

# Probing flavoured Axions in the Tail of $B_q \rightarrow \mu^+ \mu^-$

Johannes Albrecht<sup>\*a</sup>, Emmanuel Stamou<sup>†b,c</sup>,  
 Robert Ziegler<sup>‡d,e</sup>, Roman Zwicky<sup>§f</sup>

<sup>a</sup>*Fakultät für Physik, TU Dortmund, D-44221 Dortmund, Germany*

<sup>b</sup>*Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA*

<sup>c</sup>*Institut de Théorie des Phénomènes Physiques, EPFL, Lausanne, Switzerland*

<sup>d</sup>*Theoretical Physics Department, CERN, 1211 Geneva 23, Switzerland*

<sup>e</sup>*Institut für Theoretische Teilchenphysik, KIT, 76131 Karlsruhe, Germany*

<sup>f</sup>*Higgs Centre for Theoretical Physics, School of Physics and Astronomy,  
 University of Edinburgh, Edinburgh EH9 3JZ, Scotland*

November 13, 2019

## Abstract

We discuss how LHC di-muon data collected to study  $B_q \rightarrow \mu\mu$  can be used to constrain light particles with flavour-violating couplings to  $b$ -quarks. Focussing on the case of a flavoured QCD axion,  $a$ , we compute the decay rates for  $B_q \rightarrow \mu\mu a$  and the SM background process  $B_q \rightarrow \mu\mu\gamma$  near the kinematic endpoint. These rates depend on non-perturbative  $B_q \rightarrow \gamma^{(*)}$  form factors with on- or off-shell photons. The off-shell form factors –relevant for generic searches for beyond-the-SM particles— are discussed in full generality and computed with QCD sum rules for the first time. With these results, we analyse available LHCb data to obtain the sensitivity on  $B_q \rightarrow \mu\mu a$  at present and future runs. We find that the full LHCb dataset alone will allow to probe axion-coupling scales of the order of  $10^6$  GeV for both  $b \rightarrow d$  and  $b \rightarrow s$  transitions.

---

<sup>\*</sup>johannes.albrecht@cern.ch

<sup>†</sup>emmanuel.stamou@epfl.ch

<sup>‡</sup>robert.ziegler@cern.ch

<sup>§</sup>roman.zwicky@ed.ac.uk

# 1 Introduction and Motivation

Open questions in particle physics and cosmology may well be addressed by very light particles that interact only feebly with the Standard Model (SM). The prime example is the QCD axion [1, 2], which is not only predicted within the Peccei–Quinn (PQ) [3, 4] solution to the strong CP Problem, but which can also explain the Dark Matter abundance if it is sufficiently lighter than the meV scale [5–7]. In the past years much activity has been devoted towards experimental searches for the QCD axion, and multiple proposals for new experiments are underway to complement ongoing efforts to discover the axion, see Ref. [8] for a review.

While most axion searches rely on axion couplings to photons, the axion also couples to SM fermions if they are charged under the PQ symmetry. Generically, these charges constitute new sources of flavour violation, which induce flavour-violating axion couplings to fermions, which can thus be probed by precision flavour experiments. For instance, this situation arises naturally when the PQ symmetry is identified with a flavour symmetry that shapes the hierarchical structure of the SM Yukawas [9–12], therefore, connecting the strong CP problem with the SM flavour puzzle. Even in the absence of such a connection, axion models with flavour non-universal PQ charges can be easily constructed and motivated by, e.g., stellar cooling anomalies that require suppressed axion couplings to nucleons [13–15].

In the absence of explicit models, the couplings of the axion to different flavours are *a priori* unrelated, and are parametrised by a model-independent effective Lagrangian for Goldstone bosons. The flavour-violating couplings in the various quark and lepton sectors can then be constrained by experimental data, see Ref. [16] for a recent assessment of the relevant bounds in the quark sector using mainly hadron decays with missing energy. In this article we explore a novel direction to probe flavour-violating axion couplings involving  $b$ -quarks using the present and future LHC data collected to study  $B_q \rightarrow \mu\mu$ .

We therefore focus on flavour-violating  $b \rightarrow q$  transitions, which are described by the Lagrangian

$$\mathcal{L} = \frac{\partial_\mu a}{2f_a} \bar{b} \gamma^\mu (C_{bq}^V + C_{bq}^A \gamma_5) q + \text{h.c.} \equiv \partial_\mu a \bar{b} \gamma^\mu \left( \frac{1}{F_{bq}^V} + \frac{\gamma_5}{F_{bq}^A} \right) q + \text{h.c.}, \quad (1.1)$$

where  $F_{bq}^{V/A}$  are parity odd/even couplings,  $q = d, s$  and  $a$  denotes the derivatively coupled QCD axion, whose mass is inversely proportional to the axion decay constant,  $f_a$ , which suppresses all axion couplings. The decay constant has to be much larger than the electroweak scale to sufficiently decouple the axion from the SM in order to satisfy experimental constraints [17, 18]. This implies that the axion is light, with a mass much below an eV, and stable even on cosmological scales.

Therefore, two-body  $B$ -meson decays with missing energy, which closely resemble the very rare SM decays with final-state neutrino pairs that have been looked for at  $B$ -factories, stringently constrain the couplings in Eq. (1.1). The resulting constraints on the vector couplings  $F_{bq}^V$  (from  $B \rightarrow K/\pi a$  decays) and the axial-vector couplings  $F_{bq}^A$  (from  $B \rightarrow K^*/\rho a$  decays and  $B_q$  mixing) have been given in Refs. [16] (see also Refs. [19, 20]) and are summarised in Table 1. Note that constraints from neutral

	$F_{bq}^V$ [GeV]	$F_{bq}^A$ [GeV]
$bd$	$1.2 \cdot 10^8$ ( $B \rightarrow \pi a$ )	$4.8 \cdot 10^6$ ( $B - \bar{B}$ mixing)
$bs$	$3.1 \cdot 10^8$ ( $B \rightarrow Ka$ )	$1.3 \cdot 10^8$ ( $B \rightarrow K^* a$ )

**Table 1:** Lower bounds on  $F_{bq}^{V,A}$  at 90% CL from  $B$ -decays and  $B_q$ -mixing, taken from Ref. [16].

meson mixing are typically much weaker than the ones from decays to vector mesons, except in the case of  $b \rightarrow d$  transitions. This is mainly due to the lack of experimental data on  $B \rightarrow \rho\nu\bar{\nu}$  suitable for the two-body recast.

In the present work, we investigate whether the couplings in Eq. (1.1) can also be constrained at the LHC. To this end, we propose to use the three-body decays  $B_{s,d} \rightarrow \mu\mu a$ , where the muon pair originates from an off-shell photon, cf., Figure 1 (left). With the main goal of measuring the SM decay  $B_q \rightarrow \mu\mu$ , the ATLAS [21], CMS [22] and LHCb [23] collaborations have collected di-muon events with an invariant mass  $q^2$  down to roughly  $(5 \text{ GeV})^2$ . As long as no vetos on extra particles in the event are applied, these datasets can be used to constrain decays with additional particles in the final state, e.g., the radiative decay  $B_q \rightarrow \mu\mu\gamma$ , as proposed in Ref. [24]. Here, we point out that the same datasets can be used to constrain the decays  $B_q \rightarrow \mu\mu X$ , where  $X$  is a neutral, beyond-the-SM (BSM) particle with a mass that is sufficiently small to be kinematically allowed at the tail of  $B_q \rightarrow \mu\mu$ , i.e.,  $m_X \lesssim m_{B_q} - 5 \text{ GeV} \approx 300 \text{ MeV}$ . In this respect, the radiative decay  $B_q \rightarrow \mu\mu\gamma$  merely constitutes a SM background, which we take into account in our analysis. In particular, we suggest that when the measurement of  $B_q \rightarrow \mu\mu\gamma$  becomes feasible in the future, it can be directly interpreted in terms of constraining BSM particles that replace the final state photon. A similar strategy can be applied to  $s \rightarrow d$  transitions, using for example the di-muon data collected at LHCb to study  $K_S \rightarrow \mu\mu$ , cf., Ref. [25], and possibly also to  $c \rightarrow u$  transitions, i.e.,  $D \rightarrow \mu\mu$  [26].

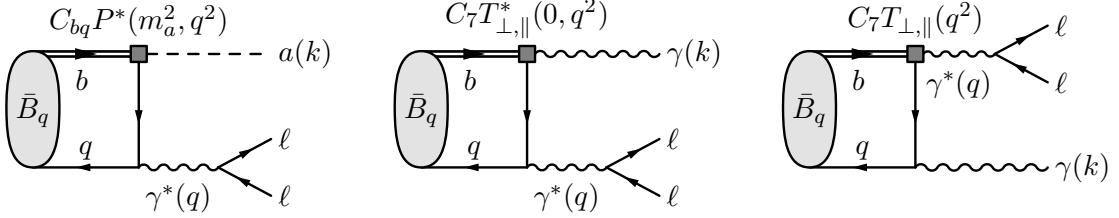
In the following we focus on the case of the invisible QCD axion,  $a$ , but our analysis can be readily extended to other particles appearing in the final state, as long as they are not vetoed in the event. In particular these could be heavy axions decaying within the detector, i.e., axion-like particles (ALPs). We expect such an analysis to be fully inclusive, that is, independent of the ALP decay mode. Similarly our proposal can be extended to constrain light vectors with flavour-violating couplings, e.g., dark photons or  $Z'$ 's. In this article we demonstrate the key elements of the analysis and perform the first sensitivity studies based on the published dataset of the LHCb collaboration. The ATLAS and CMS data can be analysed analogously.

The photon off-shell form factors are necessary for predicting branching fractions of  $B_q \rightarrow \ell\ell X$  where  $X$  is any of the above mentioned light BSM particles. We discuss the complete set of form factors, relevant for the dimension-six effective Hamiltonian, compute them with QCD sum rules and fit them to a  $z$ -expansion. In addition the off-shell basis is shown to be related to the standard  $B \rightarrow V = \rho^0, \omega, \phi \dots$  basis through a dispersion representation, which interrelates many properties of these two sets of form factors.

This article is organised as follows: In Section 2 we provide the differential rates for the axionic decay  $B_q \rightarrow \mu\mu a$  and the radiative decay  $B_q \rightarrow \mu\mu\gamma$ . In Section 3 we provide the tools necessary to perform the analysis and use available background estimates and data from LHCb's  $B_s \rightarrow \mu\mu$  measurement to evaluate the sensitivity to  $B_q \rightarrow \mu\mu a$  at present and future runs. We conclude in Section 4. The Appendix A is devoted to various aspects of the  $B \rightarrow \gamma^{(*)}$  form factors.

## 2 Differential Decay Rates

In this section we calculate the differential rates for the axionic  $B_q \rightarrow \ell\ell a$  and radiative  $B_q \rightarrow \ell\ell\gamma$  decay channels. In Figure 1, we show on the left the diagram for the axionic decay and in the centre and on the right representative diagrams for the radiative decay. The rates are differential in the lepton-pair momentum  $q \equiv p_{\ell+} + p_{\ell-}$ , and depend on non-perturbative  $B_q \rightarrow \gamma^{(*)}$  form factors with on- or off-shell photons, which we briefly introduce before presenting the differential decay rates. Finally, we evaluate the rates close to the kinematic endpoint  $(4.9 \text{ GeV})^2 \lesssim q^2 < m_{B_q}^2$ , and compare our prediction for the radiative decay to results in the literature.



**Figure 1:** The diagram to the left is the main axion process  $B_q \rightarrow \ell\ell a$  whereas the two diagrams in the centre and the right belong to the  $B_q \rightarrow \ell\ell\gamma$  background. The single and double lines stand for the  $q$  and  $b$ -quark, respectively. The left and central diagrams depend on off-shell form factors in the sense that the photon that emits the two leptons is off-shell. Diagrams in which the photon couples to  $b$ -quarks are not shown, but are analogous. Also diagrams with  $Q_{9,10}$ -operator insertions are not shown, and resemble the diagram on the right and are proportional to  $C_9 V_\perp$  and  $C_{10} V_\parallel$ .

## 2.1 The $B_q \rightarrow \gamma^*$ form factors

We describe  $B_q(p_B) \rightarrow \gamma^*(k)$  transitions with off-shell photons by a set of form factors with two arguments  $F^*(q^2, k^2) \equiv F^{B \rightarrow \gamma^*}(q^2, k^2)$ . The first argument (here  $q^2$ ) denotes the momentum transfer at the flavour-violating vertex while the second argument (here  $k^2$ ) denotes the momentum of the photon. For on-shell photons, i.e.  $k^2 = 0$ , these form factors reduce to the well-known on-shell form factors  $F(q^2) \equiv F^*(q^2, 0)$  given in Eq. (A.23). A complete<sup>1</sup> set of form factors is given by

$$\begin{aligned}
 M_5^\rho(q, k) &\equiv b_P \langle \gamma^*(k, \rho) | \bar{q} \gamma_5 b | \bar{B}_q(p_B) \rangle &= i m_{B_q} R_\perp^\rho P^*(q^2, k^2), \\
 M_V^{\mu\rho}(q, k) &\equiv b_V \langle \gamma^*(k, \rho) | \bar{q} \gamma^\mu b | \bar{B}_q(p_B) \rangle &= R_\perp^{\mu\rho} V_\perp^*(q^2, k^2), \\
 M_A^{\mu\rho}(q, k) &\equiv b_V \langle \gamma^*(k, \rho) | \bar{q} \gamma^\mu \gamma_5 b | \bar{B}_q(p_B) \rangle &= R_\parallel^{\mu\rho} V_\parallel^*(q^2, k^2) + R_\perp^{\mu\rho} V_\perp^*(q^2, k^2) + R_P^{\mu\rho} V_P^*(q^2, k^2), \\
 M_T^{\mu\rho}(q, k) &\equiv b_T \langle \gamma^*(k, \rho) | \bar{q} i q_\nu \sigma^{\mu\nu} b | \bar{B}_q(p_B) \rangle &= R_\perp^{\mu\rho} T_\perp^*(q^2, k^2), \\
 M_{T_5}^{\mu\rho}(q, k) &\equiv b_T \langle \gamma^*(k, \rho) | \bar{q} i q_\nu \sigma^{\mu\nu} \gamma_5 b | \bar{B}_q(p_B) \rangle &= - (R_\parallel^{\mu\rho} T_\parallel^*(q^2, k^2) + R_\perp^{\mu\rho} T_\perp^*(q^2, k^2)), \tag{2.1}
 \end{aligned}$$

where  $q \equiv p_B - k$  denotes the momentum transfer at the flavour-violating vertex, and we define the off-shell photon state  $\langle \gamma^*(k, \rho) |$  in Eq. (A.2). The coefficients

$$b_P \equiv \frac{m_b + m_q}{s_e e}, \quad b_V \equiv - \frac{m_{B_q}}{s_e e}, \quad b_T \equiv \frac{1}{s_e e}, \tag{2.2}$$

depend on the sign convention,  $s_e$ , for the covariant derivative  $D_\mu = \partial_\mu + s_e i Q_f e A_\mu$ . The Lorentz tensors  $R^\rho, R_{\perp, \parallel, \perp}^{\mu\rho} \equiv R^\rho(q, k), R_{\perp, \parallel, \perp}^{\mu\rho}(q, k)$  are defined in Eq. (A.3), and the matrix element satisfy the QED and the axial Ward identities

$$k_\rho M_5^\rho(q, k) = k_\rho M_{V,A,T,T_5}^{\mu\rho}(q, k) = 0, \quad q_\mu M_A^{\mu\rho}(q, k) = m_{B_q} M_5^\rho(q, k). \tag{2.3}$$

The latter implies  $P^*(q^2, k^2) = q^2/(2m_{B_q}^2) V_P^*(q^2, k^2)$  and reduces the number of independent form factors down to a total of seven. At  $q^2 = 0$  there are two further constraints

$$P^*(0, k^2) = \hat{V}_\perp^*(0, k^2), \quad T_\parallel^*(0, k^2) = (1 - \frac{k^2}{m_{B_q}^2}) T_\perp^*(0, k^2), \tag{2.4}$$

---

<sup>1</sup>The scalar form factor,  $\langle \gamma^*(k, \rho) | \bar{q} b | \bar{B}_q(p_B) \rangle$ , vanishes due to parity conservation of QCD.

where  $\hat{V}_{\mathbb{L}}^*(q^2, k^2) \equiv -q^2/(2m_{B_q}^2)V_{\mathbb{L}}^*(q^2, k^2)$  thereby reducing the form factors down to five. An extensive discussion including dispersion representations in the  $q^2$  and  $k^2$  variables, the derivation of Eq. (2.4), the limit to photon on-shell form factors, and their computation from QCD sum rules are deferred to Appendix A. The off-shell form factors in the limit of small momentum transfer at the flavor-violating vertex,  $T_{\perp, \parallel, \mathbb{L}}^*(0, q^2)$ ,  $V_{\perp, \parallel, \mathbb{L}}^*(0, q^2)$  and  $P^*(0, q^2)$  are computed in this work for the first time.<sup>2</sup> Moreover in Ref. [29], the off-shell form factor  $T_{\perp}^*(0, k^2) = F_{TV}(0, k^2)$  is evaluated using a vector-meson-dominance approximation. For the on-shell form factors  $B \rightarrow \gamma$  we use the leading-order (LO) version of the soon-to-appear next-to-LO (NLO) light-cone sum rule (LCSR) computation [30]. Note that the QCD sum rule result of the off-shell form factors can be used in the relevant kinematic region  $(4.9 \text{ GeV})^2 \lesssim q^2 < m_{B_q}^2$  since thresholds are far away. The photon on-shell form factors are more challenging in this region because the light-cone expansion breaks down. They can, however, be extrapolated to this region by using a  $B_q^*$  and  $B_1$ -pole ansatz, with the residue computed from LCSR [31], supplemented with  $z$ -expansion to account for further states.

