|h^{T_t}|^2),
\end{aligned} \tag{D2}$$

where we have used the shorthand $\hat{m}_\ell \equiv m_\ell/\sqrt{q^2}$.

1. Explicit $\overline{B} \to \overline{K}$ helicity amplitudes in terms of form factors

As for $\overline{B} \to \overline{K}^*\ell_1\overline{\ell}_2$ we quote the HAs for form factor contributions only, which allows for comparison with the literature. Form factor computations are available for low q^2 and high q^2 from LCSR [15,16] and lattice QCD [17] respectively. Contributions to long-distance processes can be found in the same references as for the K^*-meson final state (quoted in Appendix C 5). The form factor matrix elements relevant to $\overline{B} \to \overline{K}$ transition, in standard parametrization, are

$$\begin{aligned}
\langle\overline{K}(p)|\bar{s}\gamma_\mu b|\overline{B}(p_B)\rangle &= (p_B + p)_\mu f_+(q^2) + \frac{m_B^2 - m_K^2}{q^2}q_\mu(f_0(q^2) - f_+(q^2)), \\
\langle\overline{K}(p)|\bar{s}\sigma_{\mu\nu} b|\overline{B}(p_B)\rangle &= i[(p_B + p)_\mu q_\nu - (p_B + p)_\nu q_\mu]\frac{f_T(q^2)}{m_B + m_K}, \\
\langle\overline{K}(p)|\bar{s}b|\overline{B}(p_B)\rangle &= \frac{m_B^2 - m_K^2}{m_b - m_s}f_0(q^2),
\end{aligned} \tag{D3}$$

with $\langle\overline{K}(p)|\bar{s}\gamma_\mu\gamma_5 b|\overline{B}(p_B)\rangle = \langle\overline{K}(p)|\bar{s}\gamma_5 b|\overline{B}(p_B)\rangle = 0$ in QCD. The hadronic HA is defined by

$$h^X = \langle\overline{K}|\bar{s}\overline{\Gamma}^X b|\overline{B}\rangle, \tag{D4}$$

where $\overline{\Gamma}^X|_{\lambda_X\to 0}$ as in Table 3.1 with $\omega \to \overline{\omega}$, containing the full set of dimension-six operators in the effective Hamiltonian (12). We find

$$\begin{aligned}
h^V &= \frac{\sqrt{\lambda_{BK}}}{2\sqrt{q^2}}\left(\frac{2m_b}{m_B+m_K}(C_7+C_7')f_T + (C_V+C_V')f_+\right),\\
h^A &= \frac{\sqrt{\lambda_{BK}}}{2\sqrt{q^2}}(C_A+C_A')f_+,\\
h^S &= \frac{m_B^2-m_K^2}{2}\left(\frac{(C_S+C_S')}{m_b-m_s}+\frac{m_{\ell_1}-m_{\ell_2}}{q^2}(C_V+C_V')\right)f_0,\\
h^P &= \frac{m_B^2-m_K^2}{2}\left(\frac{(C_P+C_P')}{m_b-m_s}+\frac{m_{\ell_1}+m_{\ell_2}}{q^2}(C_A+C_A')\right)f_0,\\
h^T &= -i\frac{\sqrt{\lambda_{BK}}}{\sqrt{2}(m_B+m_K)}(C_T-C_T')f_T,\\
h^{T_t} &= -i\frac{\sqrt{\lambda_{BK}}}{2(m_B+m_K)}(C_T+C_T')f_T,
\end{aligned} \tag{D5}$$

where the Källén function [cf. Eq. (B1)] $\lambda_{BK} \equiv \lambda(m_B^2, m_K^2, q^2)$ replaces $\lambda_B \equiv \lambda(m_B^2, m_{K^*}^2, q^2)$ and $C_{V(A)} = C_{9(10)}$ in the standard notation used in the literature.

2. Comparison with the literature

The results for equal lepton masses (D2) do agree with the results of Ref. [37] when $\theta_\ell \to \pi - \theta_\ell$ is taken into account. This is consistent with the angular conventions. In this paper we use the same conventions as LHCb [38], which differ from the ones of [37] by the transformation stated above.

