Direct Determination of the Lepton Flavour Universality Observables $\Delta\mathcal{C}_{9,10}$ in $B^0 \to K^{*0}\ell^+\ell^-$ Decays

---

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)
vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Michele Atzeni
aus
Italien

Promotionskommission

Prof. Dr. Nicola Serra (Vorsitz)
PD Dr. Olaf Steinkamp
Dr. Katharina Müller
Prof. Rafael Silva Coutinho

Zürich, 2023

A Marta, a mia madre e a mio padre,
per avermi sostenuto durante questi anni
e avermi mostrato il valore della pazienza e del duro lavoro.

Abstract

The Standard Model of particle physics is a theory that has been remarkably successful in describing the behaviour of the microscopic particles that constitute the world around us. However, experimental evidences such as the non-zero masses of the neutrinos and the presence of dark matter and dark energy in the universe, clearly indicate that this theory needs to be extended.

Flavour changing neutral current transitions are forbidden at the tree level in the Standard Model and are therefore rare. Contributions coming from yet unobserved interactions, could be comparable in size with the Standard Model amplitudes and be observed indirectly as shifts in some physical observables with respect to their predictions. This class of decays is therefore considered an ideal venue to probe the limits of the Standard Model.

In recent years, some tensions have indeed been observed between the measurement of branching ratios and angular observables in $b \to s\ell^+\ell^-$ transitions and their corresponding predictions. The nature of these tensions, still below the discovery threshold of five standard deviations, could however be explained by large hadronic contribution difficult to estimate reliably from first principles, which could mimic the effect of a new interaction. Lepton flavour universality tests are in these regards protected by the inherent lepton flavour universality of Quantum Chromodynamics, and currently provide the cleanest benchmark to test the Standard Model in semileptonic decays. An update on the ratio of branching ratio between $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ decays provided by the LHCb experiment in 2021 has reported the first evidence of lepton flavour universality violation in flavour changing neutral currents, further sparking interest in this class of decays.

This thesis presents the first attempt to perform a simultaneous amplitude analysis of the decays $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ to extract, in a single direct measurement, the difference between the observables that encode potential non-equal couplings between muons and electrons, $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$. This lepton flavour universality test naturally combines the ratio of branching fractions and the differences of angular observables between $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays as a function of dilepton invariant mass squared, providing an unprecedented discovery potential.

The measurement is performed using the $9\,\text{fb}^{-1}$ collected by the LHCb experiment between the years 2011 and 2018, and is limited to the region of dilepton invariant mass squared between 1.1 and $7\,\text{GeV}/c^2$. While the fit results of this analysis are currently blinded, an estimation of the realistic statistical uncertainty obtained directly from a likelihood profile on data, together with the sensitivity obtained from realistic toys, is provided as the main result of this work. Additionally, an heuristic estimation of the dominant systematic uncertainties expected is also discussed, suggesting that with the data currently available the sensitivity of the measurement will be dominated by the statistical uncertainty. The result, when unblinded, will provide the most sensitive measurement of $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ obtained from $B^0 \to K^{*0}\ell^+\ell^-$ decays.

Contents

Chapter 1

Theoretical background

This chapter aims to provide a short overview of the key features of the Standard Model (SM), a theory that in the last 60 years has proven to describe particle physics phenomena with astonishing accuracy. After a short introduction where the main components of the SM are presented, it follows a more detailed and formal presentation describing how the theory is built, with a particular attention to the features more relevant for the physics of flavour. In conclusion, an outline of some of the limitations of the SM is provided.

1.1 Introduction

The Standard Model of particle physics has been one of the most successful theories of all time, capable of describing under the same formalism three of the four fundamental forces of nature: the weak, electromagnetic and strong force. The gravitational force, currently not included in the SM, has a strength that is much smaller than the other forces and has therefore no relevance for particle physics at the energy scales presently explored. After years of experimental work, culminated by the discovery of the Higgs boson in 2012[1, 2], the physics community settled on a set of particles that are considered to be fundamental, *i.e.* with no internal structure. These particles, shown schematically in Figure 1.1, can be sorted by their properties.

- **Fermions**. These are particles with spin 1/2 that follow the Fermi-Dirac statistics. They can be further divided into quark and leptons. The first ones correspond to the six quarks u, d, c, s, t and b and experience all three forces in the SM, since they have weak isospin, electric and colour charge. Leptons do not take part in strong interactions as they are colour neutral. The three charged leptons e, μ and τ participate in the electromagnetic and weak interactions, while the neutrinos ν_e, ν_μ, ν_τ only interact via the weak force. Both quark and leptons can be further categorized in three generations. The first generation constitutes the great majority of all the matter in the Universe, for this reason quarks and leptons are also referred to as matter fields. The second and third generations are exact copies of the first one, with the difference that they are heavier and thus unstable. Flavour physics is the branch of particle physics that studies transitions between these different generations.