## 2.2 The $B_q \rightarrow \ell \ell a$ differential rate

Given the effective Lagrangian in Eq. (1.1), the amplitude for  $\bar{B}_q(p_B) \rightarrow \ell^+(p_{\ell+}) \ell^-(p_{\ell-}) a(k)$  is<sup>3</sup>

$$\mathcal{A}_{\mu\mu a} = i \frac{e^2 Q_\ell}{m_{B_q} q^2} \frac{k_\mu}{F_{bq}^A} M_A^{\mu\rho}(k, q) \bar{u}^s(p_{\ell-}) \gamma_\rho v^r(p_{\ell+}) = i \frac{e^2 Q_\ell}{F_{bq}^A q^2} M_5^\rho(k, q) \bar{u}^s(p_{\ell-}) \gamma_\rho v^r(p_{\ell+}), \quad (2.5)$$

where  $\mathcal{A}_{\mu\mu a} \equiv \langle \mu\mu a | (-\mathcal{H}_{\text{eff}}) | \bar{B}_q \rangle$ ,  $q \equiv p_B - k = p_{\ell+} + p_{\ell-}$  and  $Q_\ell = -1$  denotes the lepton charge. After squaring this amplitude, summing over fermion spins, and integrating over the unobserved axion momentum, the differential rate in the invariant mass of the final-state leptons,  $q^2$ , becomes

$$\frac{d\Gamma}{dq^2}(B_q \rightarrow \ell \ell a) = \frac{\alpha^2}{48\pi m_{B_q}} \frac{\lambda_\gamma^{1/2} (\lambda_{B_q}^{(a)})^{3/2}}{|F_{bq}^A|^2} \frac{2m_\ell^2 + q^2}{q^8} |P^*(m_a^2, q^2)|^2, \quad (2.6)$$

where  $\lambda_\gamma \equiv \lambda(q^2, m_\ell^2, m_\ell^2)$ ,  $\lambda_{B_q}^{(a)} \equiv \lambda(m_{B_q}^2, q^2, m_a^2)$ , and  $\lambda(x, y, z) \equiv x^2 + y^2 + z^2 - 2xy - 2xz - 2yz$  is the Källén function. For our work it is sufficient to approximate  $m_a \rightarrow 0$ .

## 2.3 The $B_q \rightarrow \ell \ell \gamma$ differential rate

The relevant part of the effective SM Lagrangian is<sup>4</sup>

$$\mathcal{L}_{\text{SM}} = \frac{4G_F}{\sqrt{2}} V_{tb}^* V_{tq} \sum_{i=7,9,10} (C_i Q_i + C'_i Q'_i) + \text{h.c.},$$

---

<sup>2</sup> The weak annihilation process,  $B \rightarrow V\gamma^*$  matrix elements of four-quark operators, contain some of these form factors as sub processes. Weak annihilation has been computed in the SM to LO in QCD factorisation [27] and including all BSM operators in LCSR [28]. However, the discussion in our paper is more complete as even the BSM computation in Ref. [28] does not include all form factors since the  $V$ -mesons do not couple to scalar operators for instance.

<sup>3</sup>Notice the interchanged role of  $k$  and  $q$  with respect to the definition of the form factors in Eq. (2.1).

<sup>4</sup>By including the factor  $s_e$  in the definition of the operators  $Q_7, Q'_7$  we ensured that the sign of their Wilson coefficients is independent of the definition of the covariant derivative.

with

$$\begin{aligned} Q_7 &= \frac{s_e e}{16\pi^2} m_b \bar{b}_R \sigma^{\mu\nu} q_L F_{\mu\nu}, & Q'_7 &= \frac{s_e e}{16\pi^2} m_q \bar{b}_L \sigma^{\mu\nu} q_R F_{\mu\nu}, \\ Q_9 &= \frac{e^2}{16\pi^2} (\bar{b}_L \gamma^\mu q_L) (\bar{\ell} \gamma_\mu \ell), & Q'_9 &= \frac{e^2}{16\pi^2} (\bar{b}_R \gamma^\mu q_R) (\bar{\ell} \gamma_\mu \ell), \\ Q_{10} &= \frac{e^2}{16\pi^2} (\bar{b}_L \gamma^\mu q_L) (\bar{\ell} \gamma_\mu \gamma_5 \ell), & Q'_{10} &= \frac{e^2}{16\pi^2} (\bar{b}_R \gamma^\mu q_R) (\bar{\ell} \gamma_\mu \gamma_5 \ell). \end{aligned} \quad (2.7)$$

Given this Lagrangian, the amplitude for the  $\bar{B}_q(p_B) \rightarrow \ell^+(p_{\ell+}) \ell^-(p_{\ell-}) \gamma(k)$  is

$$\mathcal{A}_{\mu\mu\gamma} = -\frac{s_e e \alpha G_F}{2\sqrt{2}\pi m_{B_q}} V_{tb} V_{tq}^* \epsilon_\rho^*(k) \bar{u}^s(p_{\ell-}) \gamma_\mu \left( A_9^{\mu\rho} + A_{10}^{\mu\rho} \gamma_5 - \frac{2Q_\ell m_b m_{B_q}}{q^2} A_7^{\mu\rho} \right) v^r(p_{\ell+}), \quad (2.8)$$

where  $q \equiv p_B - k = p_{\ell+} + p_{\ell-}$  and we defined

$$\begin{aligned} A_7^{\mu\rho} &= (C_7 + \frac{m_q}{m_b} C'_7)(M_T^{\mu\rho}(q, k) + M_T^{\rho\mu}(k, q)) + (C_7 - \frac{m_q}{m_b} C'_7)(M_{T_5}^{\mu\rho}(q, k) + M_{T_5}^{\rho\mu}(k, q)), \\ A_9^{\mu\rho} &= (C_9 + C'_9) M_V^{\mu\rho}(q, k) - (C_9 - C'_9) M_A^{\mu\rho}(q, k), \\ A_{10}^{\mu\rho} &= (C_{10} + C'_{10}) M_V^{\mu\rho}(q, k) - (C_{10} - C'_{10}) M_A^{\mu\rho}(q, k). \end{aligned} \quad (2.9)$$

Above we omitted the contribution from photons radiated off final-state muons, because these are obtained from the  $B_q \rightarrow \mu\mu$  rates using PHOTOS, cf., Ref. [32]. Going slightly lower in  $q^2$  would necessitate the inclusion of broad charmonium resonances [33, 34]. For an overview of other non form-factor matrix elements see for instance Refs. [29, 33].

After integrating over the unobserved photon momentum, the differential rate for the radiative mode  $B_q \rightarrow \ell\ell\gamma$  reads

$$\frac{d\Gamma}{dq^2}(B_q \rightarrow \ell\ell\gamma) = \frac{\alpha^3 G_F^2 |\lambda_t|^2}{768\pi^4} \frac{\lambda_\gamma^{1/2} \lambda_{B_q}^{3/2}}{m_{B_q}^3 q^2} \left( c_A (|\mathcal{A}_{A_\perp}|^2 + |\mathcal{A}_{A_\parallel}|^2) + c_V (|\mathcal{A}_{V_\perp}|^2 + |\mathcal{A}_{V_\parallel}|^2) \right), \quad (2.10)$$

where  $\lambda_{B_q} \equiv \lambda(m_{B_q}^2, q^2, 0)$ ,  $c_V \equiv (q^2 + 2m_\ell^2)$ ,  $c_A \equiv (q^2 - 4m_\ell^2)$  and

$$\begin{aligned} \mathcal{A}_{V_{\perp,\parallel}} &\equiv \frac{1}{m_{B_q}} (C_9 \pm C'_9) V_{\perp,\parallel}^*(q^2, 0) + \frac{2m_b}{q^2} (C_7 \pm \frac{m_q}{m_b} C'_7) \bar{T}_{\perp,\parallel}(q^2), \\ \mathcal{A}_{A_{\perp,\parallel}} &\equiv \frac{1}{m_{B_q}} (C_{10} \pm C'_{10}) V_{\perp,\parallel}^*(q^2, 0), \end{aligned} \quad (2.11)$$

with the shorthands

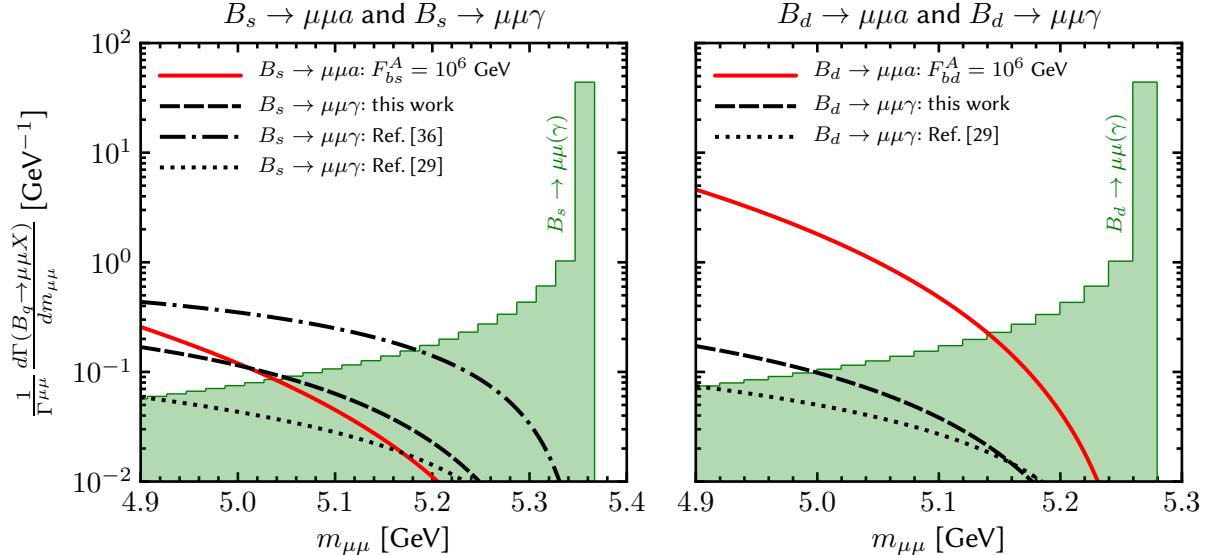
$$\begin{aligned} \bar{T}_\perp(q^2) &= T_\perp^*(q^2, 0) + T_\perp^*(0, q^2), \\ \bar{T}_\parallel(q^2) &= T_\parallel^*(q^2, 0) + T_\parallel^*(0, q^2)/(1 - q^2/m_{B_q}^2) = T_\parallel^*(q^2, 0) + T_\perp^*(0, q^2). \end{aligned} \quad (2.12)$$

The last equality relates  $T_\parallel^*(0, q^2)$  to  $T_\perp^*(0, q^2)$ , see Appendix A.1.5 and footnote 7 just before Eq. (A.1).

## 2.4 $B_q \rightarrow \mu\mu a$ and $B_q \rightarrow \mu\mu\gamma$ close to the kinematic endpoint

To illustrate the relative importance between the SM background  $B_q \rightarrow \mu\mu\gamma$  and the  $B_q \rightarrow \mu\mu a$  signal we take as a reference value for the flavour-violating coupling  $F_{bq}^A = 10^6$  GeV. In Figure 2, we show the differential rate normalised with respect to the two-body decay width

$$\frac{1}{\Gamma(B_q \rightarrow \mu\mu)} \frac{d\Gamma(B_q \rightarrow \mu\mu X)}{dm_{\mu\mu}},$$



**Figure 2:** Comparison of the axionic decay mode  $B_q \rightarrow \mu\mu a$  (red solid lines) and the radiative  $B_q \rightarrow \mu\mu\gamma$  modes (black lines). The left panel shows the  $B_s$  case while the right the  $B_d$  case. For the axion predictions  $F_{bg}^A = 10^6$  GeV is assumed as a reference value. The different black lines are the photon predictions with different form factor treatments (see legend and main text). In green are bins of the two-body  $B_q \rightarrow \mu\mu$  rate including radiation from final-state muons. To better compare the  $B_s$  and  $B_d$  cases, all rates are normalised to their respective two-body decay  $B_q \rightarrow \mu\mu$ , which is why the  $B_d \rightarrow \mu\mu a$  line appears enhanced with respect to the  $B_s \rightarrow \mu\mu a$  one.

where  $X = a, \gamma$ ,  $m_{\mu\mu}^2 \equiv q^2$ . In the left panel, we show the predictions for the  $B_s$  decays and in the right the corresponding ones for the  $B_d$  case. The binned (green) predictions are the  $B_q \rightarrow \mu\mu$  rates including photon radiation from the final-state muons using PHOTOS (see Ref. [32]). The red solid lines are the rates from the axion mode for the reference value  $F_{bg}^A = 10^6$  GeV (note that the relative enhancement between left and the right panel is due to the normalization, which carries a different CKM suppression.). The black lines are the  $B_q \rightarrow \mu\mu\gamma$  predictions when the photon does not originate from muon bremsstrahlung. They depend on the treatment of the non-perturbative input, i.e., the hadronic form factors introduced in Section 2.1. In all cases, we use the same perturbative input, namely the SM Wilson coefficients  $C_7^{\text{eff}}$ ,  $C_9^{\text{eff}}$  and  $C_{10}$  evaluated at the hadronic  $B_q$  scale. We obtain  $C_{10}$  from Ref. [32] and use `flavio` [35] to evaluate  $C_7^{\text{eff}}$  and  $C_9^{\text{eff}}$ .

We show the results of three different approaches of estimating the relevant hadronic form factors:

- **Dashed line:** the QCD sum rule form factor computation discussed in Section 2.1 and Appendix A,
- **Dotted line:** the quark-model approach of Ref. [29],
- **Dashed-dotted line:** the pole-dominance approach supplemented by experimental data and heavy-quark effective theory of Ref. [36]. It is specific to the  $B_s$  case (left panel).

The agreement of the predictions is rather crude. For  $q^2 \approx (4.9 \text{ GeV})^2$ , our prediction is about a factor of three larger than the quark model [29] and about a factor of two smaller than the pole-dominance approximation [36]. The disagreement with the quark model is not surprising as the method is designed for low  $q^2$  and, unlike in our work, no additional input is employed to constrain the residua of the leading poles near the kinematic endpoint. The agreement of the form factors themselves at lower  $q^2$ , which

we do not show, is much better. The comparison with the pole-dominance approach [36] has two major components. The difference in the  $B_q^*$ -residue and the fact that the effect of  $B_{q1}$ -resonance is neglected in Ref. [36] cf. Appendix A.4.2. While it is important to understand<sup>5</sup> the origin of the discrepancy in light of a possible measurement of the radiative decay, the discrepancy does not play a significant role in obtaining a bound on the axion couplings  $F_{bq}^A$ , which we derive in the next section.

### 3 Sensitivity at LHCb

In this section we recast the LHCb analysis of Ref. [23] to obtain an estimate for the current and future sensitivity of LHCb to probe the flavour-violating couplings  $F_{bs}^A$  and  $F_{bd}^A$ . We first discuss, in Section 3.1, how we extract the backgrounds by rescaling the original LHCb analysis, and derive the expected number of events in each bin for a given luminosity. We then describe, in Section 3.2, our statistical method and provide the recast of the present data and the sensitivity study for future runs. Our main results are summarised in Tables 2 and 3.

#### 3.1 Rescaling the LHCb analysis

The  $B_s \rightarrow \mu\mu$  analysis of LHCb in Ref. [23] makes use of datasets collected at different LHC runs, with luminosities  $\bar{\mathcal{L}}_7 = 1.0 \text{ fb}^{-1}$  from 7 TeV,  $\bar{\mathcal{L}}_8 = 2.0 \text{ fb}^{-1}$  from 8 TeV, and  $\bar{\mathcal{L}}_{13} = 1.4 \text{ fb}^{-1}$  from 13 TeV runs. Under the SM hypothesis, a total number of 62  $B_s \rightarrow \mu\mu$  events and 6.7  $B_d \rightarrow \mu\mu$  events are expected in this analysis in the full range of boosted-decision-trees (BDT) and the signal window ( $m_{\mu\mu} \in [5.2, 5.445] \text{ GeV}$ ). Since the BDT discrimination is flat one expects half of these events to pass the  $\text{BDT} > 0.5$  selection. For this BDT selection, LHCb supplies a plot with backgrounds, which we use to extract their numerical values. By combining the expected number of  $B_q \rightarrow \mu\mu$  events in the SM with the SM branching-fraction predictions, we extract a universal rescaling factor,  $r \simeq 0.079$ , via

$$\begin{aligned} N_{B_d} &= \underbrace{(\epsilon 2 f_d)}_{\equiv r} \times \overline{\text{BR}}_{[5.2 \text{ GeV} - 5.445 \text{ GeV}]}^{B_d \rightarrow \mu\mu(n\gamma)} \times \sum_{i=7,8,13} \sigma_{b,i} \bar{\mathcal{L}}_i, \\ N_{B_s} &= r \times \frac{f_s}{f_d} \times \overline{\text{BR}}_{[5.2 \text{ GeV} - 5.445 \text{ GeV}]}^{B_s \rightarrow \mu\mu(n\gamma)} \times \sum_{i=7,8,13} \sigma_{b,i} \bar{\mathcal{L}}_i. \end{aligned} \quad (3.1)$$

In these equations,  $i$  labels the  $\sqrt{s}$  run and  $\sigma_i$  is the corresponding  $b$ -quark production cross section in the acceptance of LHCb. The latter has been measured by LHCb for  $\sqrt{s} = 7, 13 \text{ TeV}$ ,  $\sigma_{b,7} = 72 \mu\text{b}$  and  $\sigma_{b,13} = 144 \mu\text{b}$  [39]. For  $\sigma_{b,8}$  we linearly rescale the 7 TeV value ( $\sigma_{b,8} = 8/7\sigma_{b,7}$ ).  $f_d$  and  $f_s$  are the fragmentation ratios of  $b$ -quarks that are produced at LHCb and fragment into  $B_d$  and  $B_s$ , respectively. We absorb  $f_d$  in the rescaling factor,  $r$ , and use the ratio  $f_s/f_d$  to obtain  $N_{B_s}$ . This ratio has been measured by the LHCb collaboration to be  $f_s/f_d = 0.259 \pm 0.015$  [40]. Finally,  $\epsilon$  summarises the experimental efficiencies and all other global rescaling factors, which we absorb into the definition of  $r$ .