APPENDIX E: $\Lambda_b \to \Lambda(\to (p,n)\pi)\ell_1\overline{\ell}_2$ ANGULAR DISTRIBUTION

The decay $\Lambda_b \to \Lambda(\to (p,n)\pi)\ell_1\overline{\ell}_2$ with a final-state proton or neutron, recently measured by the LHCb Collaboration [83], can also be considered within the generalized helicity formalism, and is particularly relevant because this decay can also be described using the effective Hamiltonian defined in A 2. In this case (5) becomes, in the rest frame of the Λ_b,

$$\begin{aligned}
&\mathcal{A}(\Omega_{\Lambda_b},\Omega_\ell,\Omega_\Lambda|\lambda_{\Lambda_b},\lambda_N,\lambda_1,\lambda_2)\\
&\sim \sum_{\lambda_\gamma,\lambda_\Lambda,J_\gamma}\delta_{\lambda_{\Lambda_b},\lambda_\gamma-\lambda_\Lambda}\mathcal{H}_{\lambda_\gamma\lambda_\Lambda}D^{\frac{1}{2}}_{\lambda_\Lambda,\lambda_N}(\Omega_\Lambda)\mathcal{N}_{\lambda_N}D^{J_\gamma}_{\lambda_\gamma,\lambda_\ell}(\Omega_\ell)\mathcal{L}_{\lambda_1\lambda_2}\\
&= \sum_{\lambda_\gamma,J_\gamma}\mathcal{H}_{\lambda_\gamma,\lambda_\gamma-\lambda_{\Lambda_b}}D^{\frac{1}{2}}_{\lambda_\gamma-\lambda_{\Lambda_b},\lambda_N}(\Omega_\Lambda)\mathcal{N}_{\lambda_N}D^{J_\gamma}_{\lambda_\gamma,\lambda_\ell}(\Omega_\ell)\mathcal{L}_{\lambda_1\lambda_2},
\end{aligned} \tag{E1}$$

where the leptonic HAs are the same as before and $\mathcal{N}_{\lambda_N}$ is the HA for the decay $\Lambda \to N\pi$ analogous to the $g_{K^*K\pi}$ factor in the $B \to K^*$ decay, this time carrying nontrivial dependence on helicities owing to the final state particle N having spin-$\frac{1}{2}$. The terms $\mathcal{H}_{\lambda_\gamma\lambda_\Lambda}$ are the HAs for the $\Lambda_b \to \Lambda$ decay and can be again expressed in the form

$$\mathcal{H}_{\lambda_\gamma\lambda_\Lambda} = \langle\Lambda(\lambda_\Lambda)|\bar{s}\overline{\Gamma}^X b|\Lambda_b(\lambda_{\Lambda_b})\rangle, \tag{E2}$$

with the $\overline{\Gamma}^X$ the same as defined in Table 3.1. The resulting angular distribution can then be expressed as

$$\begin{aligned}
K(q^2,\Omega_\Lambda,\Omega_\ell) \sim \mathrm{Re}[&\mathcal{K}_0^{0,0}\Omega_0^{0,0}(\Omega_\Lambda,\Omega_\ell) + \mathcal{K}_0^{0,1}\Omega_0^{0,1}(\Omega_\Lambda,\Omega_\ell)\\
&+ \mathcal{K}_0^{0,2}\Omega_0^{0,2}(\Omega_\Lambda,\Omega_\ell) + \mathcal{K}_0^{1,0}\Omega_0^{1,0}(\Omega_\Lambda,\Omega_\ell)\\
&+ \mathcal{K}_0^{1,1}\Omega_0^{1,1}(\Omega_\Lambda,\Omega_\ell) + \mathcal{K}_0^{1,2}\Omega_0^{1,2}(\Omega_\Lambda,\Omega_\ell)\\
&+ \mathcal{K}_1^{1,1}\Omega_1^{1,1}(\Omega_\Lambda,\Omega_\ell) + \mathcal{K}_1^{1,2}\Omega_1^{1,2}(\Omega_\Lambda,\Omega_\ell)],
\end{aligned} \tag{E3}$$