- **Bosons**. These are particles with integer spin that follow the Bose-Einstein statistics. They can be divided into gauge bosons, with spin 1, and scalar bosons with spin 0. The first ones are force mediators and correspond to the Z^0 and $W^\pm$, that mediate the weak interactions, the photon γ,

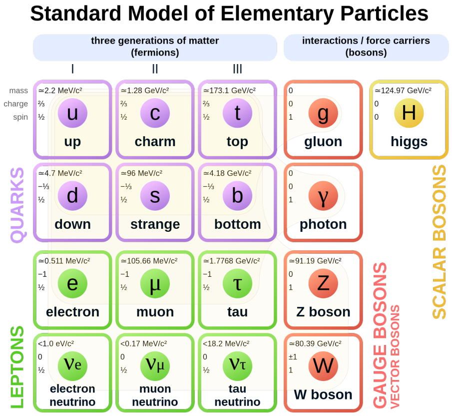


Figure 1.1: Schematic view of the fundamental particles of the Standard Model [3].

responsible for electromagnetic interactions, and the gluons, mediators of the strong force. Of the four gauge bosons only $W^{\pm}$ are not electrically neutral. The Higgs boson is the only elementary particle that is a scalar boson. This particle is responsible for the mass of all fermions and gauge bosons except for gluons and photons.

1.2 The Standard Model of particle physics

All properties outlined in Sec. 1.1 can be mathematically formalized in the context of a Quantum Field Theory (QFT), where elementary particles are described as excitations of the corresponding quantum fields. The SM can then be built starting from a free Lagrangian of the matter fields, in which the interactions are introduced by transforming a global symmetry of the theory into a local (gauge) symmetry by means of a covariant derivative; an approach usually referred to as *minimal coupling*. To perform this transformation, additional fields with spin 1 must be introduced, the so-called gauge bosons. With this prescription, the SM is then built starting from three main pillars:

- the symmetry (gauge group) of the theory: $SU(3)_C \otimes SU(2)_W \otimes U(1)_Y$, associated to the colour, weak isospin and hypercharge, respectively;
- the representations of the matter fields under those symmetries;
- the spontaneous symmetry breaking mechanism (SSB).

This last step is fundamental to make the theory phenomenologically accurate, since it is responsible for the breaking of $SU(2)_W \otimes U(1)_Y \rightarrow U(1)_{em}$, which gives rise to electromagnetism and to the dynamical generation of the mass terms for the matter fields, Z^0 and $W^\pm$ gauge bosons. The matter fields have strong and electroweak charge and are therefore grouped in the multiplets of the gauge group. The leptonic fields

$$\begin{aligned} E_{iL} =& \begin{pmatrix} \nu_{iL} = \frac{1}{2}(1-\gamma^5)\nu_i \\ e_{iL} = \frac{1}{2}(1-\gamma^5)e_i \end{pmatrix} && (1,2)_{-1/2}\,, \\ e_{iR} =& \frac{1}{2}(1+\gamma^5)e_i && (1,1)_{-1}\,, \end{aligned} \tag{1.1}$$

correspond, when considering the first generation ($i = 1$), to an electron and an electronic neutrino with negative helicity[1] and to an electron with positive helicity. These fields describe also the corresponding antiparticles: a positron and positronic antineutrino with positive helicity and a positron with negative helicity. The column on the right in Eq. 1.1 explicitly states, between parenthesis, under which representation the different fields transform with respect to the symmetries $(SU(3)_C, SU(2)_W)$, while the subscript corresponds to the value of the hypercharge linked to the $U(1)_Y$ symmetry. The matrices $\frac{1}{2}(1-\gamma^5)$ and $\frac{1}{2}(1+\gamma^5)$ correspond to the left- and right-handed projectors P_L and P_R, respectively.

The quark fields are composed of two types for each generation, up and down, grouped as

$$\begin{aligned} Q_{iL} =& \begin{pmatrix} u_{iL} \\ d_{iL} \end{pmatrix} && (3,2)_{1/6}\,, \\ u_{iR}& && (3,1)_{2/3}\,, \\ d_{iR}& && (3,1)_{-1/3}\,. \end{aligned} \tag{1.2}$$

In this notation, parity is clearly violated, since different weak quantum numbers are assigned to left- and right-handed components of the matter fields. The left-handed fields are doublets of $SU(2)_W$, while right-handed fields are singlets. In other words, the first ones interact weakly, while the second ones do not.

Starting from these premises, it is possible to write the most general renormalizable Lagrangian density $\mathcal{L}$ that contains the required fields and is compatible with Lorentz and gauge invariance:[2]

$$\begin{aligned} \mathcal{L}_{SM} = & -\frac{1}{4}\sum_{a=1}^{8} F^{a\mu\nu}F^a_{\mu\nu} - \frac{1}{4}\sum_{i=1}^{3} A^{i\mu\nu}A^i_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \\ & + (D^\mu\varphi)^\dagger D_\mu\varphi + \mu^2\varphi^\dagger\varphi - \lambda(\varphi^\dagger\varphi)^2 \\ & + i\bar{E}_{iL}\gamma^\mu D_\mu E_{iL} + i\bar{Q}_{iL}\gamma^\mu D_\mu Q_{iL} \end{aligned}$$

---

[1] The helicity of a particle is the projection of its spin onto the direction of its momentum. A particle has a positive helicity when the direction of its spin is the same as the direction of its motion. It is important to point out that this quantity is Lorentz-invariant only for massless particles, since in this case it is equivalent to their chirality.