The quantities  $\overline{\text{BR}}$ 's in Eq. (3.1) are the respective branching ratios in the signal window. This includes the effect of photon radiation from muons [32, 41], which LHCb simulates with PHOTOS. The overline in the branching-ratio prediction indicates that the partial width is divided by the width of the heavy mass eigenstate ( $\Gamma_{B_s}^H, \Gamma_{B_d}^H$ ) to obtain the branching fraction. In this way the effect of  $B_q$ -mixing is included [32, 42]. This is relevant for the  $B_s$  system, but much less so for the  $B_d$  system. This is numerically equivalent to LHCb's treatment of the effective lifetime, cf. Eq. (1) in Ref. [23]).

LHCb's  $\text{BDT} > 0.5$  selection covers the  $m_{\mu\mu} \in [4.9 \text{ GeV}, m_{B_s}]$  region in bins of 50 MeV. We apply the same universal rescaling factor,  $r$ , to rescale the predictions of all  $B_q \rightarrow \mu\mu a$  and  $B_q \rightarrow \mu\mu\gamma$  branching

<sup>5</sup> Whereas it will be challenging for lattice QCD to compute off-shell form factors, the on-shell ones have gained attention and computations are in progress [37, 38].

fractions for all  $m_{\mu\mu}$  bins. This is a good approximation as there are no triggers or similar thresholds that significantly change the rescaling over this invariant-mass range. In the next section, we present the sensitivity of this analysis to probe the flavour-violating  $F_{bs}^A$  and  $F_{bd}^A$  axion couplings in future runs of LHCb by rescaling the 13 TeV dataset. We denote the corresponding effective total luminosity by

$$\mathcal{L} = \bar{\mathcal{L}}_7 + \bar{\mathcal{L}}_8 + \mathcal{L}_{13}. \quad (3.2)$$

At a given total luminosity,  $\mathcal{L}$ , the expected number of events at a given  $m_{\mu\mu}$ -bin ( $\text{Bin}_k$ ) then is

$$\begin{aligned} N_k[F_{bs}^A, F_{bd}^A] &= N_{\text{Bin}_k}^{\text{BKG,analysis}} \frac{\text{SL}(\mathcal{L})}{\text{SL}(\bar{\mathcal{L}})} \\ &+ (\overline{\text{BR}}_{\text{Bin}_k}[B_d \rightarrow \mu\mu(n\gamma)] + \overline{\text{BR}}_{\text{Bin}_k}[B_d \rightarrow \mu\mu\gamma] + \overline{\text{BR}}_{\text{Bin}_k}[B_d \rightarrow \mu\mu a]) r \text{SL}(\mathcal{L}) \\ &+ (\overline{\text{BR}}_{\text{Bin}_k}[B_s \rightarrow \mu\mu(n\gamma)] + \overline{\text{BR}}_{\text{Bin}_k}[B_s \rightarrow \mu\mu\gamma] + \overline{\text{BR}}_{\text{Bin}_k}[B_s \rightarrow \mu\mu a]) r \frac{f_s}{f_d} \text{SL}(\mathcal{L}), \end{aligned} \quad (3.3)$$

with shorthands  $\bar{\mathcal{L}} \equiv \bar{\mathcal{L}}_7 + \bar{\mathcal{L}}_8 + \bar{\mathcal{L}}_{13} = 4.4 \text{ fb}^{-1}$  and  $\text{SL}(\mathcal{L}) \equiv \sigma_{b,7}\bar{\mathcal{L}}_7 + \sigma_{b,8}\bar{\mathcal{L}}_8 + \sigma_{b,13}(\mathcal{L} - \bar{\mathcal{L}}_7 - \bar{\mathcal{L}}_8)$ . The quantity  $N_{\text{Bin}_i}^{\text{BKG,analysis}}$  is the expected total number of background events that do not originate from the radiative decay in the given bin. We obtain  $N_{\text{Bin}_i}^{\text{BKG,analysis}}$  by digitising and integrating the plot of LHCb's  $\text{BDT} > 0.5$  selection. In Eq. (3.3) we kept separate the rate from photon emission from muons ( $B_q \rightarrow \mu\mu(n\gamma)$ ) and the rate from photon emissions from the initial state ( $B_q \rightarrow \mu\mu\gamma$ ). In principle, the amplitudes interfere but the interference is tiny close to the  $B_q$  threshold and we thus neglect it.

### 3.2 Recast and sensitivity analysis

To compute the sensitivity of the LHCb analysis in probing  $F_{bs}^A$  and  $F_{bd}^A$ , we must combine the information of all  $m_{\mu\mu}$  bins and include statistical and systematic uncertainties. We neglect the subdominant experimental systematic uncertainties but will include the theory uncertainties associated to the form factors entering the three-body rates. In what follows we always either turn on  $F_{bs}^A$  or  $F_{bd}^A$ , i.e., but will not let them float simultaneously.

Each  $m_{\mu\mu}$  bin corresponds to an independent counting experiment that obeys Poisson statistics. Exclusion limits on  $F_{bq}^A$  are then obtained from a joined Poisson (Log)Likelihood. For a sufficiently large number of events, Poisson statistics are well described by Gaussian statistics and the Poisson (Log)Likelihood is equivalent to a  $\chi^2$  function of the NP parameter, i.e.,  $F_{bq}^A$ :

$$\chi^2(F_{bq}^A) = \sum_{i,j} (N_i - N_i^{\text{obs}})(V_{\text{cov}}^{-1})_{ij}(N_j - N_j^{\text{obs}}), \quad (3.4)$$

with  $i$  numbering the bins and  $q = s, d$ .  $N_i = N_i(F_{bq}^A)$  denotes the total number of events (background plus signal) for the value  $F_{bq}^A$  in a given bin, whereas  $N_i^{\text{obs}}$  is the observed number of events. For the recast we use the actual number of events observed by LHCb, read off from Figure 1 in Ref. [23]. To project the sensitivity for future LHCb runs we set  $N_i^{\text{obs}}$  to the number of events expected in the SM. The covariance matrix,  $V_{\text{cov}}$ , incorporates statistical and systematic uncertainties in a way that we discuss below. If we neglect systematic uncertainties, this matrix is diagonal and only contains the squared Poisson variances,  $V_{\text{cov}} = V_{\text{stat}}$  with  $(V_{\text{stat}})_{ij} = \delta_{ij}N_i$ . We have explicitly checked, that for the data samples considered here, the Poisson (Log-)Likelihood is always very well approximated by the  $\chi^2$ .

To incorporate systematic/theory uncertainties we follow the commonly used approach of Ref. [43]. Theory uncertainties are then treated as Gaussian uncertainties smearing the expectation values of the underlying Poisson probability distribution functions. We can then obtain the limits on  $F_{bq}^A$  by generating

	$B_s \rightarrow \mu\mu a$	$B_d \rightarrow \mu\mu a$		
	sys+stat	stat only	sys+stat	stat only
$ F_{bq,\text{best-fit}}^A  \times 10^{-5} [\text{GeV}]$	$\chi_{\min}^2$	15.0	15.0	14.6
$ F_{bq,90\%}^A  \times 10^{-5} [\text{GeV}]$		3.8	3.8	4.5
		$> 2.2$	$> 2.3$	$> 2.8$
				$> 2.9$

**Table 2:** The results of recasting LHCb’s analysis [23] to test flavour-violating couplings of the axion to  $B_s$  ( $F_{bs}^A$ ) and  $B_d$  ( $F_{bd}^A$ ). The analysis employs a total of  $4.4 \text{ fb}^{-1}$  of data from runs at 7, 8, and 13 TeV. In the columns labelled “sys+stat” we combine statistical and theory uncertainties, while in the columns labelled “stat only” we neglect the latter. We see that presently the bounds are dominated by statistical uncertainties. When computing the  $\chi^2$  we sum over the ten first bins of the analysis, i.e.,  $m_{\mu^+\mu^-} \in [4.9 \text{ GeV}, m_{B_s}]$ . For every case we list the values of the  $\chi_{\min}^2$  and the corresponding best-fit value for  $|F_{bq}^A|$ . The values of  $\chi_{\min}^2$  should be compared with the  $\chi^2$  value of the SM,  $\chi_{\text{SM}}^2 = 15.7$ . The axion best-fit values are thus in roughly  $1\sigma$  agreement with the SM.  $|F_{bq,90\%}^A|$  are the resulting 90% CL exclusion limits.

Monte-Carlo events based on the joined Poisson likelihood after smearing the expectation values by the (correlated) systematic errors. If the measurement is well-described by Gaussian statistics (as in our case) and the systematic uncertainties are small with respect to the statistical ones, this treatment of uncertainties is equivalent to adding the statistical and systematic errors in quadrature in  $V_{\text{cov}}$ .

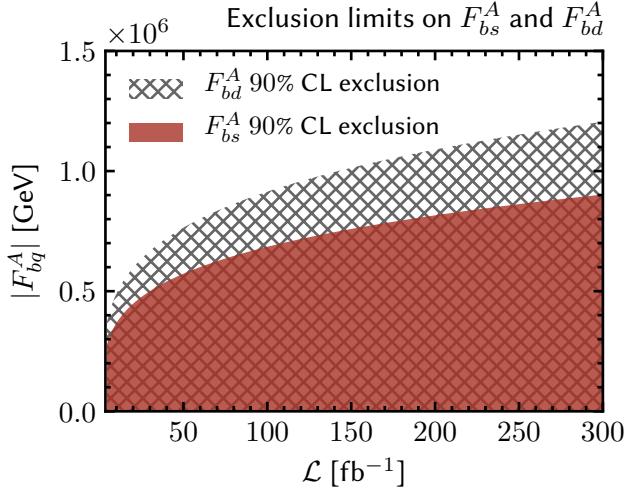
In our case the main systematic uncertainties are due to the form factors that enter the radiative  $B_q \rightarrow \mu\mu\gamma$  and the  $B_q \rightarrow \mu\mu a$  rate. Since the uncertainties in the form factors originate in part from uncertainties in input parameters like  $m_b$  and  $\langle \bar{q}q \rangle$  that are  $q^2$ -independent, the predicted number of events among different bins are correlated. Therefore, the full covariance matrix for the case in which the axion has a coupling  $F_{bq}^A$  is not diagonal and decomposes into

$$V_{\text{cov}} = V_{\text{stat}} + V_\gamma + \frac{1}{(F_{bq}^A)^4} V_a^q + \frac{1}{(F_{bq}^A)^2} V_{a-\gamma}^q. \quad (3.5)$$

Here,  $(V_{\text{stat}})_{ij} = \delta_{ij} N_i$  are the statistical uncertainties, while the matrices  $V_\gamma$ ,  $V_a^q$ , and  $V_{a-\gamma}^q$  describe the correlated errors among the predictions of various rates over the bins. Aside from trivial functional dependencies on global rescaling factors, e.g., luminosity, we can determine them once and for all by generating Monte-Carlo events in which we vary the parameters on which the form factors depend. In practice we use the mean values of the  $z$ -expansion fit (of degree four) and their covariance matrix (see Appendix A.4.3) to determine each piece of  $V_{\text{cov}}$ . Using the covariance matrices we obtain the 90% Confidence Level (CL) exclusion limit on  $|F_{bq}^A|$ , i.e.  $\chi^2(F_{bq,90\%}^A) - \chi_{\min}^2 = 1.64$ .

First, we recast the observed data of LHCb’s analysis [23] in which  $\mathcal{L} = \bar{\mathcal{L}} = 4.4 \text{ fb}^{-1}$ . The measurement is dominated by statistical uncertainties, but for purposes of illustration we show both the bounds when combining statistical and systematic theory errors and the bounds when only the statistical uncertainty is included. In the  $\chi^2$  we include the first ten bins of the LHCb analysis. The observed data are in good agreement with the SM expectation. Indeed, we find that the  $\chi^2$  of the SM divided by the ten degrees of freedom of the  $\chi^2$  (d.o.f.) is  $\chi_{\text{SM}}^2/\text{d.o.f.} = 1.6$ . The best-fit points for the axion lies roughly  $1\sigma$  off the SM. In Table 2 we list the best-fit points with their corresponding  $\chi_{\min}^2$ , as well as the resulting 90% CL exclusion limits on  $|F_{bs}^A|$  and  $|F_{bd}^A|$ .

Next we make projections for future runs of LHCb. As discussed in Section 3.1, to this end we rescale the 13 TeV events assuming LHCb will collect a total of  $300 \text{ fb}^{-1}$ . To compute the sensitivity we assume that LHCb will observe exactly the number of events expected from the SM. Therefore, the best-fit point always corresponds to observing zero events from axion decays and  $\chi_{\min}^2 = 0$ . For the projection study



**Figure 3:** Projected sensitivity of LHCb to probe the flavour-violating axion couplings  $F_{bs}^A$  (filled red region) and  $F_{bd}^A$  (hatched region) as a function of the total integrated luminosity. Shown are the 90% CL exclusion limits assuming that the observed number of events will be the same as predicted in the SM hypothesis.

$\mathcal{L}$ [fb $^{-1}$ ]	$B_s \rightarrow \mu\mu a$		$B_d \rightarrow \mu\mu a$	
	$ F_{bs}^A  \times 10^{-5}$ [GeV] sys+stat	$ F_{bs}^A  \times 10^{-5}$ [GeV] stat only	$ F_{bd}^A  \times 10^{-5}$ [GeV] sys+stat	$ F_{bd}^A  \times 10^{-5}$ [GeV] stat only
10	3.7	3.7	4.9	5.0
30	5.0	5.1	6.7	6.8
50	5.7	5.8	7.6	7.8
100	6.8	6.9	9.1	9.3
300	9.0	9.2	12	12

**Table 3:** Projected 90% CL exclusion limits on the flavour-violating couplings of the axion to  $B_s$  ( $F_{bs}^A$ ) and  $B_d$  ( $F_{bd}^A$ ) as a function of the integrated luminosity at LHCb. In the columns labelled “sys+stat” we combine statistical and theory uncertainties, while in the columns labelled “stat only” we neglect the latter.

we present the results both when only statistical uncertainties are considered and when they are folded with the correlated theory uncertainties. In Figure 3 we show the resulting 90% CL exclusion limit on  $|F_{bs}^A|$  (left panel) and  $|F_{bd}^A|$  (right panel) as a function of the total luminosity. In addition, the limits for some indicative luminosities are listed in Table 3.

Note that the limit from the actual recast is weaker than the expected limit under the background-only hypothesis. More precisely, if we consider the case  $\mathcal{L} = 4.4 \text{ fb}^{-1}$  and set  $N_i^{\text{obs}} = N_i^{\text{SM}}$  (as we do for the projection study) we find for the statistics-only case  $|F_{bs,90\%}^A| < 2.7 \cdot 10^5 \text{ GeV}$  and  $|F_{bd,90\%}^A| < 3.5 \cdot 10^5 \text{ GeV}$ . In comparison, the corresponding exclusion limits of the recast (table 3) are slightly weaker. The origin of this difference is mainly an excess of roughly 10 events in the first bin of the current LHCb  $B_s \rightarrow \mu\mu$  analysis, which can be fitted by the best-fit point of an axion signal. However, as discussed in the recast the excess is not statistically significant and the best-fit point of the axion is within  $1\sigma$  of the SM.

## 4 Summary and Outlook

In this article we have proposed a novel method to probe flavour-violating couplings of the QCD axion to  $b$ -quarks at the LHC, exploiting the di-muon datasets collected for the  $B_q \rightarrow \mu\mu$  analyses. To this end, we have computed the relevant differential decay rates for the decay of a  $B_q$ -meson to muons and an axion  $B_q \rightarrow \mu\mu a$  [Eq. (2.6)] and the radiative decay  $B_q \rightarrow \mu\mu\gamma$  [Eq. (2.10)], which is a background to the former process.

These rates depend on non-perturbative  $B_q \rightarrow \gamma^{(*)}$  form factors, which we have discussed from a general viewpoint, computed with QCD sum rules (at zero flavour-violating momentum transfer), and fitted to a  $z$ -expansion in Appendices A.1, A.3 and A.4, respectively. To the best of our knowledge this is the first discussion of the complete set of form factors, for the dimension-six effective Hamiltonian  $\mathcal{H}_{\text{eff}}^{b \rightarrow (d,s)}$ , supplemented with an explicit computation of all form factors. Besides being useful for axion searches these form factors are also the ingredients for other light BSM particle (e.g. dark photon) searches. In addition, we have exposed the relation between the introduced basis and the standard  $B \rightarrow V$  basis through the dispersion representation in Appendix A.2, which interrelates form-factor properties of the two bases.