where $\Omega_\Lambda = (0,\theta_\Lambda,0)$ and $\Omega_\ell = (\phi,\theta_\ell,-\phi)$. A theoretical angular analysis of this decay has been performed in [41,84]; in terms of the functions defined in [41], the $\mathcal{K}_m^{l_\Lambda,l_\ell}$ above are

$$\begin{aligned}
\mathcal{K}_0^{0,0} &= \frac{1}{3}(K_{1cc}+2K_{1ss}), \qquad \mathcal{K}_0^{0,1} = K_{1c},\\
\mathcal{K}_0^{0,2} &= \frac{2}{3}(K_{1cc}-K_{1ss}), \qquad \mathcal{K}_0^{1,0} = \frac{1}{3}(K_{2cc}+2K_{2ss}),\\
\mathcal{K}_0^{1,1} &= K_{2c}, \qquad \mathcal{K}_0^{1,2} = \frac{2}{3}(K_{2cc}-K_{2ss}),\\
\mathcal{K}_1^{1,1} &= K_{3s}+iK_{4s}, \qquad \mathcal{K}_1^{1,2} = \frac{1}{\sqrt{3}}(K_{3sc}+iK_{4sc}).
\end{aligned} \tag{E4}$$

These results can also be compared with those found in [24]; it follows that the MoM will be equally useful in future angular analyses of this decay.

APPENDIX F: CHANGES IN CONVENTIONS AND PRESENTATION

Notational changes with respect to the arXiv version 1, aimed at clarifying the underlying structure, are as follows: (i) results are presented for $\overline{B} \to \overline{K}^{(*)}\ell_1\overline{\ell}_2$ rather than the conjugate decay, (ii) we use $C_T^{(\prime)}$ Wilson coefficients in place of $C_{T(5)}$ for the tensor operators cf. Appendix A 2 for details, (iii) the angular distribution (C1) is presented in terms of g_i in place of J_i in order to emphasize the differences of angular convention of this paper and the theory community (as discussed in Appendix C 2), (iv) lepton HAs are presented in the A, V rather than L, R basis and (v) timelike HAs are absorbed into scalar and pseudoscalar HAs. In addition we provide a *Mathematica* notebook, entitled notebookGHZ. nb, containing the results presented in Appendix C 4 for the decay mode $\overline{B} \to \overline{K}^*(\to K\pi)\ell_1\overline{\ell}_2$ for nonequal lepton masses [29].

[1] M. Jacob and G. Wick, On the general theory of collisions for particles with spin, Ann. Phys. (N.Y.) **7**, 404 (1959).

[2] J. D. Richman, An Experimenter's Guide to the Helicity Formalism (unpublished).

[3] H. E. Haber, Spin formalism and applications to new physics searches, arXiv:hep-ph/9405376.

[4] R. Kutschke, An Angular Distribution Cookbook (unpublished).

[5] P. Ball and R. Zwicky, $B_{d,s} \to \rho, \omega, K^*, \phi$ decay form-factors from light-cone sum rules revisited, Phys. Rev. D **71**, 014029 (2005).

[6] A. Khodjamirian, T. Mannel, and N. Offen, Form-factors from light-cone sum rules with B-meson distribution amplitudes, Phys. Rev. D **75**, 054013 (2007).

[7] M. Beneke, T. Feldmann, and D. Seidel, Systematic approach to exclusive $B \to V l^+ l^-, V\gamma$ decays, Nucl. Phys. **B612**, 25 (2001).

[8] T. Feldmann and J. Matias, Forward backward and isospin asymmetry for $B \to K^* \ell^+ \ell^-$ decay in the standard model and in supersymmetry, J. High Energy Phys. 01 (2003) 074.

[9] H. Asatryan, H. Asatrian, C. Greub, and M. Walker, Calculation of two loop virtual corrections to $b \to s \ell^+ \ell^-$ in the standard model, Phys. Rev. D **65**, 074004 (2002).