[2] For simplicity, the term in the gluon fields $-\frac{\theta_s}{64\pi^2}\sum_{a=1}^{8} F^{a\mu\nu}\epsilon_{\mu\nu\rho\sigma}F^{a\rho\sigma}$ has been ignored since, in the SM, the coeffiecient θ_s is compatible with zero [4].

$$+ i\bar{e}_{iR}\gamma^{\mu}D_{\mu}e_{iR} + i\bar{u}_{iR}\gamma^{\mu}D_{\mu}u_{iR} + i\bar{d}_{iR}\gamma^{\mu}D_{\mu}d_{iR}$$
$$- \Lambda_{e,ij}[\bar{E}_{iL}\varphi]e_{jR} - \Lambda_{d,ij}[\bar{Q}_{iL}\varphi]d_{jR} - \Lambda_{u,ij}[\bar{Q}_{iL}\varphi^{c}]u_{jR} + h.c. \tag{1.3}$$

The first row contains the kinematic terms that describe the free propagation in space-time and self-interaction of the strong, weak and hypercharge gauge bosons fields, defined as

$$G^{a}_{\mu}(8,1)_{0}, \quad A^{a}_{\mu}(1,3)_{0}, \quad B_{\mu}(1,1)_{0}\,, \tag{1.4}$$

and the corresponding field strengths given by

$$\begin{aligned} F^{a\mu\nu} &= \partial^{\mu}G^{a\nu} - \partial^{\nu}G^{a\mu} - g_{s}f^{abc}G^{b\mu}G^{c\nu}\,, \\ A^{i\mu\nu} &= \partial^{\mu}A^{i\nu} - \partial^{\nu}A^{i\mu} - g\epsilon^{abc}A^{b\mu}A^{c\nu}\,, \\ B^{\mu\nu} &= \partial^{\mu}B^{\nu} - \partial^{\nu}B^{\mu}\,, \end{aligned} \tag{1.5}$$

where g_s (g) is the strong (electroweak) coupling constant, and f^{abc} (ϵ^{abc}) are the structure constants of $SU(3)_C$ ($SU(2)_W$). The second row contains the Higgs sector, expressed as a function of the scalar field $\varphi(1,2)_{1/2}$, and responsible for the spontaneous breaking of the electroweak symmetry. The third and fourth rows contain all the matter fields (leptons and quarks) coupled minimally to the gauge fields through the covariant derivative

$$D_{\mu}\psi_{i} = \partial_{\mu}\psi_{i} + igA^{i}_{\mu}T^{i}\psi_{i} + iyg'B_{\mu}\psi_{i} + ig_{s}G^{a}_{\mu}t^{a}_{s}\psi_{i}, \tag{1.6}$$

where

- t^a_s are the $SU(3)_C$ generators (3×3 Gell-Mann matrices $\lambda_a/2$ for triplets, 0 for singlets),
- T^i are the $SU(2)_W$ generators (2×2 Pauli matrices $\sigma_i/2$ for doublets, 0 for singlets),
- y is the $U(1)_Y$ hypercharge and g' is the weak hypercharge coupling constant.

The fifth row contains the Yukawa couplings between the matter fields and the Higgs boson. Notice that the field $\varphi^c = i\sigma^2\varphi^*$ is used instead of φ^* in the Lagrangian density. The advantage of this notation is related to the fact that φ^c corresponds to a Higgs field with hypercharge of $-\frac{1}{2}$ rather than $\frac{1}{2}$, keeping manifest the invariance under hypercharge of the expression. This term is also responsible for the breaking of the *global flavour symmetry* of the gauge part of the Lagrangian and for the transition between different fermion generations.

The Lagrangian density, written in this very compact form, completely describes the SM, and it clearly shows the symmetries of the theory before the SSB takes place.

1.2.1 SSB, Higgs and the gauge fields

The phenomenon of spontaneous breaking of a symmetry happens when a physical system, whose dynamics is invariant with respect to a continuous symmetry group $\mathcal{G}$, does not have a unique state of minimum energy (void state or configuration) but a continuos family of configurations of minimal energy that is invariant only with respect to a subgroup of $\mathcal{G}$. In the case of the Higgs field, its void configuration can be found by studying the minima of the potential introduced in Eq. 1.3

$$V(\varphi) = -\mu^2\varphi^\dagger\varphi + \lambda(\varphi^\dagger\varphi)^2 \to \varphi^\dagger\varphi = \frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}\,. \tag{1.7}$$

This relationship fixes the modulus of the complex vector at two components φ, partially breaking the symmetry of the potential. Without loss of generality, the state of minimum energy for the Higgs fields can be chosen as

$$\varphi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{1.8}$$

Expanding the Higgs field φ around these minima and using the gauge transformation $\varphi \to e^{i\alpha^j(x)\frac{\sigma^j}{2}+i\beta(x)/2}\varphi$ (*unitary gauge*) to remove three of the four components of φ, the fields can be written as

$$\varphi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{1.9}$$

The potential term of the Higgs then becomes:

$$V(h) = \lambda v^2 h^2 + \lambda v h^3 + \frac{\lambda}{4}h^4, \tag{1.10}$$

which describes a massive field of mass $m_h^2 = 2\lambda v^2$, characterized by cubic and quartic self-interaction vertices. As a consequence, the couplings between the Higgs and the gauge bosons are simplified to:

$$(D^\mu\varphi)^\dagger D_\mu\varphi = \frac{1}{2}\partial^\mu h \partial_\mu h + \frac{v^2}{8}\left[g^2 A^1_\mu A^{1\mu} + g^2 A^2_\mu A^{2\mu}\right. \tag{1.11}$$

$$\left. + \left(gA^3_\mu - g'B_\mu\right)\left(gA^{3\mu} - g'B^\mu\right)\right]\left(1 + \frac{h}{v}\right)^2. \tag{1.12}$$