With these decay rates we performed a recast using available LHCb data and estimated the sensitivity to  $B_q \rightarrow \mu\mu a$  at present and future runs, taking into account the SM background  $B_q \rightarrow \mu\mu\gamma$ . We find that present data constrain the relevant axion couplings  $F_{bd}^A$  ( $F_{bs}^A$ ) to be larger than  $2.8$  ( $2.2$ )  $\cdot 10^5$  GeV at 90% CL [Table 2], while the full LHCb dataset will probe scales of the order of  $10^6$  GeV in both  $b \rightarrow d$  and  $b \rightarrow s$  transitions [Table 3].

For stable axions, these results should be compared with the ones derived from  $B$ -meson decays with missing energy. In the case of  $b \rightarrow s$  transitions, the data from the BaBar collaboration on  $B \rightarrow K^*\nu\bar{\nu}$  provide constraints that are roughly two orders of magnitude stronger than the ones from our LHCb recast of  $B_s \rightarrow \mu\mu a$ , cf. Table 1. For the case of  $b \rightarrow d$  transitions, the BaBar constraints are roughly of the same order than the ones that LHCb can obtain in upcoming runs. Nevertheless, the combination with the corresponding ATLAS and CMS analyses of  $B_q \rightarrow \mu\mu$  may improve the bounds significantly.

While it is remarkable that the LHC can play a role in constraining couplings of the QCD axion, the analysis of  $B_q \rightarrow \mu\mu a$  that we have presented here can be relevant for other extensions of the SM with light neutral particles with flavour-violating couplings. Since the  $B_q \rightarrow \mu\mu a$  analysis is inclusive, it can be extended to search for light BSM particles even if they decay within the detector. For example, an ALP that decays promptly to, for instance, photons may be subject to cuts on additional photons in the analyses of  $B \rightarrow K(a \rightarrow \gamma\gamma)$  at the  $B$ -factories and thus evade detection, while it would be kept in the  $B_q \rightarrow \mu\mu(a \rightarrow \gamma\gamma)$  samples at the LHC. Therefore, the analysis that we have presented here complements axion searches in rare meson decays with missing energy at  $B$ -factories, and can play an important role in constraining flavour-violating couplings of light particles.

## Acknowledgments

We are grateful to Tadeusz Janowski for providing the  $z$ -expansion fits to the form factors. We thank Ben Pullin, Mikolai Misiak, and Uli Nierste for very useful discussions. This research was supported by the Munich Institute for Astro- and Particle Physics (MIAPP) of the DFG Excellence Cluster Origins ([www.origins-cluster.de](http://www.origins-cluster.de)). E. Stamou and R. Ziegler thank the Galileo Galilei Institute for Theoretical Physics for the hospitality and the INFN for partial support during the initial stages of this work. J. Albrecht gratefully acknowledges support of the European Research Council, ERC Starting Grant: PRECISION 714536. R. Zwicky is supported by an STFC Consolidated Grant, ST/P0000630/1.

## A The $B_q \rightarrow \gamma^*$ Form Factors

The standard  $B_q \rightarrow V$  matrix elements (ME), where  $V = \rho^0, \omega, \phi \dots$  is a vector meson, hold some analogy with the  $B_q \rightarrow \gamma^*$  ones. However, the difference is that the analogue of the vector meson mass is the photon off-shell momentum which is a variable rather than a constant. Hence the MEs are functions of two variables and this leads to a more involved analytic structure. In this paper we restricted ourselves to the kinematic region  $q^2 \in [(4.9 \text{ GeV})^2, m_{B_q}^2]$ , where the form factors (FFs) can be expected to dominate over long-distance contributions.

This appendix is structured as follows. Firstly, we define and state relation and limits of the FFs in Section A.1, the link with the  $B \rightarrow V$  basis is discussed in Section A.2, the QCD sum rule computation of the off-shell FFs follows in Section A.3 and finally we turn to the FF-parametrisation and fits in Section A.4. Note, that sections A.1, A.2 and A.4 are independent of the method of computation. In an updated version we plan to provide the  $z$ -expansion data as ancillary files including plots of FFs.

Form Factor	$P^*(0, k^2)$	$T_{  }^*(0, k^2) = T_{\perp}^*(1 - \hat{k}^2)(0, k^2)$	$V, T_{\perp,   }(q^2) = V^*, T_{\perp,   }^*(q^2, 0)$
Mode	$B_s \rightarrow \ell\ell a$	$B_q \rightarrow \ell\ell\gamma$	$B_s \rightarrow \ell\ell\gamma$
Poles $q^2, k^2$	$m_\phi^2, m_Y^2$	$m_\phi^2, m_Y^2$	$m_{B_q^*}^2, m_{B_{1q}}^2$
Defined in Eqs.	(2.1,A.1)	(2.1,A.1)	(A.23)
Graph in Figure 1	(left)	(centre)	(right)
Other notation	—	$F_{TV}(0, k^2) = F_{TA}(0, k^2)$ [29, 44]	$F_{V,A}, F_{TV,A}(q^2, 0)$ [29, 44]

**Table 4:** Overview of FFs referencing definitions, graphs, and analytic structure. The latter defines the region of validity of the computation. Long-distance contributions are relevant in other kinematic regions [29, 33]. For  $B_d \rightarrow \gamma^*$ ,  $m_\phi^2$  is to be replaced by  $m_{\rho,\omega}^2$  above.

### A.1 Definition of $B \rightarrow \gamma^{(*)}$ form factors

We introduce a complete set of *off-shell* FFs which is related to the standard  $B \rightarrow V$  basis [45, 46] via dispersion relations, cf. Section A.2. On a technical level this appendix extends previous work [29, 44], in that we discuss the full set of seven vector and tensor FFs and not only those needed for the SM transition. The complete basis is for example useful for other invisible particle searches such as the dark photon. The off-shell FFs are not to be confused with the on-shell FFs which have received more attention in the literature [29, 30, 44, 47]. An overview of the on- and off-shell FFs used for this paper are shown in the diagrams in Figure 1 and contrasted in Table 4.

#### A.1.1 The complete basis of seven off-shell form factors $F^*(q^2, k^2)$

We introduce the FFs with two momentum squares  $q^2$  and  $k^2$  collectively as  $F^*(q^2, k^2) \equiv F^{B \rightarrow \gamma^*}(q^2, k^2)$ . The first argument (here  $q^2$ ) denotes the momentum transfer at flavour-violating vertex while the second argument (here  $k^2$ ) denotes the momentum of the photon emitted at low energies.

We introduce a new off-shell basis via a dispersion representation based on the standard  $B \rightarrow V$  basis [45, 46]. Below we state the basis before turning to the construction in Section A.2. The absence of unphysical singularities in the matrix element enforces relations between FFs which we discuss in some detail. We will refer to this circumstance as “regularity” for short.

The complete set of FFs were already introduced in the main text in Eq. (2.1) and reproduced here for

convenience<sup>6,7</sup>

$$\begin{aligned}
M_5^\rho &\equiv b_P \langle \gamma^*(k, \rho) | \bar{q} \gamma_5 b | \bar{B}_q(p_B) \rangle &= i m_{B_q} R^\rho P^*(q^2, k^2) \\
M_V^{\mu\rho} &\equiv b_V \langle \gamma^*(k, \rho) | \bar{q} \gamma^\mu b | \bar{B}_q(p_B) \rangle &= + R_\perp^{\mu\rho} V_\perp^*(q^2, k^2), \\
M_A^{\mu\rho} &\equiv b_V \langle \gamma^*(k, \rho) | \bar{q} \gamma^\mu \gamma_5 b | \bar{B}_q(p_B) \rangle &= + (R_\parallel^{\mu\rho} V_\parallel^*(q^2, k^2) + R_{\mathbb{L}}^{\mu\rho} V_{\mathbb{L}}^*(q^2, k^2) + R_P^{\mu\rho} V_P^*(q^2, k^2)), \\
M_T^{\mu\rho} &\equiv b_T \langle \gamma^*(k, \rho) | \bar{q} i q_\nu \sigma^{\mu\nu} b | \bar{B}_q(p_B) \rangle &= + R_\perp^{\mu\rho} T_\perp^*(q^2, k^2), \\
M_{T_5}^{\mu\rho} &\equiv b_T \langle \gamma^*(k, \rho) | \bar{q} i q_\nu \sigma^{\mu\nu} \gamma_5 b | \bar{B}_q(p_B) \rangle &= - (R_\parallel^{\mu\rho} T_\parallel^*(q^2, k^2) + R_{\mathbb{L}}^{\mu\rho} T_{\mathbb{L}}^*(q^2, k^2)), 
\end{aligned} \tag{A.1}$$

where  $b_P \equiv \left(\frac{m_b+m_q}{s_e e}\right)$ ,  $b_V \equiv \left(-\frac{m_{B_q}}{s_e e}\right)$ ,  $b_T \equiv \left(\frac{1}{s_e e}\right)$ , the momentum transfer is  $q \equiv p_B - k$  and the off-shell photon state  $\langle \gamma^*(k, \rho) |$  is defined through

$$\langle \gamma^*(k, \rho) | O(0) | B \rangle \equiv -i e s_e \int d^4 x e^{ik \cdot x} \langle 0 | T j^\rho(x) O(0) | B \rangle, \tag{A.2}$$

where  $j^\rho = \sum_f Q_f \bar{f} \gamma^\rho f$  is the electromagnetic current.

$$\begin{aligned}
R_\perp^{\mu\rho} &\equiv \varepsilon^{\mu\rho\beta\gamma} q_\beta k_\gamma, & R_\parallel^{\mu\rho} &\equiv \frac{i}{2} (1 - \hat{q}^2) (m_{B_q}^2 G^{\mu\rho} - \frac{(q+2k)^\mu R^\rho}{1 - \hat{k}^2}), \\
R_{\mathbb{L}}^{\mu\rho} &\equiv \frac{i}{2} (q^\mu - \frac{\hat{q}^2 (q+2k)^\mu}{1 - \hat{k}^2}) R^\rho, & R_P^{\mu\rho} &\equiv \frac{i}{2} q^\mu R^\rho, & R^\rho &\equiv q^\rho - \frac{k \cdot q}{k^2} k^\rho, 
\end{aligned} \tag{A.3}$$

are Lorentz tensors with convenient properties (cf. below) and hereafter

$$\hat{k}^2 \equiv \frac{k^2}{m_{B_q}^2}, \quad \hat{q}^2 \equiv \frac{q^2}{m_{B_q}^2}. \tag{A.4}$$

The photon transverse tensor,  $k^\alpha G_{\alpha\beta} = 0$ , is

$$G_{\alpha\beta} \equiv g_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2}, \tag{A.5}$$

and it is noted that  $R^\rho = q_\mu G^{\mu\rho}$  (A.3).

The QED Ward identity holds off-shell in the form

$$k_\rho M_{V,A,T,T_5}^{\mu\rho} = 0, \tag{A.6}$$

without contact term since the weak operator is neutral in the total electric charge. Note that Eq. (A.6) is automatically satisfied in our parametrisation since  $k_\rho R_{\perp,\parallel,\mathbb{L},P}^{\mu\rho} = 0$ . The non-singlet axial Ward identity for  $M_A^{\mu\rho}$  assumes the form

$$q_\mu M_A^{\mu\rho} = m_{B_q} M_5^\rho, \tag{A.7}$$

---

<sup>6</sup>The conventions are  $\gamma_5 = i \gamma^0 \gamma^1 \gamma^2 \gamma^3$ ,  $g = \text{diag}(1, -1, -1, -1)$ ,  $\langle 0 | \bar{q} \gamma^\mu \gamma_5 b | \bar{B}_q(p_B) \rangle = i p_B^\mu f_{B_q}$ ,  $D_\mu = \partial_\mu + s_e i Q_f e A_\mu$  and  $\varepsilon_{0123} = 1$ . These conventions fix the phase of the  $B_s$ - and the  $\gamma$ -state together with  $\langle \gamma | A_\mu | 0 \rangle = \epsilon_\mu^*$ . For  $s_e = 1$  these phase conventions render the  $B_q \rightarrow \gamma$  FFs positive.

<sup>7</sup>Whereas  $M_{T_5}^{\rho\mu}$  in [Eq. 4] in [33], and similarly in [29], is incomplete it remains sufficient within the SM as there  $\epsilon_\mu^*(q) M_{T_5}^{\rho\mu}$  and  $q^2 \rightarrow 0$  annihilate the  $T_{\mathbb{L}}$ -contribution. However, the correct substitution reads  $T_\parallel(0, q^2)|_{[33]} \rightarrow T_\parallel(0, q^2)/(1 - \hat{q}^2)$  since the normalisation differs slightly.

which in turn holds without contact term since the electromagnetic current is invariant under non-singlet axial rotations. Eq. (A.7), upon using  $q_\mu R_{\perp,\parallel,\mathbb{L}}^{\mu\rho} = 0$ , implies that

$$V_P^*(q^2, k^2) = \frac{2}{\hat{q}^2} P^*(q^2, k^2). \quad (\text{A.8})$$

Regularity enforces constraints on the FFs defined in (A.1).<sup>8</sup> There are two constraints at  $q^2 = 0$  and  $k^2 = m_{B_q}^2$  respectively. The Ward identity (A.8) enforces

$$P^*(0, k^2) = \hat{V}_{\mathbb{L}}^*(0, k^2), \quad (\text{A.9})$$

where  $\hat{V}_{\mathbb{L}}^*$  is implicitly defined by

$$V_{\mathbb{L}}^*(q^2, k^2) \equiv -\frac{2}{\hat{q}^2} \hat{V}_{\mathbb{L}}^*(q^2, k^2). \quad (\text{A.10})$$

The second constraint is

$$T_{\parallel}^*(0, k^2) = (1 - \hat{k}^2) T_{\perp}^*(0, k^2). \quad (\text{A.11})$$

The two constraints at  $k^2 = m_{B_q}^2$  are

$$\begin{aligned} (1 - \hat{q}^2) V_{\parallel}^*(q^2, m_{B_q}^2) + \hat{q}^2 V_{\mathbb{L}}^*(q^2, m_{B_q}^2) &= 0, \\ (1 - \hat{q}^2) T_{\parallel}^*(q^2, m_{B_q}^2) + \hat{q}^2 T_{\mathbb{L}}^*(q^2, m_{B_q}^2) &= 0. \end{aligned} \quad (\text{A.12})$$

Whereas the constraints (A.9) and (A.12) are merely imposed on by the FF-parametrisation, (A.11) is of mostly algebraic origin cf. Section A.1.5 for the derivation.

### A.1.2 The four photon on-shell form factors $F(q^2) \equiv F^*(q^2, 0)$

We next turn to the case where the low-energy photon is on-shell;  $k^2 = 0$ . We introduce the commonly used shorthand

$$F(q^2) \equiv F^*(q^2, 0), \quad \text{for } F \in \{P, V_{\perp,\parallel,\mathbb{L}}, T_{\perp,\parallel,\mathbb{L}}\}, \quad (\text{A.13})$$

(or  $F^{B \rightarrow \gamma}(q^2) \equiv F^{B \rightarrow \gamma^*}(q^2, 0)$ ). The basic physics idea is that the absence of the photon's zero helicity component implies the vanishing the pseudoscalar FF and the zero helicity part of the vector FFs. We may define the helicity amplitude by

$$\mathcal{A}_{\lambda\lambda'}^X = M_X^{\rho\mu} \epsilon_{\rho}^*(k, \lambda) \epsilon_{\mu}(q, \lambda'), \quad (\text{A.14})$$

and then

$$\lim_{k^2 \rightarrow 0} \mathcal{A}_{00}^A \propto V_{\parallel} - V_{\mathbb{L}}, \quad \lim_{k^2 \rightarrow 0} \mathcal{A}_{0t}^A \propto P, \quad \lim_{k^2 \rightarrow 0} \mathcal{A}_{00}^{T_5} \propto T_{\parallel} - T_{\mathbb{L}}, \quad (\text{A.15})$$

which can be derived using the explicit parametrisation

$$\begin{aligned} k &= (\sqrt{k^2 + v^2}, 0, 0, v), & \epsilon^*(k, 0) &= (v, 0, 0\sqrt{k^2 + v^2})/\sqrt{k^2}, \\ q &= (\sqrt{q^2 + v^2}, 0, 0, -v), & \epsilon^*(q, 0) &= (-v, 0, 0\sqrt{q^2 + v^2})/\sqrt{q^2}, \end{aligned} \quad (\text{A.16})$$

---

<sup>8</sup>The two constraints (A.9,A.11) have well-known analogues in  $B \rightarrow V$  which are stated in Section A.2. A similar constraint to (A.11) was reported in Ref. [29] and we comment in the same section in what way it differs from ours.

in the  $B_q$ -meson restframe and  $v \equiv |\vec{k}| = \lambda^{1/2}(m_{B_q}^2, q^2, k^2)/(2m_{B_q})$ . Second, in the limit  $k^2 \rightarrow 0$ ,  $\epsilon^*(k, 0) \propto k$ , and this enforces,

$$\lim_{k^2 \rightarrow 0} \epsilon_\rho^*(k, 0) M_{A, T_5}^{\mu\rho} = 0 , \quad (\text{A.17})$$

since it is equivalent to the QED Ward identity (A.6). Eqs. (A.15,A.17) lead to the following constraints

$$V_{\parallel}(q^2) = V_{\mathbb{L}}(q^2) , \quad T_{\parallel}(q^2) = T_{\mathbb{L}}(q^2) , \quad P(q^2) = 0 , \quad (\text{A.18})$$

and reduces the seven FFs of Eq. (A.1) to four. Alternatively one can infer the constraints (A.18) from the regularity of the matrix elements as  $k^2 \rightarrow 0$ . The regularity condition and the helicity arguments are clearly related as one would expect.