[10] M. Dimou, J. Lyon, and R. Zwicky, Exclusive chromomagnetism in heavy-to-light FCNCs, Phys. Rev. D **87**, 074008 (2013).

[11] J. Lyon and R. Zwicky, Isospin asymmetries in $B \to (K^*, \rho) \to l^+ l^-$ and $B \to K l^+ l^-$ in and beyond the standard model, Phys. Rev. D **88**, 094004 (2013).

[12] A. Khodjamirian, T. Mannel, A. Pivovarov, and Y.-M. Wang, Charm-loop effect in $B \to K^{(*)} \ell^+ \ell^-$ and $B \to K^* \gamma$, J. High Energy Phys. 09 (2010) 089.

[13] P. Ball, G. W. Jones, and R. Zwicky, $B \to V\gamma$ beyond QCD factorisation, Phys. Rev. D **75**, 054004 (2007).

[14] F. Muheim, Y. Xie, and R. Zwicky, Exploiting the width difference in $B_s \to \phi\gamma$, Phys. Lett. B **664**, 174 (2008).

[15] P. Ball and R. Zwicky, New results on $B \to \pi, K, \eta$ decay formfactors from light-cone sum rules, Phys. Rev. D **71**, 014015 (2005).

[16] A. Khodjamirian, T. Mannel, and Y. Wang, $B \to K \ell^+ \ell^-$ decay at large hadronic recoil, J. High Energy Phys. 02 (2013) 010.

[17] C. Bouchard, G. P. Lepage, C. Monahan, H. Na, and J. Shigemitsu (HPQCD Collaboration), Standard Model Predictions for $B \to K \ell^+ \ell^-$ with Form Factors from Lattice QCD, Phys. Rev. Lett. **111**, 162002 (2013).

[18] G. Hiller and R. Zwicky, (A)symmetries of weak decays at and near the kinematic endpoint, J. High Energy Phys. 03 (2014) 042.

[19] D. Das, G. Hiller, M. Jung, and A. Shires, The $\overline{B} \to \overline{K} \pi \ell \ell$ and $\overline{B}_s \to \overline{K} K \ell \ell$ distributions at low hadronic recoil, J. High Energy Phys. 09 (2014) 109.

[20] L. Hofer and J. Matias, Exploiting the symmetries of P and S wave for $B \to K^* \mu^+ \mu^-$, J. High Energy Phys. 09 (2015) 104.

[21] S. Descotes-Genon and J. Virto, Time dependence in $B \to V \ell \ell$ decays, J. High Energy Phys. 04 (2015) 045.

[22] F. Kruger, L. M. Sehgal, N. Sinha, and R. Sinha, Angular distribution and CP asymmetries in the decays $\overline{B} \to K^- \pi^+ e^- e^+$ and $\overline{B} \to \pi^- \pi^+ e^- e^+$, Phys. Rev. D **61**, 114028 (2000).

[23] A. Faessler, T. Gutsche, M. Ivanov, J. Korner, and V. E. Lyubovitskij, The exclusive rare decays $B \to \mathrm{K}(\mathrm{K}*)\overline{\ell}\ell$ and $B_c \to \mathrm{D}(\mathrm{D}*)\overline{\ell}\ell$ in a relativistic quark model, Eur. Phys. J. direct **4**, 1 (2002).

[24] F. Beaujean, M. Chrzaszcz, N. Serra, and D. van Dyk, Extracting angular observables without a likelihood and applications to rare decays, Phys. Rev. D **91**, 114012 (2015).

[25] R. Aaij *et al.* (LHCb Collaboration), Test of Lepton Universality Using $B^+ \to K^+ \ell^+ \ell^-$ Decays, Phys. Rev. Lett. **113**, 151601 (2014).

[26] R. Aaij *et al.* (LHCb Collaboration), Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0} \mu^+ \mu^-$, Phys. Rev. Lett. **111**, 191801 (2013).