The purely quadratic terms in the second part of Eq. 1.11 can be interpreted as mass terms produced from the interaction of the Higgs field with the electroweak gauge fields. The mass spectrum of these fields after the spontaneous symmetry breaking appears clearly when performing the following transformations

$$\begin{aligned} W^{\pm} &= \frac{1}{\sqrt{2}}(A^1 \mp iA^2) \quad &\rightarrow \quad m_W^2 &= \frac{v^2}{4}g^2, \\ Z &= \frac{1}{\sqrt{g^2+g'^2}}(gA^3 - g'B) \quad &\rightarrow \quad m_Z^2 &= \frac{v^2}{4}(g^2+g'^2), \\ A &= \frac{1}{\sqrt{g^2+g'^2}}(g'A^3 + gB) \quad &\rightarrow \quad m_\gamma^2 &= 0. \end{aligned} \tag{1.13}$$

In this base, the Lagrangian term then takes the following simple form:

$$(D^\mu \varphi)^\dagger D_\mu \varphi = \frac{1}{2}\partial^\mu h \partial_\mu h + \left[m_W^2 W^+ W^- + \frac{m_Z^2 Z^2}{2} \right] (1 + \frac{2h}{v} + \frac{h^2}{v^2}) \,. \tag{1.14}$$

By recognizing that the field component that stays massless corresponds to the electromagnetic field, A_μ, it is possible to define the electric charge Q and the electromagnetic coupling constant e as

$$Q = T^3 + Y \quad \text{and} \quad e = \frac{gg'}{\sqrt{g^2+g'^2}} \,, \tag{1.15}$$

and rewrite the electroweak part of the covariant derivative as

$$D_\mu \supset i\frac{g}{\sqrt{2}}\left(W_\mu^+ T^+ + W_\mu^- T^-\right) + i\frac{g}{\cos\theta_W}\left(T^3 - \sin^2\theta_W Q\right) Z_\mu + ieQA_\mu \,, \tag{1.16}$$

where $\sin\theta_W = \frac{g'}{\sqrt{g'^2+g^2}}$, $\cos\theta_W = \frac{g}{\sqrt{g'^2+g^2}}$ and θ_W is usually referred to as the Weinberg mixing angle.

1.2.2 Yukawa couplings

The last row in Eq. 1.3 is also modified by the choice of unitary gauge, becoming

$$-\mathcal{L}_{\text{Yuk}} = \tfrac{v}{2}\Lambda_{e,ij}\bar{e}_{iL}e_{jR}\left(1+\tfrac{h}{v}\right) + \tfrac{v}{2}\Lambda_{d,ij}\bar{d}_{iL}d_{jR}\left(1+\tfrac{h}{v}\right) + \tfrac{v}{2}\Lambda_{u,ij}\bar{u}_{iL}u_{jR}\left(1+\tfrac{h}{v}\right) + h.c. \tag{1.17}$$

The matrices $\Lambda_{e,ij}$, $\Lambda_{d,ij}$ and $\Lambda_{u,ij}$ are generic complex matrices in *flavour space* that can allow interactions between different lepton and quark generations. This expression can be further simplified with the diagonalization of these matrices by means of a bi-unitary transformation $\Lambda = VDW$, where V and W are unitary[3] matrices. By performing the following transformations:

$$\begin{aligned} \Lambda_e &= V_e \,\text{diag}(\lambda_e, \lambda_\mu, \lambda_\tau)\, W_e \,, \\ \Lambda_u &= V_u \,\text{diag}(\lambda_u, \lambda_c, \lambda_t)\, W_u \,, \\ \Lambda_d &= V_d \,\text{diag}(\lambda_d, \lambda_s, \lambda_b)\, W_d \,, \end{aligned} \tag{1.18}$$

---

[3] A matrix U is unitary if $U^\dagger U = UU^\dagger = I$.

and

$$\begin{aligned} e'_R &= W_e e_R, \quad d'_R = W_d d_R, \quad u'_R = W_u u_R\,, \\ e'_L &= V_e^\dagger e_L, \quad d'_L = V_d^\dagger d_L, \quad u'_L = V_u^\dagger u_L\,, \end{aligned} \tag{1.19}$$

Eq. 1.17 can be rewritten in the simplified form

$$-\mathcal{L}_{\text{Yuk}} = \tfrac{v}{2}\lambda_{e,i}\bar{e}_{iL}e_{iR}\left(1+\tfrac{h}{v}\right) + \tfrac{v}{2}\lambda_{d,i}\bar{d}_{iL}d_{iR}\left(1+\tfrac{h}{v}\right) + \lambda_{u,i}\bar{u}_{iL}u_{iR}\left(1+\tfrac{h}{v}\right) + h.c. \tag{1.20}$$

The fields redefined in Eq. 1.19 diagonalize the quadratic part of the SM and are for this reason also called *mass eigenstates* of the fermion fields. As it can be seen, only 9 real parameters survived: $m_{f,i} = \frac{v}{2}\lambda_{f,i}$, where $f = e, d, u$ and $i = 1, 2, 3$. Additionally, since the Yukawa matrices Λ have been diagonalized, no interaction between the different fermion generations is possible in this part of the Lagrangian.