For completeness we give the explicit  $k^2 \rightarrow 0$  basis [30, 33]<sup>9</sup>

$$\begin{aligned} M_V^\mu &\equiv b_V \langle \gamma(\epsilon(k)) | \bar{q} \gamma^\mu b | \bar{B}_q(p_B) \rangle &= +P_\perp^\mu V_\perp(q^2) , \\ M_A^\mu &\equiv b_V \langle \gamma(\epsilon(k)) | \bar{q} \gamma^\mu \gamma_5 b | \bar{B}_q(p_B) \rangle &= +P_\parallel^\mu V_\parallel(q^2) , \\ M_T^\mu &\equiv b_T \langle \gamma(\epsilon(k)) | \bar{q} i q_\nu \sigma^{\mu\nu} b | \bar{B}_q(p_B) \rangle &= +P_\perp^\mu T_\perp(q^2) , \\ M_{T_5}^\mu &\equiv b_T \langle \gamma(\epsilon(k)) | \bar{q} i q_\nu \sigma^{\mu\nu} \gamma_5 b | \bar{B}_q(p_B) \rangle &= -P_\parallel^\mu T_\parallel(q^2) , \end{aligned} \quad (\text{A.19})$$

where

$$P_\perp^\mu \equiv \varepsilon^{\mu\rho\beta\gamma} \epsilon_\rho^* q_\beta k_\gamma \quad P_\parallel^\mu \equiv i (q \cdot k \epsilon^*{}^\mu - q \cdot \epsilon^* k^\mu) , \quad (\text{A.20})$$

and are related to the  $R$ -tensors by

$$P_\perp^\mu = \epsilon_\rho^* R_\perp^{\mu\rho} , \quad P_\parallel^\mu = \epsilon_\rho^* (R_\parallel^{\mu\rho} + R_{\mathbb{L}}^{\mu\rho}) . \quad (\text{A.21})$$

**Relation to the standard  $B \rightarrow V$  basis** We consider it worthwhile to comment on some aspects in the standard basis of  $B \rightarrow V$  FFs e.g. [46]. The  $k^2 \rightarrow 0$  limit is then akin to  $m_V \rightarrow 0$ . The relations  $V_{\parallel}(q^2) - V_{\mathbb{L}}(q^2) = T_{\parallel}(q^2) - T_{\mathbb{L}}(q^2) = 0$  implies

$$\begin{aligned} V_2^{B \rightarrow V}(q^2) &= (1 - \hat{q}^2) V_3^{B \rightarrow V}(q^2) + O(m_V) , \\ T_2^{B \rightarrow V}(q^2) &= (1 - \hat{q}^2) T_3^{B \rightarrow V}(q^2) + O(m_V) . \end{aligned} \quad (\text{A.22})$$

Such relations were noted previously. Firstly, in the  $B \rightarrow V$  context in Ref. [48] in Appendix A and around Eq. [5] in Ref. [28], where it is argued that the relation has to hold in order to cancel a kinematic  $1/m_V$ -factor. Second for  $B \rightarrow \gamma$  ( $m_V = 0$ ) they were previously reported in Ref. [44] as a consequence of regularity.

### A.1.3 The five form factors $F^*(0, k^2)$ at zero flavour-violating momentum transfer

In the process  $B_q \rightarrow \ell\ell X$ , with  $X$  a light BSM particle, the limit in which the flavour-violating momentum transfer goes to zero, i.e.,  $q^2 = 0$ , corresponds to the case of zero or small mass of  $X$ . In this limit the two constraints in Eqs. (A.9) and (A.11) reduce the number of independent FFs from seven to five.

---

<sup>9</sup> The charged FF  $B_u \rightarrow \gamma^{(*)}$  is similar but comes with a non gauge invariant contact term for the axial vector structure. This contact term is canceled by the photon emission of the lepton [30].

The matrix elements, at  $q^2 = 0$ , read

$$\begin{aligned}
M_V^{\mu\rho} \equiv b_V \langle \gamma^*(k, \rho) | \bar{q} \gamma^\mu b | \bar{B}_q(p_B) \rangle &= + R_\perp^{\mu\rho} V_\perp^*(0, k^2), \\
M_A^{\mu\rho} \equiv b_V \langle \gamma^*(k, \rho) | \bar{q} \gamma^\mu \gamma_5 b | \bar{B}_q(p_B) \rangle &= + (R_\parallel^{\mu\rho} V_\parallel^*(0, k^2) + i \frac{(2k+q)^\mu}{1-\hat{k}^2} R^\rho P^*(0, k^2)), \\
M_T^{\mu\rho} \equiv b_T \langle \gamma^*(k, \rho) | \bar{q} i q_\nu \sigma^{\mu\nu} b | \bar{B}_q(p_B) \rangle &= + R_\perp^{\mu\rho} T_\perp^*(0, k^2), \\
M_{T_5}^{\mu\rho} \equiv b_T \langle \gamma^*(k, \rho) | \bar{q} i q_\nu \sigma^{\mu\nu} \gamma_5 b | \bar{B}_q(p_B) \rangle &= - (R_\parallel^{\mu\rho} (1-\hat{k}^2) T_\perp^*(0, k^2) + \frac{i}{2} q^\mu R^\rho T_\perp^*(0, k^2)), \quad (\text{A.23})
\end{aligned}$$

where  $P^*(0, k^2) = \hat{V}_\perp^*(0, k^2)$  (A.9), and  $2k \cdot q|_{q^2=0} = m_{B_q}^2 - k^2$  have been used. At  $q^2 = 0$  the constraints (A.12) imply

$$V_\parallel^*(0, m_{B_q}^2) = 2P^*(0, m_{B_q}^2), \quad T_\parallel^*(0, m_{B_q}^2) = 0. \quad (\text{A.24})$$

With  $T_\perp^*(0, m_{B_q}^2)$  finite the last constraint is obeyed trivially by (A.11).

#### A.1.4 Counting form factors

type \ $J^P$	#	$1^-$	$1^+$	$1^+$	$0^-$
$F^*(q^2, k^2)$	7	$V_\perp^*(q^2, k^2)$	$V_\parallel^*(q^2, k^2)$	$\hat{V}_\perp^*(q^2, k^2)$	$P^*(q^2, k^2)$
		$T_\perp^*(q^2, k^2)$	$T_\parallel^*(q^2, k^2)$	$T_\perp^*(q^2, k^2)$	---
$F(q^2) \equiv$	4	$V_\perp(q^2)$	$V_\parallel(q^2)$	$V_\perp(q^2) = V_\parallel(q^2)$	$P(q^2) = 0$
$F^*(q^2, 0)$		$T_\perp(q^2)$	$T_\parallel(q^2)$	$T_\perp(q^2) = T_\parallel(q^2)$	---
$F^*(0, k^2)$	5	$V_\perp^*(0, k^2)$	$V_\parallel^*(0, k^2)$	$\hat{V}_\perp^*(0, k^2) = P^*(0, k^2)$	$P^*(0, k^2)$
		$T_\perp^*(0, k^2)$	$T_\parallel^*(0, k^2) = T_\perp^*(0, k^2)(1-\hat{k}^2)$	$T_\perp^*(0, k^2)$	---

**Table 5:** The  $J^P = 0^+$  FF vanishes by parity conservation of QCD. Generally, there are seven independent  $F^*(q^2, k^2)$  FFs (light-blue) with two constraints  $\hat{V}_\perp^*(0, k^2) = P^*(0, k^2)$  (A.9) and  $T_\parallel^*(0, k^2) = (1-\hat{k}^2)T_\perp^*(0, k^2)$  (A.11). For the photon on-shell case,  $F(q^2) \equiv F^*(q^2, 0)$ , there are four independent FFs (light-red) and the reduction is due to the absence of the photon 0-helicity component. At zero flavour-violating momentum there are five independent FFs (light-green), due to the two constraints mentioned above. For the computation of the  $B \rightarrow \ell\ell\gamma$  SM rate, the following five FFs are sufficient  $\{V_{\perp,\parallel}(q^2), T_{\perp,\parallel}(q^2), T_\perp^*(0, k^2)\}$ .

Since the last few section were a bit involved with many steps we summarise the classification in Table 5. In general there are seven FFs for the  $B \rightarrow 1^-$  transition. In the photon on-shell case this reduces to four because the photon comes with two polarisations only. In the case of zero flavour-violating momentum transfer the two general constraints (A.9,A.11) and reduces this number from seven to five.

#### A.1.5 Derivation of $T_\parallel^{B \rightarrow \gamma^*}(0, k^2) = (1-\hat{k}^2)T_\perp^{B \rightarrow \gamma^*}(0, k^2)$

At last we turn to the derivation of the relation (A.11). To do so one has to uncontract  $q^\nu$  in (A.1). We first write an uncontracted  $B \rightarrow V$  matrix element

$$b_T \langle V(\eta(k)) | \bar{q} \sigma^{\mu\nu} \gamma_5 b | \bar{B}_q(p) \rangle = x_0 \eta^* \cdot p \frac{k^{[\mu} p^{\nu]}}{q \cdot k} + x_1 \eta^{*\mu} k^\nu + x_2 \eta^{*\mu} p^\nu \equiv \eta^{*\alpha} x_\alpha^{\mu\nu}, \quad (\text{A.25})$$

with shorthands  $x_i = x_i(q^2, k^2)$ ,  $p = p_B$ ,  $\eta$  is the polarisation vector of a massive vector boson and square brackets denote antisymmetrisation in the respective indices. The corresponding uncontracted  $B \rightarrow \gamma^*$  matrix element then reads

$$M_{T_5}^{\mu\nu\rho} \equiv b_T \langle \gamma^*(k, \rho) | \bar{q} \sigma^{\mu\nu} \gamma_5 b | \bar{B}_q(p) \rangle = X_0 R^\rho \frac{k^{[\mu} p^{\nu]}}{q \cdot k} + X_1 g^{\rho[\mu} k^{\nu]} + X_2 (g^{\rho[\mu} p^{\nu]} - \frac{k^\rho}{k^2} k^{[\mu} p^{\nu]}) , \quad (\text{A.26})$$

where technically

$$M_{T_5}^{\mu\nu\rho} = c G^{\rho\alpha} x_\alpha^{\mu\nu} , \quad (\text{A.27})$$

with  $c$  some  $i$ -independent kinematic function ( $X_i = cx_i$ ) which is irrelevant for our purposes. The appearance of the tensor  $G^{\rho\alpha}$  can be understood from the viewpoint of a dispersion relation cf. Section A.2. Regularity enforces at  $k \cdot q \propto \text{tol} - \hat{k}^2 - \hat{q}^2 \rightarrow 0$ ,

$$X_0(q^2, m_{B_q}^2 - q^2) = 0 , \quad (\text{A.28})$$

and at  $k^2 \rightarrow 0$  we have

$$X_0(q^2, 0) + X_2(q^2, 0) = 0 . \quad (\text{A.29})$$

These two constraints are generally valid.

We may make the connection with our basis by identifying

$$M_{T_5}^{\mu\rho} = iq_\nu M_{T_5}^{\mu\nu\rho} , \quad M_T^{\mu\rho} = -\frac{i}{2}(iq_\nu) \epsilon^{\mu\nu}_{\mu'\nu'} M_{T_5}^{\mu'\nu'\rho} , \quad (\text{A.30})$$

to obtain

$$\begin{aligned} T_\perp^*(q^2, k^2) &= -(X_1(q^2, k^2) + X_2(q^2, k^2)) , \\ T_\parallel^*(q^2, k^2) &= -\frac{1 - \hat{k}^2}{1 - \hat{q}^2} (X_1(q^2, k^2) + X_2(q^2, k^2)) + \frac{\hat{q}^2}{1 - \hat{q}^2} (X_1(q^2, k^2) - X_2(q^2, k^2)) , \\ T_L(q^2, k^2) &= (X_2(q^2, k^2) - X_1(q^2, k^2)) + 2\frac{1 - \hat{k}^2}{1 - \hat{k}^2 + \hat{q}^2} X_0(q^2, k^2) . \end{aligned} \quad (\text{A.31})$$

There are two consequences of this equation. Since  $X_1$  and  $X_2$  are free from poles at  $q^2 = 0$  on gets (A.11),

$$T_\parallel^*(0, k^2) = (1 - \hat{k}^2) T_\perp^*(0, k^2) , \quad (\text{A.32})$$

and by inserting (A.29) into  $T_L^*$  one deduces that  $T_L(q^2) = T_\parallel(q^2)$  which we derived earlier cf. (A.18). This confirms the earlier observation that the regularity conditions in  $k^2 \rightarrow 0$  are equivalent to the previously mentioned helicity argument. The derivation of relations (A.32) achieves the purpose of this section.

## A.2 Relation of the $B \rightarrow \gamma^*$ - and $B \rightarrow V$ -basis through the dispersion relation

In this appendix we make the link between the  $B \rightarrow V$ - and the  $B \rightarrow \gamma^*$ -FFs through the dispersion relations. This is an instructive exercise and we will be able to recover properties of the  $B \rightarrow \gamma^*$  FFs from the  $B \rightarrow V$ -ones. Our argumentation remains true if one considers any intermediate state (e.g. two pseudoscalar particles in a  $P$ -wave) as long as its quantum number,  $J^{PC} = 1^{--}$ , is equal to the one of the photon. This is the case since the properties follows from the general decomposition and the fact that any such state can be interpolated by the electromagnetic current in the LSZ formalism. In addition the dispersion representation might be useful for improving the fit ansatz of these FFs.

For our purposes it is convenient to first write the  $B \rightarrow V$  FFs [45, 46] in the following form<sup>10</sup>

$$\begin{aligned} c_V^{(q)} \langle V(k, \eta) | \bar{q} \gamma^\mu (1 \mp \gamma_5) b | \bar{B}(p_B) \rangle (-m_{B_q}) &= P_1^\mu \bar{\mathcal{V}}_1^{B \rightarrow V}(q^2) \pm \sum_{i=2,3,P} P_i^\mu \bar{\mathcal{V}}_i^{B \rightarrow V}(q^2), \\ c_V^{(q)} \langle V(k, \eta) | \bar{q} i q_\nu \sigma^{\mu\nu} (1 \pm \gamma_5) b | \bar{B}(p_B) \rangle &= P_1^\mu T_1^{B \rightarrow V}(q^2) \pm \sum_{i=2,3} P_i^\mu T_i^{B \rightarrow V}(q^2), \end{aligned} \quad (\text{A.33})$$

where  $\eta$  is the vector meson polarisation,  $P_i^\mu$  are Lorentz vectors

$$\begin{aligned} P_P^\mu &= i(\eta^* \cdot q) q^\mu, & P_1^\mu &= 2\epsilon_{\alpha\beta\gamma}^\mu \eta^{*\alpha} k^\beta q^\gamma, \\ P_2^\mu &= i(1 - \hat{k}^2) \{ m_{B_q}^2 \eta^{*\mu} - \frac{(\eta^* \cdot q)}{1 - \hat{k}^2} (k + p_B)^\mu \}, & P_3^\mu &= i(\eta^* \cdot q) \{ q^\mu - \frac{\hat{q}^2}{1 - \hat{k}^2} (k + p_B)^\mu \}, \end{aligned} \quad (\text{A.34})$$

where in (A.33) but not (A.34)  $k^2 = m_V^2$  is assumed, and

$$c_\omega^{(d)} = -c_{\rho^0}^{(d)} = \sqrt{2}, \quad c_\phi^{(s)} = 1, \quad (\text{A.35})$$

( $c_\omega^{(u)} = c_{\rho^0}^{(u)} = \sqrt{2}$  are note used) take into account the composition of the vector mesons' wave-functions,  $|\rho_0[\omega]\rangle \simeq (|\bar{u}u\rangle \mp |\bar{d}d\rangle)/\sqrt{2}$  and  $|\phi\rangle \simeq |\bar{s}s\rangle$ . The correspondence of  $\bar{\mathcal{V}}_{1,2,3,P}^{B \rightarrow V}$  with the more traditional FFs  $A_{0,1,2,3}^{B \rightarrow V}$  and  $V^{B \rightarrow V}$  is as follows

$$\begin{aligned} \bar{\mathcal{V}}_P^{B \rightarrow V}(q^2) &= \frac{2\hat{m}_V}{\hat{q}^2} A_0^{B \rightarrow V}(q^2), & \bar{\mathcal{V}}_1^{B \rightarrow V}(q^2) &= \frac{V^{B \rightarrow V}(q^2)}{1 + \hat{m}_V}, & \bar{\mathcal{V}}_2^{B \rightarrow V}(q^2) &= \frac{A_1^{B \rightarrow V}(q^2)}{1 - \hat{m}_V}, \\ \bar{\mathcal{V}}_3^{B \rightarrow V}(q^2) &= \left( \frac{1 - \hat{m}_V}{\hat{q}^2} A_2^{B \rightarrow V}(q^2) - \frac{1 + \hat{m}_V}{\hat{q}^2} A_1^{B \rightarrow V}(q^2) \right) \equiv \frac{-2\hat{m}_V}{\hat{q}^2} A_3^{B \rightarrow V}(q^2), \end{aligned} \quad (\text{A.36})$$

where  $\hat{m}_V \equiv m_V/m_{B_q}$ . The analogue of the two ( $q^2 = 0$ ) constraints (A.9,A.11) are

$$A_3^{B \rightarrow V}(0) = A_0^{B \rightarrow V}(0), \quad T_1^{B \rightarrow V}(0) = T_2^{B \rightarrow V}(0), \quad (\text{A.37})$$

respectively. The constraint (A.12) does not apply since  $m_V^2 \neq m_{B_q}^2$ .