[27] R. Aaij *et al.* (LHCb Collaboration), Angular Analysis of the $B^0 \to K^{*0} \mu^+ \mu^-$ Decay (to be published).

[28] S. L. Glashow, D. Guadagnoli, and K. Lane, Lepton Flavor Violation in B Decays?, Phys. Rev. Lett. **114**, 091801 (2015).

[29] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.93.054008, where we provide the mathematica notebook notebookGHZ.nb, containing explicit expressions for the angular distribution and helicity amplitudes given in Eqs. (C12) and (C15).

[30] A. S. Dighe, I. Dunietz, and R. Fleischer, Extracting CKM phases and $B_s - \overline{B}_s$ mixing parameters from angular distributions of nonleptonic B decays, Eur. Phys. J. C **6**, 647 (1999).

[31] C.-D. Lu and W. Wang, Analysis of $B \to K_J^*(\to K\pi)\mu^+\mu^-$ in the higher kaon resonance region, Phys. Rev. D **85**, 034014 (2012).

[32] R. Aaij *et al.* (LHCb Collaboration), Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0} \mu^+ \mu^-$, J. High Energy Phys. 08 (2013) 131.

[33] K. Olive *et al.* (Particle Data Group Collaboration), Review of particle physics, Chin. Phys. C **38**, 090001 (2014).

[34] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, Optimizing the basis of $B \to K^* \ell^+ \ell^-$ observables in the full kinematic range, J. High Energy Phys. 05 (2013) 137.

[35] W. Altmannshofer, P. Ball, A. Bharucha, A. J. Buras, D. M. Straub, and M. Wick, Symmetries and asymmetries of $B \to K^* \mu^+ \mu^-$ decays in the standard model and beyond, J. High Energy Phys. 01 (2009) 019.

[36] C. Bobeth, G. Hiller, and D. van Dyk, General analysis of $\overline{B} \to \overline{K}^{(*)} \ell^+ \ell^-$ decays at low recoil, Phys. Rev. D **87**, 034016 (2013).

[37] C. Bobeth, G. Hiller, and G. Piranishvili, Angular distributions of $\overline{B} \to \overline{K} \overline{\ell} \ell$ decays, J. High Energy Phys. 12 (2007) 040.

[38] R. Aaij *et al.* (LHCb Collaboration), Angular analysis of charged and neutral $B \to K \mu^+ \mu^-$ decays, J. High Energy Phys. 05 (2014) 082.

[39] U. Egede, M. Patel, and K. A. Petridis, Method for an unbinned measurement of the, q^2 dependent decay amplitudes of $\bar{B}^0 \to \bar{K}^{*0} \mu^+ \mu^-$ decays, arXiv:1504.00574.

[40] K. Hecht, *Quantum Mechanics* (Springer, New York, 2000).

[41] P. Boer, T. Feldmann, and D. van Dyk, Angular analysis of the decay $\Lambda_b \to \Lambda(\to N\pi)\ell^+\ell^-$, J. High Energy Phys. 01 (2015) 155.

[42] D. Becirevic and A. Tayduganov, Impact of $B \to K_0^* \ell^+ \ell^-$ on the new physics search in $B \to K^* \ell^+ \ell^-$ decay, Nucl. Phys. **B868**, 368 (2013).

[43] T. Blake, U. Egede, and A. Shires, The effect of S-wave interference on the $B^0 \to K^{*0} \ell^+ \ell^-$ angular observables, J. High Energy Phys. 03 (2013) 027.

[44] R. Zwicky, Endpoint symmetries of helicity amplitudes, arXiv:1309.7802.

[45] J. Gratrex, M. Hopfer, and R. Zwicky (to be published).

[46] G. Hiller and F. Kruger, More model independent analysis of $b \to s$ processes, Phys. Rev. D **69**, 074020 (2004).

[47] C. Bobeth, G. Hiller, and G. Piranishvili, Angular distributions of $\overline{B} \to K \overline{\ell} \ell$ decays, J. High Energy Phys. 12 (2007) 040.