1.2.3 Charged and neutral currents

With the updated definition of the covariant derivative shown in Eq. 1.16, it is possible to partially reorganize the terms describing the interaction between the matter and gauge fields that appear in the third and fourth line of Eq. 1.3. Neutral currents can be written as

$$\mathcal{L}_{\text{NC}} = -gZ_\mu J_n^\mu - eJ_{em,\mu}A^\mu, \tag{1.21}$$

where

$$\begin{aligned} J_n^\mu =& \frac{1}{\cos\theta_W}\Big[(-\frac{1}{2}+\sin^2\theta_W)\bar{e}_L\gamma^\mu e_L + \frac{1}{2}\bar{\nu}_L\gamma^\mu\nu_L \\ &+ (\frac{1}{2}-\frac{2}{3}\sin^2\theta_W)\bar{u}_L\gamma^\mu u_L + (-\frac{1}{2}+\frac{1}{3}\sin^2\theta_W)\bar{d}_L\gamma^\mu d_L \\ &+ \sin^2\theta_W\bar{e}_R\gamma^\mu e_R - \frac{2}{3}\sin^2\theta_W\bar{u}_R\gamma^\mu u_R + \frac{1}{3}\sin^2\theta_W\bar{d}_R\gamma^\mu d_R\Big]\,, \\ J_{em}^\mu =& -\bar{e}\gamma^\mu e + \frac{2}{3}\bar{u}\gamma^\mu u - \frac{1}{3}\bar{d}\gamma^\mu d\,, \end{aligned} \tag{1.22}$$

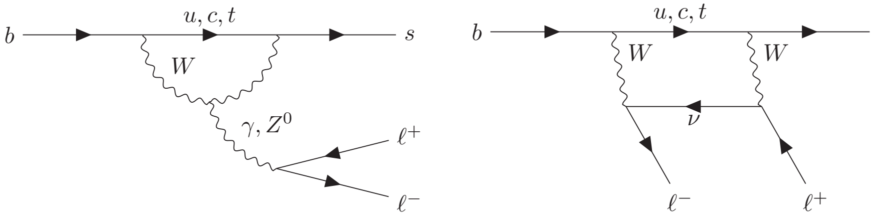
while charged currents are described by

$$\mathcal{L}_{\text{CC}} = -g(J_\mu^+ W^{+\mu} + J_\mu^- W^{-\mu})\,, \tag{1.23}$$

where

$$\begin{aligned} J^{+\mu} &= \frac{1}{\sqrt{2}}(\bar{\nu}_L\gamma^\mu e_L + \bar{u}_L\gamma^\mu d_L)\,, \\ J^{-\mu} &= \frac{1}{\sqrt{2}}(\bar{e}_L\gamma^\mu\nu_L + \bar{d}_L\gamma^\mu u_L)\,. \end{aligned} \tag{1.24}$$

The change of basis in Eqs. 1.18 and 1.19, which greatly simplified the Yukawa sector, must be propagated to the remaining parts of $\mathcal{L}_{SM}$. In the SM, the only



(a) “Penguin” diagram (b) “Box” diagram

Figure 1.2: Feynman diagrams corresponding to the “penguin” and “box” loop level diagrams used to describe the $b \to s\,\ell^+\ell^-$ FCNC transition in the SM.

terms affected are the charged current interactions between matter fields and gauge boson of Eq. 1.24, which transform into:

$$\begin{aligned} J^{+\mu} &= \frac{1}{\sqrt{2}}(\bar{\nu}_L \gamma^\mu V_e e'_L + \bar{u}'_L \gamma^\mu V_u^\dagger V_d d'_L)\,, \\ J^{-\mu} &= \frac{1}{\sqrt{2}}(\bar{e}'_L \gamma^\mu V_e^\dagger \nu_L + \bar{d}'_L \gamma^\mu V_d^\dagger V_u u'_L)\,. \end{aligned} \tag{1.25}$$

In the leptonic charge current, the matrix V_e can be removed simply by the redefinition $\nu_L = V_e \nu'_L$, and no mixing between leptons of different generations is possible. On the contrary, in the quarkonic charged current, the unitary matrix $V_{CKM} = V_u^\dagger V_d$, called Cabibbo-Kobayashi-Maskawa matrix (CKM), cannot be removed and transitions between different quark generations are possible. It is important to stress that, even though the CKM matrix appears explicitly in charged current interactions, it originated in the Yukawa sector, specifically due to the misalignment in flavour space of the matrices Λ_d and Λ_u.