As stated above the relation between the FFs becomes apparent in the dispersion representation (cf. the textbook [49] or the recent review [50]). A specific example is chosen for illustration,<sup>11</sup>

$$\begin{aligned} M_{T_5}^{\mu\rho} &= -ib_T \int d^4x e^{ik \cdot x} \langle 0 | T j^\rho(x) \bar{q} i q_\nu \sigma^{\mu\nu} \gamma_5 b(0) | \bar{B}_q \rangle \\ &= \sum_{i=2,3} \int_{u_{\text{low}}}^{\infty} \frac{\rho_{T_i}(q^2, u) du}{u - k^2 - i0} (-G^\rho{}_\alpha) P_i^{\mu\alpha} + \text{subtractions}, \end{aligned} \quad (\text{A.38})$$

where  $u_{\text{low}} = (m_{P_1} + m_{P_2})^2$ , and  $V \rightarrow P_1 + P_2$  is the lowest decay channel (e.g.  $\rho^0 \rightarrow \pi^+ + \pi^-$  for  $B_q = B_d$ ). The dispersion relation requires one subtraction (cf. discussion further below). Note, that the appearance of the tensor  $G^\rho{}_\alpha$  (A.5) goes hand in hand with the QED Ward identity constraint.

---

<sup>10</sup>Below  $\bar{\mathcal{V}} = (-m_{B_q})\mathcal{V}_i$  absorbs the factor on the left-hand side into the definition. This renders the  $\bar{\mathcal{V}}$  FFs dimensionless.

<sup>11</sup>In order to distinguish the various dispersion representations throughout this paper, we use the variables  $(s, t, u)$  for the momenta  $(p_B^2, q^2, k^2)$ .

In order to further illustrate we resort to the narrow width approximation (NWA) which can be improved by introducing a finite decay width or better multiparticle states of stable particles (cf. remark at beginning this section). In this NWA  $G_{\rho\lambda}$  appears through the sum of polarisation vectors

$$\sum_{\lambda=-1,0,1} \eta_\rho^*(k, \lambda) \eta_\alpha(k, \lambda) = \left( -g_{\rho\alpha} + \frac{k_\rho k_\alpha}{m_V^2} \right) |_{m_V^2=k^2} = (-G_{\rho\alpha}) , \quad (\text{A.39})$$

and the spectral or discontinuity function  $\rho_{T_i}(q^2, u)$  assumes the simple form

$$\rho_{T_i}(q^2, u) = \delta(u - m_\rho^2) r_{T_i}^\rho(q^2) + \delta(u - m_\omega^2) r_{T_i}^\omega(q^2) + \dots , \quad (\text{A.40})$$

where the dots stand for higher states in the spectrum. The residua  $r_{T_i}^V$  are then given by

$$r_{T_i}^V = -m_V f_V^{\text{em}} / |c_V^{(d,s)}|^2 T_i^{B \rightarrow V}(q^2) , \quad (\text{A.41})$$

where  $f_V^{\text{em}}$  is a conveniently normalised matrix element

$$(c_V^{(d,s)})^* \langle 0 | j_\mu | V(k, \eta) \rangle = m_V f_V^{\text{em}} \eta_\mu , \quad (\text{A.42})$$

of the electromagnetic current and the vector meson. In particular

$$f_{\rho_0}^{\text{em}} = (Q_d - Q_u) f_{\rho_0} = -f_{\rho_0} , \quad f_\omega^{\text{em}} = (Q_d + Q_u) f_\omega = \frac{1}{3} f_\omega , \quad f_\phi^{\text{em}} = Q_s f_\phi = -\frac{1}{3} f_\phi . \quad (\text{A.43})$$

Rewriting our parametrisation (A.1), in compact form,

$$M_{T_5}^{\mu\rho} = \sum_{J=\parallel, \perp} R_J^{\mu\rho} T_J^{B \rightarrow \gamma^*}(q^2, k^2) , \quad (\text{A.44})$$

and equating with (A.38) we are able to identify the two bases

$$R_J^{\mu\rho} \omega_{Ji}(q^2, k^2) = P_i^{\mu\alpha} (-G_\alpha^\rho) , \quad (\text{A.45})$$

where  $\omega_{Ji}$  is a matrix with diagonal entries

$$\omega_{\perp 1}(q^2, k^2) \equiv 2 , \quad \omega_{\parallel 2}(q^2, k^2) \equiv 2 \frac{1 - \hat{k}^2}{1 - \hat{q}^2} , \quad \omega_{\perp 3}(q^2, k^2) \equiv 2 , \quad \omega_{PP}(q^2, k^2) \equiv 2 , \quad (\text{A.46})$$

and all others set to zero. Of course we could have chosen any other basis at the cost of having a non-diagonal  $\omega$ -matrix but we feel that this is a economic way.

Let us make the dispersion representation more concrete for which we first need to clarify what the subtraction terms mean in (A.38). Unlike the  $B \rightarrow V$  FFs, the  $B \rightarrow \gamma^*$  ones require a single subtraction. This can be inferred from the asymptotic behaviour of LO perturbation theory which is  $\ln k^2$  (cf. the explicit results in Section A.3.2).<sup>12</sup> The asymptotic behaviour of  $B \rightarrow V$  FFs is  $F^{B \rightarrow V} \propto 1/k^2$  and therefore does not require a subtraction. Finally we may write

$$T_J^{B \rightarrow \gamma^*}(q^2, k^2) = T_J^{B \rightarrow \gamma^*}(q^2, k_0^2) + \omega_{Ji}(k^2 - k_0^2) \int_{u_{\text{low}}}^{\infty} \frac{\rho_{T_i}(q^2, u) du}{(u - k_0^2 - i0)(u - k^2 - i0)} , \quad (\text{A.47})$$

---

<sup>12</sup>This also holds when resumming the LO expressions cf. Ref. [51] for the correlation functions in question. For the tensor correlation function to converge one would need more flavours ( $N_c = 3$ ) than asymptotic freedom allows as can be inferred from the  $\bar{q}q$ -correlator example in the appendix.

where  $J = \perp, \mathbb{L}, P$  and  $i = 1, 3, P$  with  $\omega_{Ji}$  defined above and the same formula applies for  $T_J^* \rightarrow V_J^*$  and  $\rho_{T_i} \rightarrow \rho_{\bar{\mathcal{V}}_i}$ . The  $T, V_{\parallel}$ -FFs are a bit more involved. One defines

$$\Delta T_{\parallel}^{B \rightarrow \gamma^*}(q^2, k^2) = \frac{T_{\parallel}^{B \rightarrow \gamma^*}(q^2, k^2) - T_{\parallel}^*(q^2, m_{B_q}^2)}{\omega_{\parallel 2}(q^2, k^2)}, \quad (\text{A.48})$$

and then the correct dispersion relation reads

$$\Delta T_{\parallel}^{B \rightarrow \gamma^*}(q^2, k^2) = \Delta T_{\parallel}^{B \rightarrow \gamma^*}(q^2, k_0^2) + (k^2 - k_0^2) \int_{u_{\text{low}}}^{\infty} \frac{\rho_{T_2}(q^2, u) du}{(u - k_0^2 - i0)(u - k^2 - i0)}. \quad (\text{A.49})$$

The same applies again for  $V_{\parallel}^*$  with the substitutions  $T_{\parallel}^* \rightarrow V_{\parallel}^*$  and  $\rho_{T_2} \rightarrow \rho_{\bar{\mathcal{V}}_2}$ . The analogy with (A.49) is restored if one divides the latter equation by  $\omega_{Ji}$ .

For the sake of clarity, we give a few examples of FFs in the  $k^2$ -dispersion representation (A.49) which illustrates some of its properties:

$$\begin{aligned} V_{\perp}^{B_d \rightarrow \gamma^*}(q^2, k^2) &= V_{\perp}^{B_d \rightarrow \gamma^*}(q^2, k_0^2) - \frac{\omega_{\perp 1}}{2}(k^2 - k_0^2) \left( \frac{f_{\rho}^{\text{em}} \bar{\mathcal{V}}_1^{B_d \rightarrow \rho}(q^2)}{(m_{\rho}^2 - k^2)(m_{\rho}^2 - k_0^2)} + \frac{f_{\omega}^{\text{em}} \bar{\mathcal{V}}_1^{B_d \rightarrow \omega}(q^2)}{(m_{\omega}^2 - k^2)(m_{\omega}^2 - k_0^2)} + \dots \right), \\ \hat{V}_{\mathbb{L}}^{B_s \rightarrow \gamma^*}(q^2, k^2) &= \hat{V}_{\mathbb{L}}^{B_s \rightarrow \gamma^*}(q^2, k_0^2) - \omega_{\mathbb{L} 3}(k^2 - k_0^2) \left( \frac{m_{\phi}}{m_{B_q}} \frac{m_{\phi} f_{\phi}^{\text{em}} A_3^{B_s \rightarrow \phi}(q^2)}{(m_{\phi}^2 - k^2)(m_{\phi}^2 - k_0^2)} + \dots \right), \\ P^{B_s \rightarrow \gamma^*}(q^2, k^2) &= P^{B_s \rightarrow \gamma^*}(q^2, k_0^2) - \omega_{PP}(k^2 - k_0^2) \left( \frac{m_{\phi}}{m_{B_q}} \frac{m_{\phi} f_{\phi}^{\text{em}} A_0^{B_s \rightarrow \phi}(q^2)}{(m_{\phi}^2 - k^2)(m_{\phi}^2 - k_0^2)} + \dots \right), \\ T_{\perp}^{B_s \rightarrow \gamma^*}(q^2, k^2) &= T_{\perp}^{B_s \rightarrow \gamma^*}(q^2, k_0^2) - \omega_{\perp 1}(k^2 - k_0^2) \left( \frac{m_{\phi} f_{\phi}^{\text{em}} T_1^{B_s \rightarrow \phi}(q^2)}{(m_{\phi}^2 - k^2)(m_{\phi}^2 - k_0^2)} + \dots \right). \end{aligned} \quad (\text{A.50})$$

Above the  $k^2 + i0$  prescription has been dropped for brevity and  $|c_{\rho^0}|^2 = |c_{\omega}|^2 = 2$  has been used.

These formulae show that properties of the  $B \rightarrow \gamma^*$ - and  $B \rightarrow V$ -FFs imply each other. For example, the  $B \rightarrow V$  constraint  $A_0^{B \rightarrow V}(0) = A_3^{B \rightarrow V}(0)$  (A.37) implies the constraint  $P^*(0, k^2) = \hat{V}_{\mathbb{L}}^*(0, k^2)$  (A.9). The algebraic relation (A.11) follows from (A.49), if  $T_{\parallel}(0, m_{B_q}^2) = 0$  holds which in turn follows from (A.12). In the  $SU(3)_F$  limit  $m_u = m_d = m_s$ ,  $m_V$  and  $f_V$  are degenerate and (A.43),  $f_{\rho}^{\text{em}}/|c_{\rho}|^2 + f_{\omega}^{\text{em}}/|c_{\omega}|^2 = f_{\phi}^{\text{em}}/|c_{\phi}|^2$  and  $F^{B_d \rightarrow \rho} = F^{B_d \rightarrow \omega} = F^{B_s \rightarrow \phi}$  which finally implies  $F^{B_d \rightarrow \gamma^*} = F^{B_s \rightarrow \gamma^*}$  as expected. These relations can be turned around since they hold for any  $k^2$ , they necessarily hold at each point of the spectrum and thus for the  $B \rightarrow \gamma^*$  properties imply the  $B \rightarrow V$  FF properties.

Moreover, the examples reveal that the slope of the FF are positive which is the choice by convention. This is the case since  $r_{\phi} > 0$  and  $r_{\rho} > |r_{\omega}| > 0$ . At last let us note that a particularly convenient form for  $P^*$  can be obtained

$$P^{B_s \rightarrow \gamma^*}(q^2, k^2) = -2k^2 \left( \frac{1}{m_{B_q}} \frac{f_{\phi}^{\text{em}} A_0^{B_s \rightarrow \phi}(q^2)}{(m_{\phi}^2 - k^2)} + \dots \right), \quad (\text{A.51})$$

if ones chooses the subtraction point  $k_0^2 = 0$  where the FF vanishes.

### A.3 Explicit results of the off-shell form factor

#### A.3.1 QCD sum rule for the off-shell form factors $P^*(0, k^2)$ , $T_{\perp,\parallel}^*(0, k^2)$ and $V_{\perp,\parallel}^*(0, k^2)$

The FFs are computed using QCD sum rules [52]. The starting point is the correlation function of the form

$$\begin{aligned}\Pi_{\mu\rho}^{V-A}(p_B, q) &\equiv -i^2(b_V s_e e) \int_{x,y} e^{-ip_B \cdot x} e^{ik \cdot y} \langle 0 | T j_\rho(y) J_{B_q}(x) \bar{q} \gamma^\mu (1 - \gamma_5) b(0) | 0 \rangle \\ &= R_{\mu\rho}^\perp \Pi_\perp^V - (R_{\mu\rho}^\parallel \Pi_\parallel^A + R_{\mu\rho}^\parallel \Pi_\parallel^A + R_{\mu\rho}^P \Pi_P^A) + C_{\mu\rho}(q^2),\end{aligned}\quad (\text{A.52})$$

where  $(b_V s_e e) = -m_{B_q}$ ,  $\Pi^{V,A} = \Pi^{V,A}(q^2, p_B^2, k^2)$  are analytic functions in three variables and the Lorentz structures  $R_{\mu\rho}$  are defined in (A.3). Gauge invariance, again, holds in the simplest form  $k^\rho \Pi_{\mu\rho}^V(p_B, q) = 0$  since we work with electrically neutral states. The term  $C_{\mu\rho}(q^2)$  is a contact term but of no relevance for our purposes as independent of  $p_B^2$ . It is the correction to the naive non-singlet axial Ward identity (A.7). The operator  $J_{B_q} \equiv (m_b + m_q) \bar{b} i \gamma_5 q$  is the interpolating operator for the  $B_q$ -meson with matrix element  $\langle \bar{B}_q | J_{B_q} | 0 \rangle = m_{B_q}^2 f_{B_q}$ .

The QCD sum rule is then obtained by evaluating (A.52) in the operator product expansion (OPE) (cf. Figure 4) and equating it to the dispersion representation. The OPE consists of a perturbative part and a condensate part for which we include only the quark condensate. The OPE is convergent, in a pragmatic sense, for momenta  $p_B^2, q^2 < O(m_b \Lambda)$  and  $k^2 < -\Lambda^2$  with  $\Lambda \simeq 500$  MeV a typical hadronic scale. The perturbative part is evaluated with the help of FeynCalc [53, 54]. We neglect light quark masses i.e.  $m_d = m_s = 0$ .

The dispersion representation of  $\Pi_\perp^V$  reads

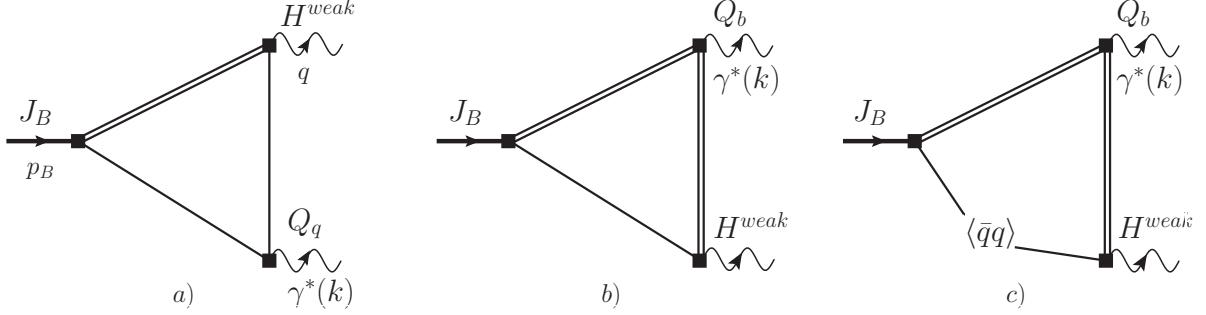
$$\Pi_\perp^V(p_B^2, q^2, k^2) = \frac{1}{\pi} \int_0^\infty \frac{\text{Im}[\Pi_\perp^V(s, q^2, k^2)] ds}{s - p_B^2 - i0} = \frac{m_{B_q}^2 f_{B_q} V_\perp^{B \rightarrow \gamma^*}(q^2, k^2)}{m_{B_q}^2 - p_B^2 - i0} + \dots, \quad (\text{A.53})$$

where the dots stand for higher resonances and multiparticle states. Moreover the NWA for the  $B$ -meson has been assumed. The FFs are then extracted via the standard procedures of Borel transformation and approximating the ‘‘higher states’’ contribution by the perturbative integral [52]. The latter is exponentially suppressed

$$V_\perp^{B \rightarrow \gamma^*}(q^2, k^2) = \frac{1}{m_{B_q}^2 f_{B_q}} \int_{m_b^2}^{s_0} e^{(m_{B_q}^2 - s)/M^2} \rho_{V_\perp^*}(s, q^2, k^2) ds, \quad (\text{A.54})$$

due to the Borel transform in  $p_B^2$ . Note, that the contact term  $C_{\mu\rho}(q^2)$ , which can appear as a subtraction constant in the dispersion relation, vanishes for definite under the Borel transform. Above  $\pi \rho_\perp^V(s, q^2, k^2) = \text{Im}[\Pi_\perp^V(s, q^2, k^2)]$  and  $M^2$  is the Borel mass. If we were able to compute  $\rho_\perp^V$  exactly then  $V_\perp(q^2)$ , obtained from (A.54), would be independent of the Borel mass and it therefore serves as a quality measure of the sum rule. Other FFs are obtained in exact analogy.