[48] J.-T. Wei *et al.* (Belle Collaboration), Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for $B \to K^{(*)} \ell^+ \ell^-$, Phys. Rev. Lett. **103**, 171801 (2009).

[49] J. Lees *et al.* (*BABAR* Collaboration), Measurement of branching fractions and rate asymmetries in the rare decays $B \to K^{(*)} l^+ l^-$, Phys. Rev. D **86**, 032012 (2012).

[50] B. Gripaios, M. Nardecchia, and S. Renner, Composite leptoquarks and anomalies in B-meson decays, J. High Energy Phys. 05 (2015) 006.

[51] D. Becirevic, S. Fajfer, and N. Kosnik, Lepton flavor nonuniversality in $b \to s \ell^+ \ell^-$ processes, Phys. Rev. D **92**, 014016 (2015).

[52] C. Niehoff, P. Stangl, and D. M. Straub, Violation of lepton flavour universality in composite Higgs models, Phys. Lett. B **747**, 182 (2015).

[53] A. Crivellin, L. Hofer, J. Matias, U. Nierste, S. Pokorski, and J. Rosiek, Lepton-flavour violating B decays in generic Z' models, Phys. Rev. D **92**, 054013 (2015).

[54] A. Greljo, G. Isidori, and D. Marzocca, On the breaking of lepton flavor universality in B decays, J. High Energy Phys. 07 (2015) 142.

[55] R. Alonso, B. Grinstein, and J. M. Camalich, Lepton universality violation and lepton flavor conservation in B-meson decays, J. High Energy Phys. 10 (2015) 184.

[56] I. de Medeiros Varzielas and G. Hiller, Clues for flavor from rare lepton and quark decays, J. High Energy Phys. 06 (2015) 072.

[57] A. Crivellin, G. D'Ambrosio, and J. Heeck, Addressing the LHC flavor anomalies with horizontal gauge symmetries, Phys. Rev. D **91**, 075006 (2015).

[58] L. Calibbi, A. Crivellin, and T. Ota, Effective Field Theory Approach to $b \to s \ell \ell^{(\prime)}$, $B \to K^{(*)} \nu \overline{\nu}$ and $B \to D^{(*)} \tau \nu$ with Third Generation Couplings, Phys. Rev. Lett. **115**, 181801 (2015).

[59] A. Crivellin, G. D'Ambrosio, and J. Heeck, Explaining $h \to \mu^{\pm} \tau^{\mp}$, $B \to K^* \mu^+ \mu^-$ and $B \to K \mu^+ \mu^- / B \to K e^+ e^-$ in a Two-Higgs-Doublet Model with Gauged $L_\mu - L_\tau$, Phys. Rev. Lett. **114**, 151801 (2015).

[60] T. Blake, T. Gershon, and G. Hiller, Rare b hadron decays at the LHC, Annu. Rev. Nucl. Part. Sci. **65**, 113 (2015).

[61] A. Guevara, G. L. Castro, P. Roig, and S. Tostado, Long-distance weak annihilation contribution to $B^{\pm} \to (\pi^{\pm}, K^{\pm}) \ell^+ \ell^-$ decays, Phys. Rev. D **92**, 054035 (2015).

[62] T. Huber, T. Hurth, and E. Lunghi, Inclusive $B \to X_s \ell^+ \ell^-$: Complete angular analysis and a thorough study of collinear photons, J. High Energy Phys. 06 (2015) 176.

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

[64] R. Aaij *et al.* (LHCb Collaboration), Differential branching fractions and isospin asymmetries of $B \to K^{(*)} \mu^+ \mu^-$ decays, J. High Energy Phys. 06 (2014) 133.

[65] S. Descotes-Genon, J. Matias, and J. Virto, Understanding the $B \to K^* \mu^+ \mu^-$ anomaly, Phys. Rev. D **88**, 074002 (2013).

[66] W. Altmannshofer and D. M. Straub, New physics in $B \to K^* \mu \mu$?, Eur. Phys. J. C **73**, 2646 (2013).