Charged current interactions are the only processes through which a quark can change its flavour in the SM. As a consequence, flavour changing neutral currents (FCNC) are forbidden at the tree level and can only occur at loop level, as a combination of charged currents and other interactions. These transitions are mediated by the so-called “penguin” or “box” diagrams, two examples of which are shown in Figure 1.2. Being loop suppressed, these interactions are expected to have small branching ratios. This makes FCNC an ideal venue to search for potential New Physics (NP) contributions, since the interference between SM and NP amplitudes could be observed indirectly as a shift of the measured observables with respect to the corresponding SM prediction. It is for this reason that FCNC in $b \to s\,\ell^+\ell^-$ transitions have been chosen as the main topic of this thesis.

1.2.4 CKM matrix

The CKM matrix is a generic unitary 3×3 complex matrix that depends on three real rotational angles, θ_{ij}, and a single phase, δ, sole responsible for CP violation in the SM. The standard parametrization of the CKM matrix[5] is

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -s_{23}c_{12} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{1.26}$$

with $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. The off-diagonal elements of the CKM matrix show a strongly hierarchical pattern: $|V_{cd}|$ and $|V_{us}|$ are close to 0.22, the elements $|V_{ts}|$ and $|V_{cb}|$ are of order $4 \cdot 10^{-2}$, whereas $|V_{ub}|$ and $|V_{td}|$ are of order $5 \cdot 10^{-3}$. The Wolfenstein parametrization[6] is a convenient way to exhibit this hierarchy:

$$\begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.27}$$

where $\lambda = |V_{us}| \approx 0.22$, A, ρ and η are free parameters of $\mathcal{O}(1)$.

The unitarity of the CKM matrix implies the following relations between its elements:

$$\sum_{k=1..3} V^*_{ik}V_{ki} = 1, \qquad \sum_{k=1..3} V^*_{ik}V_{kj\neq i} = 0\,, \tag{1.28}$$

Among the second type of relations, the one obtained for $i = 1$ and $j = 3$, namely

$$\frac{V_{ud}V^*_{ub}}{V_{cd}V^*_{cb}} + \frac{V_{td}V^*_{tb}}{V_{cd}V^*_{cb}} + 1 = 0\,, \tag{1.29}$$

is particularly interesting since it involves the sum of three terms all of the same order in λ. For this reason, it is usually represented as a triangle in the complex plane, referred to as *unitarity triangle*, and shown in Figure 1.3 on the left. Both angles and sides of the unitarity triangles are observable quantities that can be determined experimentally, with multiple decays that can be used to access the same parameter. The possibility to overconstrain the CKM elements is an important goal of flavour physics, since it could possibly reveal new sources of flavour symmetry breaking. On the right, Figure 1.3 shows the global fit performed on the parameters of the unitarity triangle by the `CKMfitter` collaboration [7](see Ref. [8] for a Bayesian approach), including results from charmless semileptonic B decays ($|V_{ub}|$), mass differences in the B_d (Δm_d) and B_s (Δm_s) systems, CP violation in the neutral kaon system (ϵ_K) and in the B_d systems ($\sin 2\beta$), and the combined constraints on α and γ from various B decays. As can be seen the available phase space for non-SM contributions in flavour changing transitions is small.

1.2.5 Standard model after SSB

The SM, ignoring neutrino masses, depends on 19 parameters: three coupling constants (g_s, g and g'), a parameter that governs CP violation in the strong

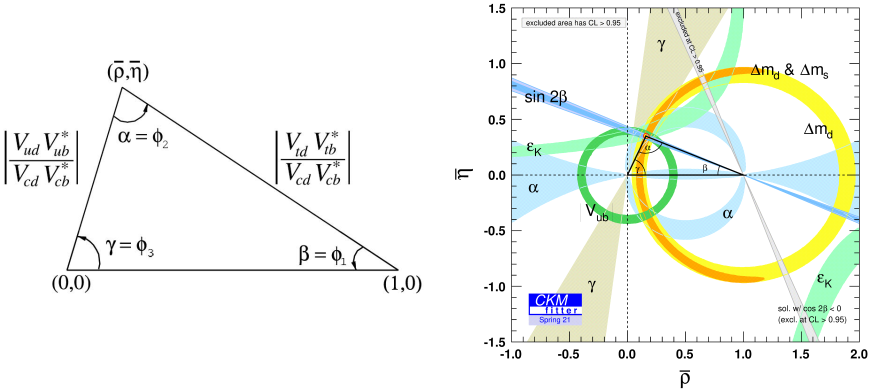


Figure 1.3: On the left, a schematic view of the CKM unitarity triangle. On the right, the result of the global fit performed on the parameters of the unitarity triangle (see Ref. [7]). The individual constraints used in the global fit and described in the text are also shown.

sector, two parameters for the Higgs potential (μ^2[4] and λ[5]), nine Yukawa parameters corresponding to the masses of six quarks and three charged leptons, four parameters from the CKM matrix, three angles and a phase.

In addition to the symmetries of the gauge group, the SM is invariant under CPT, as required by any relativistic theory. Charge conjugation and parity are violated fundamentally by the chiral nature of the electroweak interactions, while the violation of CP is more subtle: it only happens in the weak sector due to the complex phase of the CKM matrix. This is made possible by the fact that the number of quark generations is greater than two. Additionally, there are four accidental [6] symmetries in the SM: leptonic number conservation, separately for each lepton generation, and baryon number conservation. When accounting for neutrino mixing, the first three are broken into generic lepton number conservation.