Before stating the results of the computations let us turn to the issue of analytic continuation. We would like to employ our FFs in the Minkowski region  $k^2 > 0$ , whereas the OPE is convergent for  $k^2 < -\Lambda^2$ . The convergence is broken by thresholds at  $k^2 = 4m_q^2$  which signal long-distance effects corresponding to  $\rho/\omega(\phi)$ -like resonances cf. Table 4. The standard procedure is to analytically continue into the Minkowski region and use the FF for say  $k^2 > 4 \text{ GeV}^2$  which is far enough from the lowest lying narrow resonances. For  $k^2 > 4 \text{ GeV}^2$  the resonances are broad and disappear into the continuum. Under such circumstances local quark-hadron duality is usually assumed to be a reasonable approximation. In



**Figure 4:** Figures for off-shell form factors  $V$ ,  $T_{\perp,\parallel,\perp}^*$  and  $P^*$ . The single/double lines denote the  $q/b$ -quark respectively. The diagrams on the left and right are the perturbative- and the one on the left is the quark condensate-type. The quark condensate diagram corresponding to fig a) is not shown. It is proportional to  $\langle \bar{q}q \rangle / k^2$  and implicitly assumes  $k^2 \neq 0$ . In the  $k^2 \rightarrow 0$  limit these diagrams are replaced by the photon distribution amplitude e.g. [30]. In the sum rule method the  $B$ -meson is projected out via a dispersion relation in the variable  $p_B^2$  giving access to the matrix element of the off-shell form factor  $F^*(q^2, k^2)$ . The momentum assignment corresponds to the so-called ‘theory convention’.

our region of use  $k^2 \in [(4.9 \text{ GeV})^2, m_{B_{d,s}}^2]$  there are no narrow resonances in the  $k^2$ -channel.<sup>13</sup> On a pragmatic level it is best to implement the  $V_{\perp}^{B \rightarrow \gamma^*}(q^2, k^2 + i0)$ -prescription in the process of analytic continuation by deforming the path in (A.54) from  $\int_{m_b^2}^{s_0} ds \rightarrow \int_{\gamma} ds$ , where  $\gamma$  is a path in the upper-half plane starting at  $m_b^2$  and ending at  $s_0$ . One may for instance choose a semi-circle in the upper half-plane. This prescription leads to numerical stability. Clearly our computation remains valid and useful for  $D^0 \rightarrow \gamma^*$  FFs with replacements  $B_q \rightarrow D^0$  and  $m_b \rightarrow m_c$ .

---

<sup>13</sup>In fact one should able to use these computation up to  $k^2 \simeq 4m_b^2 - O(10 \text{ GeV}^2)$ .

### A.3.2 Explicit results for the off-shell form factors from QCD sum rules

The explicit FFs are found to be

$$\begin{aligned}
P^*(0, k^2) &= \frac{e^{m_{Bq}^2/M^2}}{f_{Bq} m_{Bq}^2} Q_b \left( \int_{m_b^2}^{s_0} e^{-s/M^2} \rho_{P^*}(s, 0, k^2) ds - \frac{2}{\bar{m}_{Bq}} \langle \bar{q}q \rangle e^{-m_b^2/M^2} \right) + O(\alpha_s, m_q), \\
T_\perp^*(0, k^2) &= \frac{e^{m_{Bq}^2/M^2}}{f_{Bq} m_{Bq}^2} Q_b \left( \int_{m_b^2}^{s_0} e^{-s/M^2} \rho_{T_\perp^*}(s, 0, k^2) ds + \left(1 - \frac{1}{\bar{k}^2}\right) \langle \bar{q}q \rangle e^{-m_b^2/M^2} \right) + O(\alpha_s, m_q), \\
T_{\mathbb{L}}^*(0, k^2) &= \frac{e^{m_{Bq}^2/M^2}}{f_{Bq} m_{Bq}^2} Q_b \left( \int_{m_b^2}^{s_0} e^{-s/M^2} \rho_{T_{\mathbb{L}}^*}(s, 0, k^2) ds + \left(1 - \frac{1}{\bar{k}^2}\right) \langle \bar{q}q \rangle e^{-m_b^2/M^2} \right) + O(\alpha_s, m_q), \\
V_\perp^*(0, k^2) &= \frac{e^{m_{Bq}^2/M^2}}{f_{Bq} m_{Bq}^2} Q_b \left( \int_{m_b^2}^{s_0} e^{-s/M^2} \rho_{V_\perp^*}(s, 0, k^2) ds + \bar{m}_{Bq} \left(1 - \frac{1}{\bar{k}^2}\right) \langle \bar{q}q \rangle e^{-m_b^2/M^2} \right) + O(\alpha_s, m_q), \\
V_{\parallel}^*(0, k^2) &= \frac{e^{m_{Bq}^2/M^2}}{f_{Bq} m_{Bq}^2} Q_b \left( \int_{m_b^2}^{s_0} e^{-s/M^2} \rho_{V_{\parallel}^*}(s, 0, k^2) ds + \bar{m}_{Bq} \left(1 + \frac{1}{\bar{k}^2}\right) \langle \bar{q}q \rangle e^{-m_b^2/M^2} \right) + O(\alpha_s, m_q),
\end{aligned} \tag{A.55}$$

where  $\bar{m}_{Bq} \equiv m_{Bq}/m_b$ ,  $\bar{s} \equiv s/m_b^2$ ,  $\bar{k}^2 \equiv k/m_b^2$  and the perturbative densities

$$\rho_i \equiv \frac{N_c m_b}{4\pi^2 (\bar{s} - \bar{k}^2)^3} \hat{\rho}_i, \tag{A.56}$$

are given by

$$\begin{aligned}
\hat{\rho}_{P^*}(s, 0, k^2) &= 2\bar{k}^2/\bar{m}_{Bq} \{ L_q [(\bar{s} - \bar{k}^2) - 1] + L_b \}, \\
\hat{\rho}_{T_\perp^*}(s, 0, k^2) &= [\bar{k}^4(1 - \bar{s}) + 2\bar{k}^2(\bar{s} - 1) - (\bar{s} - 1)\bar{s}^2] + L_q [\bar{k}^2] + L_b [-\bar{k}^2], \\
\hat{\rho}_{T_{\mathbb{L}}^*}(s, 0, k^2) &= (\bar{s} - 1)(5\bar{k}^2 - \bar{s})(\bar{s} - \bar{k}^2) + L_q [\bar{k}^2/(\bar{s} - \bar{k}^2)(4\bar{k}^4 + 7\bar{k}^2 + (5 - 4\bar{s})\bar{s})] + \\
&\quad L_b [\bar{k}^2/(\bar{s} - \bar{k}^2)(\bar{k}^2(8\bar{s} - 7) - \bar{s}(8\bar{s} + 5))], \\
\hat{\rho}_{V_\perp^*}(s, 0, k^2) &= \bar{m}_{Bq} \{ L_q [\bar{k}^4 - \bar{k}^2(\bar{s} - 2)] + L_b [\bar{k}^2(\bar{s} - 2) - \bar{s}^2] \}, \\
\hat{\rho}_{V_{\parallel}^*}(s, 0, k^2) &= \bar{m}_{Bq} \{ 2\bar{k}^2/\bar{s}^2(\bar{s} - 1)(\bar{s} - \bar{k}^2)^2 + L_q [\bar{k}^2/\bar{s}(\bar{k}^4 - 2\bar{k}^2(\bar{s} - 1) + \bar{s}^2 - 2\bar{s} + 2)] + \\
&\quad L_b [(\bar{k}^4\bar{s} - 2\bar{k}^2(\bar{s}^2 - \bar{s} + 1) + (\bar{s} - 2)\bar{s}^2)/s] \}, \tag{A.57}
\end{aligned}$$

with

$$L_q \equiv \ln \left( \frac{\bar{k}^2}{\bar{s}(1 + \bar{k}^2 - \bar{s})} \right), \quad L_b \equiv \ln \left( \bar{s} - \bar{k}^2 + \frac{\bar{k}^2}{\bar{s}} \right). \tag{A.58}$$

The logarithms  $L_q$  and  $L_b$  in (A.58) lead to imaginary parts in the FFs for  $k^2 > 4m_q^2 = 0$  and  $k^2 > 4m_b^2$  respectively. These expression are consistent with the  $B \rightarrow V\ell\ell$  weak annihilation computation detailed in appendix of Ref. [28] cf. footnote 2 for further remarks. Note, there is no singularity at  $k^2 = s$  when

expanded properly. The condensate contributions could be written in terms of the densities  $\rho_i$  as well. The backward substitution  $e^{-m_b^2/M^2} \langle \bar{q}q \rangle \rightarrow \langle \bar{q}q \rangle \delta(s - m_b^2)$  achieves this task.

A few comments on interpreting the results. The  $k^2 \rightarrow 0$  limit is not well-defined for the condensates. In that limit the condensates originating from quark lines attached to the photon are replaced by a photon distribution amplitude which makes the FFs computation more involved. However, the perturbative part remains well-defined in that limit. Hence the latter must contribute positively to the  $\{T_{\perp,\parallel}(0,0), V_{\perp,\parallel}(0,0)\}$  by convention which can be verified indeed by using  $Q_b = -1/3$  and sending  $\hat{k}^2 \rightarrow 0$ . The  $q^2$  constraints (A.9,A.11) are obeyed exactly by the sum rules and are assumed as we do not show  $\hat{V}_{\perp}(0, k^2)$  and  $T_{\parallel}(0, k^2)$ ; they are simply redundant. The constraints at  $k^2 = m_{B_q}^2$  (A.24) are obeyed for the correlation functions with  $k^2 = p_B^2$ . However they do not hold exactly for the FFs as  $p_B^2 \simeq m_{B_q}^2$  within approximation of the Borel procedure. We have checked that these relations hold to within 2% where for the last one we compare to a value of the FF at  $q^2 = 10 \text{ GeV}^2$ . In the fits we have implemented these constraints as they are important to cancel the poles present in the Lorentz structures  $R_{\parallel}^{\mu\rho}$  and  $R_{\perp}^{\mu\rho}$ .

The expressions could be improved by  $m_q \neq 0$ , adding the gluon condensate and radiative corrections. The first two are expected to be rather small effects since  $1 \gg m_q/\Lambda_{\text{QCD}}$  and  $m_b \langle \bar{q}q \rangle \gg \langle G^2 \rangle$ . On the other hand, radiative corrections could be sizeable and would of course reduce the scale uncertainty considerably.

For the numerical input we use the MS-bar mass  $m_b = 4.18(4) \text{ GeV}$  and  $\langle \bar{q}q \rangle_{\mu=1 \text{ GeV}} = -(0.24(1) \text{ GeV})^3$  (e.g. [55]) and

$$\{M^2, s_0\}_{B_d} = \{5(2), 36(2)\} \text{ GeV}^2, \quad \{M^2, s_0\}_{B_s} = \{6(2), 37(2)\} \text{ GeV}^2$$

The values of the Borel parameter and the continuum threshold are roughly consistent (2–3% in the region of interest) with the formally exact relation

$$m_{B_q}^2 = -e^{m_{B_q}^2/M^2} \frac{d}{d(1/M^2)} e^{-m_{B_q}^2/M^2} \ln P^*(k^2, M^2, s_0). \quad (\text{A.59})$$

Imposing this constraint is equivalent to extremising in the Borel parameter [46]. For the decay constant  $f_{B_q}$  we use the  $\alpha_s^0$ -result in [56] with similar Borel parameter and continuum threshold. This corresponds to a 11% reduction w.r.t. the NLO  $f_B$ . For the uncertainty analysis we use the same procedure as in Ref. [46] with some more detail in the fit-section.

## A.4 Dispersion relation and fit ansatz for form factors

### A.4.1 Extending the $B \rightarrow \gamma$ on-shell form factors into the $q^2 \simeq m_B^2$ -region

The  $B \rightarrow \gamma$  on-shell FFs  $F(q^2) \equiv F^*(q^2, 0)$  (A.23) are the LO part of a NLO computation [30], computed with light-cone sum rules. The region of validity of the computation is the previously mentioned  $q^2 < m_b \Lambda$  which is just outside our region of interest  $q^2 \in [(4.9 \text{ GeV})^2, m_{B_{d,s}}^2]$ . Progress can be made with the help of the generally valid dispersion representation in the flavour violating momentum transfer  $q^2$

$$\begin{aligned} V_{\perp}(q^2) &= \frac{1}{\pi} \int_{\text{cut}}^{\infty} \frac{\text{Im}[V_{\perp}(t)] dt}{t - q^2 - i0} = \frac{r_{V_{\perp}}}{1 - q^2/m_{B_q^*}^2} + \dots, \\ V_{\parallel}(q^2) &= \frac{1}{\pi} \int_{\text{cut}}^{\infty} \frac{\text{Im}[V_{\parallel}(t)] dt}{t - q^2 - i0} = \frac{r_{V_{\parallel}}}{1 - q^2/m_{B_{1q}}^2} + \dots, \end{aligned} \quad (\text{A.60})$$

	$r_{V_\perp} \propto g_{B_q^* B_q \gamma}$	$r_{T_\perp} \propto g_{B_q^* B_q \gamma}$	$r_{V_\parallel} \propto g_{B_{1q} B_q \gamma}$	$r_{T_\parallel} \propto g_{B_{1q} B_q \gamma}$
$B_d \rightarrow \gamma$	0.166(18)	0.159(19)	0.083(8)	0.155(17)
$B_s \rightarrow \gamma$	0.154(17)	0.144(19)	0.078(8)	0.141(16)

**Table 6:** The residua of the poles at  $k^2 = m_{B_q^*}^2, m_{B_{1q}}^2$  proportional to the on-shell matrix element which are preliminary LO numbers of an NLO computation [31].

where the dots stand for higher resonances and multiparticle states. The values and quantum numbers of the resonances are collected in Table 7. The dispersion relation of the other FFs are analogous. The residua are related to the  $B_q^* \rightarrow B_q \gamma$  and  $B_{1q} \rightarrow B_q \gamma$  on-shell matrix elements respectively. Unfortunately they are not known from experiment.<sup>14</sup> They can be extracted from the same sum rule as the FFs themselves by applying a double dispersion relation to interpolate for the  $B_q^*$ - and  $B_{1q}$ -meson respectively. We take the LO result of this residue from [31], collected in Table 6.

Here, we make the link to the predictions of Ref. [36], for which a single  $m_{B_s^*}$ -pole approximation was employed to estimate the FFs for the radiative decay. The single-pole approximation is expected to give a reasonable approximation around the pole provided the residue is known sufficiently well. By identifying the defining matrix elements of the residue (cf. Eq.(6) in Ref. [36]) we find the relation

$$|r_{V_\perp^{B_s \rightarrow \gamma}}| = |\mu| |f_{B_s}|_{[36]} \simeq 0.265 , \quad (\text{A.61})$$

with  $f_{B_s} = 227\text{MeV}$  [55] the standard decay constant and  $|\mu|$  defines the strength of the on-shell matrix element in Ref. [36]. The authors of Ref. [36] determine  $|\mu| = 1.13\text{GeV}^{-1}$  in an effective-theory approach valid at leading order in  $1/m_{b,c}$  using experimental data from  $D^{*+} \rightarrow D^+ \gamma$  and  $D^{*0} \rightarrow D^+ \pi^-$ . They neglect the pole of the  $B_{s1}$  meson (cf. Table 7) and thus we cannot compare the  $|r_{V_\parallel}|$  residua to theirs. Given the methods employed on both sides the discrepancy of 0.154(17) and 0.265 is not too surprising. Whereas the former is LO in the coupling with preliminary error analysis, the latter is subject to  $1/m_c$  corrections which might well be sizeable.

At last let us mention that we performed a non-trivial test of the identification in Eq. (A.61). Approximating our FF-expression to the pole part, inserting it into our rate in Eq. (2.10), and then comparing to the rate in Ref. [36] (cf. their Eq. (25)), we can confirm that Eq. (A.61) is consistent with both rates. This is a strong hint of the correctness of the treatment in our work and theirs.

#### A.4.2 The dispersion representation of the $B \rightarrow \gamma^*$ off-shell form factors

The assumed  $q^2 = 0$  is well below the various  $m_B^2$ -type poles and does not affect the computation. However, in the variable  $k^2$  there are the previously mentioned  $\rho/\omega(\phi)$ - and  $\Upsilon$ -resonances (cf. Table 7) which are far away from our region of interest  $k^2 \simeq m_B^2$  and therefore have little impact. If one wanted to fit the FFs at lower  $k^2$  then a dispersion ansatz, e.g. (A.50), could be combined with the  $z$ -expansion.

#### A.4.3 Fit ansatz and $z$ -expansion

The procedure to fit the FFs and how to include the correlation of uncertainties largely follows Ref. [46]. Based on the previous part of this section let us first motivate the fit-ansatz before summarising the

<sup>14</sup>The width of the  $B_{d,s}^*$ -mesons are unknown and the  $B_{(d,s)1}$  mesons are dominated by the strong decays to  $B_{(d,s)} \pi$ .