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

[68] A. Bozek *et al.* (Belle Collaboration), Observation of $B^+ \to \overline{D}^{*0} \tau^+ \nu_\tau$ and evidence for $B^+ \to \overline{D}^0 \tau^+ \nu_\tau$ at Belle, Phys. Rev. D **82**, 072005 (2010).

[69] J. Lees *et al.* (*BABAR* Collaboration), Measurement of an excess of $\overline{B} \to D^{(*)} \tau^- \overline{\nu}_\tau$ decays and implications for charged Higgs bosons, Phys. Rev. D **88**, 072012 (2013).

[70] R. Aaij *et al.* (LHCb Collaboration), Measurement of the Ratio of Branching Fractions $\mathcal{B}(\overline{B}^0 \to D^{*+} \tau^- \overline{\nu}_\tau) / \mathcal{B}(\overline{B}^0 \to D^{*+} \mu^- \overline{\nu}_\mu)$, Phys. Rev. Lett. **115**, 111803 (2015).

[71] B. Dey, Angular analyses of exclusive $\overline{B} \to X \ell_1 \ell_2$ with complex helicity amplitudes, Phys. Rev. D **92**, 033013 (2015).

[72] F. Kruger and J. Matias, Probing new physics via the transverse amplitudes of $B_0 \to K^{*,0}(\to K^- \pi^+) \ell^+ \ell^-$ at large recoil, Phys. Rev. D **71**, 094009 (2005).

[73] C. Kim and T. Yoshikawa, Systematic analysis of $B \to K \pi \ell^+ \ell^-$ decay through angular decomposition, arXiv:0711.3880.

[74] A. K. Alok, A. Dighe, D. Ghosh, D. London, J. Matias, M. Nagashima, and A. Szynkman, New-physics contributions to the forward-backward asymmetry in $B \to K^* \mu^+ \mu^-$, J. High Energy Phys. 02 (2010) 053.

[75] A. K. Alok, A. Datta, A. Dighe, M. Duraisamy, D. Ghosh, and D. London, New physics in $b \to s \mu^+ \mu^-$: CP-conserving observables, J. High Energy Phys. 11 (2011) 121.

[76] N. Serra (private communication).

[77] A. Bharucha, D. M. Straub, and R. Zwicky, $B \to V \ell^+ \ell^-$ in the Standard Model from Light-Cone Sum Rules, arXiv:1503.05534.

[78] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, Lattice QCD calculation of form factors describing the rare decays $B \to K^* \ell^+ \ell^-$ and $B_s \to \phi \ell^+ \ell^-$, Phys. Rev. D **89**, 094501 (2014).

[79] B. Grinstein and D. Pirjol, Exclusive rare $B \to K^* \ell^+ \ell^-$ decays at low recoil: Controlling the long-distance effects, Phys. Rev. D **70**, 114005 (2004).

[80] M. Beylich, G. Buchalla, and T. Feldmann, Theory of $B \to K^{(*)} \ell^+ \ell^-$ decays at high q^2: OPE and quark-hadron duality, Eur. Phys. J. C **71**, 1635 (2011).

[81] R. Aaij *et al.* (LHCb Collaboration), Observation of a Resonance in $B^+ \to K^+ \mu^+ \mu^-$ Decays at Low Recoil, Phys. Rev. Lett. **111**, 112003 (2013).

[82] C. Hambrock, G. Hiller, S. Schacht, and R. Zwicky, $B \to K^\star$ form factors from flavor data to QCD and back, Phys. Rev. D **89**, 074014 (2014).

[83] R. Aaij *et al.* (LHCb Collaboration), Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda^0 \mu^+ \mu^-$ decays, J. High Energy Phys. 06 (2015) 115.

[84] T. Gutsche, M. A. Ivanov, J. G. Korner, V. E. Lyubovitskij, and P. Santorelli, Rare baryon decays $\Lambda_b \to \Lambda l^+ l^- (l = e, \mu, \tau)$ and $\Lambda_b \to \Lambda \gamma$: Differential and total rates, lepton- and hadron-side forward-backward asymmetries, Phys. Rev. D **87**, 074031 (2013).