Even if it has proven to be one of the most accurate theories ever built, the SM cannot be the fundamental Lagrangian of Nature, since it fails to provide answers for some of its peculiar characteristics and it does not explain all the experimental observations currently available, the most important of which are summarized in the following:

- *Gravity* - Gravity is currently not included in the SM, but it is only natural to think that these two theories are just two faces of a more general Grand

---

[4] This is the only dimensional parameter of the SM. All the others are adimensional.

[5] Different from the parameter in V_{CKM} in the Wolfenstein parametrization.

[6] Accidental symmetries correspond to global symmetries of an effective theory (SM) that are not imposed as constraints in the determination of the Lagrangian density, but emerge as a consequence of gauge and Lorentz invariance.

Unified Theory (GUT).

- *Three generations* - The number of lepton and quark generations is chosen to fit the experimental evidence: there is no fundamental reason bounding the number of generations to three. Additional generations are currently excluded by the study of Higgs decays [9] and by the measurement of the decay width of the Z boson at LEP [10].

- *The hierarchy problem* - The observed value of the Higgs mass is much smaller than the Plank mass. This is unexpected since this mass is not protected by any symmetry and it should get quantum corrections proportional to the fundamental scale.

- *Matter-Antimatter asymmetry* - While matter is ubiquitous, antimatter can only be briefly produced in collision experiments or in the atmosphere. However, during the Big Bang, the same amount of matter and antimatter must have been created out of the energy available in the early Universe. The amount of CP violation allowed in the SM does not seem sufficient to explain the current abundance of matter over antimatter.

- *Neutrino masses* - Neutrinos are massless in the SM. However, recent experimental measurements [11, 12] have shown that oscillations between different generations are possible, implying a non-zero difference in the masses of neutrinos. Neutrino masses can be added to the SM by introducing right-handed neutrinos (mirroring the mechanism already seen in the quark sector) or by many other mechanisms as reported for example in Ref. [13].

- *Flavour problem* - A large mass hierarchy is present between the lepton and quark generations, spanning over six orders of magnitude between the electron and top quark. If the masses of the neutrinos are included, the mass difference between the lightest and heaviest fermion increases of six additional orders. No explanation for such dispersion is present in the SM.

- *Dark matter and Dark energy* - Looking at the Universe from Earth, only 5% of the mass-energy out there is described by the SM. Of the remaining, few things are known: it interacts gravitationally but does not interact with photons. From this point of view, it seems that the great majority of matter and energy in the Universe are described by a set of rules just waiting to be unravelled.

For these (and more) reasons, the Standard Model is regarded by most physicist as an *effective theory*, a low energy approximation of a more fundamental theory above the TeV scale. Currently, two main strategies seem to be able to improve our understanding of the possible extensions of the SM to be considered: direct and indirect searches. The first ones aim to produce new particles *on-shell* and study their decays within the detector. They have the advantage of a clear interpretation but are limited by the centre-of-mass energy of the experiment.

The second ones exploit the fact that new particles can participate in some decays as additional quantum fluctuations, slightly modifying the expected value of an observable with respect to the SM prediction. The new particles can therefore leave a trace on some specific interactions even if the energy involved in the process is lower than the energy necessary to produce them. The disadvantage is that, to see such small contribution, high precision from both theory predictions and experiments is required. However, such feats have been proven possible in the past, where the existence of the charm and top quark had been predicted to improve the description of the data in the $K_L \to \mu^+\mu-$ decay and in B^0-$\overline{B}^0$ oscillations, respectively. Particularly interesting for this purpose is the study of decays that are greatly suppressed or prohibited in the SM: in that case NP effects could compete or be the only mechanism through which a certain process can be observed.

Chapter 2

Rare $B^0 \to K^{*0}\ell^+\ell^-$ decays as probes for New Physics

Flavour changing neutral current decays are considered rare in the SM due to their branching ratios of the order of 10^{-6}. As discussed in Ch. 1, they are forbidden at tree level in the SM and can only occur at loop-level. Additionally, the CKM matrix is approximately diagonal, further suppressing these generation-changing transitions. These decays are expected to be particularly sensitive to some NP scenarios, since the additional NP amplitudes could compete in size with the SM ones, producing a sizable shift in some of the observables of the decay. This chapter aims to provide an overview of the theoretical and experimental knowledge currently available on FCNC in B decays and is divided as follows: Section 2.1 outlines the difficulties of describing these decays with the full SM Lagrangian of Eq. 1.3, and introduces the formalism of effective field theories in the context of rare $b \to s\,\ell^+\ell^-$ transitions; Section 2.2 exploits the effective Lagrangian obtained to describe the differential decay rate of $B^0 \to K^{*0}\ell^+\ell^-$ decays and Section 2.3 concludes this chapter describing the measurements currently available for this channel and how they relate to each other.

2.1 Weak interactions as an Effective Field Theory

Weak decays of B mesons are characterized by three main different energy scales μ:

- $\mu \approx M_W(\mathcal{O}(80\,\mathrm{GeV}/c^2))$ - The scale at which the weak transition between quarks responsible for the decay of mesons takes place.
- $\mu \approx m_b(\mathcal{O}(5\,\mathrm{GeV}/c^2))$ - The typical energy scale of the initial and final state of the particles involved in the decay.
- $\mu \approx \Lambda_{QCD}(\mathcal{O}(200\,\mathrm{MeV}/c^2))$ - The typical energy scale at which the confinement effects binding the quarks into the mesons need to be taken into account.