	$m_{B_q} = m_{0-}$	$m_{B^*} = m_{1-} = m_\perp$	$m_{B_{q1}} = m_{1+} = m_\parallel$
$b \rightarrow s$	5.367 GeV	5.415 GeV	5.829 GeV
$b \rightarrow d, u$	5.280 GeV	5.325 GeV	5.724 GeV

**Table 7:** Lowest resonances masses for Form Factors [55]. The mass  $m_{0-}$  is the  $B$ -meson mass in  $B_q \rightarrow \gamma$ . The Form Factors  $V, T_\perp$  and  $V, T_\parallel$  are associated with the resonances  $J^P = 1^-$  and  $1^+$  respectively.

essence of the  $z$ -expansion. There are four on-shell FFs and at  $q^2 = 0$  there are five off-shell FFs,

$$\begin{aligned} \text{on-shell: } & \{V_{\perp,\parallel}^{B \rightarrow \gamma}(q^2), T_{\perp,\parallel}^{B \rightarrow \gamma}(q^2)\}, \\ \text{off-shell: } & \{P^{B \rightarrow \gamma^*}(0, k^2), V_{\perp,\parallel}^{B \rightarrow \gamma^*}(0, k^2), T_{\perp,\parallel}^{B \rightarrow \gamma^*}(0, k^2)\}. \end{aligned} \quad (\text{A.62})$$

- The on-shell FFs are parameterised

$$F_n^{B \rightarrow \gamma}(q^2) = \frac{1}{1 - q^2/m_R^2} \left( \alpha_{n0} + \sum_{k=1}^N \alpha_{nk} (z(q^2) - z(0))^k \right), \quad (\text{A.63})$$

using the knowledge of the presence of the first pole  $m_R$  (A.60), cf., Table 7. The remaining part in brackets are supposed to take into account higher states in the spectrum. Specifically the  $\alpha_{nk}$ -coefficients are to be determined from a fit and  $z(q^2)$  is defined further below. The constraint of the residue, cf. (A.60) and Table 6, is implemented by

$$r_{V_\perp} = \alpha_{V_\perp 0} + \sum_{k=1}^N \alpha_{V_\perp k} (z(m_{B_q^*}^2) - z(0))^k, \quad (\text{A.64})$$

and similarly for other FFs. Further to that the constraint (A.11) is imposed by

$$T_\perp^{B \rightarrow \gamma}(0) = T_\parallel^{B \rightarrow \gamma}(0) \Leftrightarrow \alpha_{T_\perp 0} = \alpha_{T_\parallel 0}. \quad (\text{A.65})$$

- The off-shell FFs are simply parameterised by

$$F_n^{B \rightarrow \gamma^*}(0, k^2) = \left( \alpha_{n0} + \sum_{k=1}^N \alpha_{nk} (z(k^2) - z(0))^k \right), \quad (\text{A.66})$$

The constraint  $V_\parallel^*(0, m_{B_q^*}^2) = 2P^*(0, m_{B_q^*}^2)$  (A.24) is imposed

$$\alpha_{V_\parallel^* 0} + \sum_{k=1}^N \alpha_{V_\parallel^* k} (z(m_{B_q^*}^2) - z(0))^k = 2(\alpha_{P^* 0} + \sum_{k=1}^N \alpha_{P^* k} (z(m_{B_q^*}^2) - z(0))^k). \quad (\text{A.67})$$

The fit-ansatz (A.66) could easily be improved including the information on the  $\rho/\omega$  ( $\phi$ )-like resonances from the dispersion representation (A.50).<sup>15</sup>

<sup>15</sup>The extension to fit the two-variable FF  $F_n^{B \rightarrow \gamma^*}(q^2, k^2)$  is not straightforward but one would best proceed by building an ansatz from a double dispersion relation in  $q^2$  and  $k^2$  and in addition force the constraints (A.9,A.11,A.12).

Let us now describe the  $z$ -expansion in order to remain self-consistent. The function  $z(t)$  is defined by

$$z(t) = \frac{\sqrt{t_+ - t} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - t} + \sqrt{t_+ - t_0}}, \quad (\text{A.68})$$

where  $t_0 \equiv t_+(1 - \sqrt{1 - t_-/t_+})$  and  $t_{\pm} \equiv (m_{B_q} \pm m_\rho)^2$ . The  $\rho$ -mass,  $m_\rho = 770$  MeV, is just a arbitrary reference scale and the values of  $m_{B_q}$  are given in Table 7.

The coefficients  $\alpha_{nk}$  are determined by fitting  $N = 200$  random points at each integer value of  $q_i^2$  (in  $\text{GeV}^2$ -units) in a specific interval. Uncertainties in input parameters,  $p \pm \delta p$ , as for example  $m_b$ , are accounted for by sampling them with a normal distribution  $N(p, \delta p)$ , which accounts for the same correlations as in Ref. [46]. The  $N = 200$  random samples of  $F_I = F_i(q_j^2)$ , where  $I = (i, j)$  denotes the collective index for the FF-type and the momentum, determine the  $(ij) \times (ij)$  covariance matrix

$$C_{IJ} = \langle F_I F_J \rangle - \langle F_I \rangle \langle F_J \rangle. \quad (\text{A.69})$$

Angle brackets denote the average over random samples. The coefficients  $\alpha_{nk}$  are then found by minimising the function

$$\chi^2(\{\alpha\}) = \sum_{IJ} (F_I^{\text{sample}} - F_I^{\text{fit}}(\{\alpha\})) C_{IJ}^{-1} (F_J^{\text{sample}} - F_J^{\text{fit}}(\{\alpha\})), \quad (\text{A.70})$$

for each random sample, where the correlation matrix remains constant for all samples. The fitted values of  $\alpha$  are then averaged over all the samples and errors are calculated from the standard deviation, which is justified because each of the samples are statistically independent.

- The computation of the four on-shell FFs [30] are limited to roughly  $q^2 < 14 \text{ GeV}^2$ . The 200 sample points are generated for each integer interval in  $q^2 \in [-5, 14] \text{ GeV}^2$  to which the  $\alpha_n$ 's are then fitted to the ansatz (A.66).
- Since we only need the off-shell FFs in the region  $k^2 \in [(4.9 \text{ GeV})^2, m_{B_{d,s}}^2] \text{ GeV}^2$  we restrict our fitting procedure to this region.

## References

- [1] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223–226.
- [2] F. Wilczek, *Problem of Strong  $p$  and  $t$  Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279–282.
- [3] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.
- [4] R. D. Peccei and H. R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791–1797.
- [5] J. Preskill, M. B. Wise and F. Wilczek, *Cosmology of the Invisible Axion*, *Phys. Lett.* **B120** (1983) 127–132.
- [6] L. F. Abbott and P. Sikivie, *A Cosmological Bound on the Invisible Axion*, *Phys. Lett.* **B120** (1983) 133–136.

- [7] M. Dine and W. Fischler, *The Not So Harmless Axion*, *Phys. Lett.* **B120** (1983) 137–141.
- [8] P. W. Graham, I. G. Irastorza, S. K. Lamoreaux, A. Lindner and K. A. van Bibber, *Experimental Searches for the Axion and Axion-Like Particles*, *Ann. Rev. Nucl. Part. Sci.* **65** (2015) 485–514, [[1602.00039](#)].
- [9] F. Wilczek, *Axions and Family Symmetry Breaking*, *Phys. Rev. Lett.* **49** (1982) 1549–1552.
- [10] L. Calibbi, F. Goertz, D. Redigolo, R. Ziegler and J. Zupan, *Minimal axion model from flavor*, *Phys. Rev.* **D95** (2017) 095009, [[1612.08040](#)].
- [11] Y. Ema, K. Hamaguchi, T. Moroi and K. Nakayama, *Flaxion: a minimal extension to solve puzzles in the standard model*, *JHEP* **01** (2017) 096, [[1612.05492](#)].
- [12] F. Björkeroth, L. Di Luzio, F. Mescia and E. Nardi,  *$U(1)$  flavour symmetries as Peccei-Quinn symmetries*, *JHEP* **02** (2019) 133, [[1811.09637](#)].
- [13] L. Di Luzio, F. Mescia, E. Nardi, P. Panci and R. Ziegler, *The Astrophobic Axion*, [1712.04940](#).
- [14] F. Björkeroth, L. Di Luzio, F. Mescia, E. Nardi, P. Panci and R. Ziegler, *Axion-electron decoupling in nucleophobic axion models*, [1907.06575](#).
- [15] K. Saikawa and T. T. Yanagida, *Stellar cooling anomalies and variant axion models*, [1907.07662](#).
- [16] J. Martin Camalich, M. Pospelov, H. Vuong, R. Ziegler and J. Zupan, *to appear*.
- [17] R. D. Peccei, *The Strong CP Problem*, *Adv. Ser. Direct. High Energy Phys.* **3** (1989) 503–551.
- [18] M. S. Turner, *Windows on the Axion*, *Phys. Rept.* **197** (1990) 67–97.
- [19] J. L. Feng, T. Moroi, H. Murayama and E. Schnapka, *Third generation familons,  $b$  factories, and neutrino cosmology*, *Phys. Rev.* **D57** (1998) 5875–5892, [[hep-ph/9709411](#)].
- [20] F. Björkeroth, E. J. Chun and S. F. King, *Flavourful Axion Phenomenology*, *JHEP* **08** (2018) 117, [[1806.00660](#)].
- [21] ATLAS collaboration, M. Aaboud et al., *Study of the rare decays of  $B_s^0$  and  $B^0$  mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector*, *JHEP* **04** (2019) 098, [[1812.03017](#)].
- [22] CMS collaboration, S. Chatrchyan et al., *Measurement of the  $B_s^0 \rightarrow \mu^+ \mu^-$  Branching Fraction and Search for  $B^0 \rightarrow \mu^+ \mu^-$  with the CMS Experiment*, *Phys. Rev. Lett.* **111** (2013) 101804, [[1307.5025](#)].
- [23] LHCb collaboration, R. Aaij et al., *Measurement of the  $B_s^0 \rightarrow \mu^+ \mu^-$  branching fraction and effective lifetime and search for  $B^0 \rightarrow \mu^+ \mu^-$  decays*, *Phys. Rev. Lett.* **118** (2017) 191801, [[1703.05747](#)].
- [24] F. Dettori, D. Guadagnoli and M. Reboud,  $B_s^0 \rightarrow \mu^+ \mu^- \gamma$  from  $B_s^0 \rightarrow \mu^+ \mu^-$ , *Phys. Lett.* **B768** (2017) 163–167, [[1610.00629](#)].
- [25] A. A. Alves Junior et al., *Prospects for Measurements with Strange Hadrons at LHCb*, *JHEP* **05** (2019) 048, [[1808.03477](#)].
- [26] LHCb collaboration, R. Aaij et al., *Search for the rare decay  $D^0 \rightarrow \mu^+ \mu^-$* , *Phys. Lett.* **B725** (2013) 15–24, [[1305.5059](#)].

- [27] M. Beneke, T. Feldmann and D. Seidel, *Systematic approach to exclusive  $B \rightarrow Vl^+l^-$ ,  $V\gamma$  decays*, *Nucl. Phys.* **B612** (2001) 25–58, [[hep-ph/0106067](#)].
- [28] J. Lyon and R. Zwicky, *Isospin asymmetries in  $B \rightarrow (K^*, \rho)\gamma/\ell^+\ell^-$  and  $B \rightarrow K\ell^+\ell^-$  in and beyond the standard model*, *Phys. Rev.* **D88** (2013) 094004, [[1305.4797](#)].
- [29] A. Kozachuk, D. Melikhov and N. Nikitin, *Rare FCNC radiative leptonic  $B_{s,d} \rightarrow \gamma l^+l^-$  decays in the standard model*, *Phys. Rev.* **D97** (2018) 053007, [[1712.07926](#)].
- [30] T. Janowski, B. Pullin and R. Zwicky,  *$B_{d,u,s} \rightarrow \gamma$  Form Factors from LCSR at NLO – to appear*.
- [31] T. Janowski, B. Pullin and R. Zwicky,  *$B_q^*, B_{1q} \rightarrow B_q\gamma$  from LCSR at NLO – to appear*.
- [32] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou and M. Steinhauser,  *$B_{s,d} \rightarrow \ell^+\ell^-$  in the Standard Model with Reduced Theoretical Uncertainty*, *Phys. Rev. Lett.* **112** (2014) 101801, [[1311.0903](#)].
- [33] D. Guadagnoli, M. Reboud and R. Zwicky,  *$B_s$  to  $l^+ l^- \gamma$  as a Test of Lepton Flavor Universality*, [[1708.02649](#)].
- [34] J. Lyon and R. Zwicky, *Resonances gone topsy turvy - the charm of QCD or new physics in  $b \rightarrow sl^+\ell^-$ ?*, [[1406.0566](#)].
- [35] D. M. Straub, *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*, [[1810.08132](#)].
- [36] Y. G. Aditya, K. J. Healey and A. A. Petrov, *Faking  $B_s \rightarrow \mu^+\mu^-$* , *Phys. Rev.* **D87** (2013) 074028, [[1212.4166](#)].
- [37] C. Kane, C. Lehner, S. Meinel and A. Soni, *Radiative leptonic decays on the lattice*, in *37th International Symposium on Lattice Field Theory (Lattice 2019) Wuhan, Hubei, China, June 16-22, 2019*, 2019, [[1907.00279](#)].
- [38] C. T. Sachrajda, M. Di Carlo, G. Martinelli, D. Giusti, V. Lubicz, F. Sanfilippo et al., *Radiative corrections to semileptonic decay rates*, in *37th International Symposium on Lattice Field Theory (Lattice 2019) Wuhan, Hubei, China, June 16-22, 2019*, 2019, [[1910.07342](#)].
- [39] LHCb collaboration, R. Aaij et al., *Measurement of the  $b$ -quark production cross-section in 7 and 13 TeV pp collisions*, *Phys. Rev. Lett.* **118** (2017) 052002, [[1612.05140](#)].
- [40] LHCb collaboration, LHCb-CONF-2013-011, *Updated average  $f_s/f_d$   $b$ -hadron production fraction ratio for 7 TeV pp collisions*.
- [41] A. J. Buras, J. Girrbach, D. Guadagnoli and G. Isidori, *On the Standard Model prediction for  $BR(B_{s,d} \rightarrow \mu^+\mu^-)$* , *Eur. Phys. J.* **C72** (2012) 2172, [[1208.0934](#)].
- [42] A. J. Buras, R. Fleischer, J. Girrbach and R. Knegjens, *Probing New Physics with the  $B_s \rightarrow \mu^+\mu^-$  Time-Dependent Rate*, *JHEP* **07** (2013) 77, [[1303.3820](#)].
- [43] R. D. Cousins and V. L. Highland, *Incorporating systematic uncertainties into an upper limit*, *Nucl. Instrum. Meth.* **A320** (1992) 331–335.
- [44] F. Kruger and D. Melikhov, *Gauge invariance and form-factors for the decay  $B \rightarrow \gamma\ell^+\ell^-$* , *Phys. Rev.* **D67** (2003) 034002, [[hep-ph/0208256](#)].

- [45] M. Wirbel, B. Stech and M. Bauer, *Exclusive Semileptonic Decays of Heavy Mesons*, *Z. Phys.* **C29** (1985) 637.
- [46] A. Bharucha, D. M. Straub and R. Zwicky,  *$B \rightarrow V\ell^+\ell^-$  in the Standard Model from light-cone sum rules*, *JHEP* **08** (2016) 098, [[1503.05534](#)].
- [47] D. Melikhov and N. Nikitin, *Rare radiative leptonic decays  $B(d, s) \rightarrow l^+l^-\gamma$* , *Phys. Rev.* **D70** (2004) 114028, [[hep-ph/0410146](#)].
- [48] M. Dimou, J. Lyon and R. Zwicky, *Exclusive Chromomagnetism in heavy-to-light FCNCs*, *Phys. Rev.* **D87** (2013) 074008, [[1212.2242](#)].
- [49] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 2005.
- [50] R. Zwicky, *A brief Introduction to Dispersion Relations and Analyticity*, in *Proceedings, Quantum Field Theory at the Limits: from Strong Fields to Heavy Quarks (HQ 2016): Dubna, Russia, July 18-30, 2016*, pp. 93–120, 2017, [1610.06090](#), DOI.
- [51] V. Prochazka and R. Zwicky, *Finiteness of two- and three-point functions and the renormalization group*, *Phys. Rev.* **D95** (2017) 065027, [[1611.01367](#)].
- [52] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, *QCD and Resonance Physics. Theoretical Foundations*, *Nucl. Phys.* **B147** (1979) 385–447.
- [53] R. Mertig, M. Bohm and A. Denner, *FEYN CALC: Computer algebraic calculation of Feynman amplitudes*, *Comput. Phys. Commun.* **64** (1991) 345–359.
- [54] V. Shtabovenko, R. Mertig and F. Orellana, *New Developments in FeynCalc 9.0*, *Comput. Phys. Commun.* **207** (2016) 432–444, [[1601.01167](#)].
- [55] PARTICLE DATA GROUP collaboration, M. e. a. Tanabashi, *Review of particle physics*, *Phys. Rev. D* **98** (Aug, 2018) 030001.
- [56] M. Jamin and B. O. Lange,  *$f(B)$  and  $f(B(s))$  from QCD sum rules*, *Phys. Rev.* **D65** (2002) 056005, [[hep-ph/0108135](#)].