The presence of these widely separated energy scales greatly complicates the calculations in the full SM Lagrangian because large logarithms of the type $\log(M_W/\Lambda_{QCD})$ may appear, breaking the ordinary perturbation theory. This problem can be greatly simplified by deriving an effective field theory (EFT) at low-energy that allows to efficiently calculate the amplitudes of weak decays by taking into account the dynamical effects at all scales.

The fundamental idea is that, in the description of a physical system, it is enough to focus on the degrees of freedom relevant at the energy scale under consideration, while the degrees of freedom relative to much smaller scales do not appear explicitly and their integrated effect can be considered instead. Consider the generic amplitude $\langle f|i\rangle$, written using the path integral formalism [14]:

$$\langle f|i\rangle = \int \mathcal{D}\vec{\varphi} e^{\frac{i}{\hbar} S_{\text{full}}(\vec{\varphi})} , \tag{2.1}$$

where $\vec{\varphi}$ is the set of fields that enter the full theory described by the action S_{full}, and $\int \mathcal{D}\vec{\varphi}$ represents an integration from the initial configuration of the fields $\vec{\varphi}_i$ to the final configuration $\vec{\varphi}_f$. The fields $\vec{\varphi}$ can be divided between heavy $(\vec{\varphi}_H)$ and light $(\vec{\varphi}_L)$ fields, depending on the energy scale of the process considered. The integration of the heavy fields can then be formally performed, obtaining:

$$\langle f|i\rangle = \int \mathcal{D}\vec{\varphi}_L \mathcal{D}\vec{\varphi}_H e^{\frac{i}{\hbar} S_{\text{full}}(\vec{\varphi}_L, \vec{\varphi}_H)} = \int \mathcal{D}\vec{\varphi}_L e^{\frac{i}{\hbar} S_{\text{eff}}(\vec{\varphi}_L)} , \tag{2.2}$$

where the dependence of the heavy field has disappeared explicitly but is encoded in the effective interactions inside the action $S_{\text{eff}}(\vec{\varphi}_L)$, which only depends on the light fields. While this procedure is conceptually easy to understand, integrations as the one shown in Eq. 2.2 are rarely feasible and the use of a more general procedure, known as the Operator Product Expansion (OPE) [15], is preferred. The idea is to identify a complete basis of local operators $\mathcal{O}_i$ compatible with the symmetries of the system and expand the effective Lagrangian as:

$$\mathcal{L}_{\text{eff}} = -\sum_i C_i(\mu)\mathcal{O}_i , \tag{2.3}$$

where the $C_i(\mu)$ couplings, called *Wilson coefficients*, in general depend on the scale μ at which the theory needs to be used. This expansion would technically run over an infinite sum of operators, however, it converges rapidly when rearranged based on the mass dimension of the local operators involved. It can be shown [16] that for weak decays of B mesons, limiting the expansion at all the operators of dimension six provides already a very good approximation.

The effective Lagrangian can then be rewritten as:

$$\mathcal{L}_{\text{eff}} = -4\frac{G_F}{\sqrt{2}} \sum_i \mathcal{C}_i^{(6)}(\mu)\mathcal{O}_i^{(6)} , \tag{2.4}$$

where the subscript referring to the dimensionality of the operator will be dropped in the rest of the text. It is important to notice that the scale μ acts as a separator of short- and long- distance effects; while the first ones are included in the Wilson coefficients, the second ones are left as explicit degrees of freedom in the EFT.

Once the set of operators for the expansion has been obtained, the set of Wilson coefficients that describes the full theory at the energy scale μ needs to be determined. This is done by performing a *matching procedure* between the effective and full theory at the electroweak scale $\mu \approx M_W$, where all amplitudes can be calculated in perturbation theory, and a set of initial conditions for the

Wilson coefficients $\mathcal{C}_i(M_W)$ can be obtained. These initial conditions are then used to determine the evolution of the $\mathcal{C}_i(\mu)$ from the electroweak scale down to the energy scale of the physical process $\mu \approx m_b$, using the renormalization-group equations (RGE) [17–19]. The last remaining step corresponds to the evaluation of the matrix elements of the effective Lagrangian for the hadronic process, that involves energy scales between the mass of the mesons and Λ_{QCD}. It is important to stress that NP contributions can only contribute to the first step of this procedure and only by either changing the initial conditions of the Wilson coefficients at $\mu \approx M_W$ or introducing new operators that describe interactions forbidden in the SM. This corresponds to either a shift on the observed Wilson coefficients with respect to the expected SM values, e.g. $\mathcal{C}_i = \mathcal{C}_i^{SM} + \mathcal{C}_i^{NP}$ or to the necessity of including another operator and Wilson coefficient to better describe the data. The remaining steps are insensitive to the physics at smaller scales by construction.

In the specific case of $b \to s\,\ell^+\ell^-$ transitions, the effective Lagrangian in the SM can be split in two terms [20]:

$$\mathcal{L}_{\text{eff}}^{s\ell+\gamma} = \frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^*(\mathcal{C}_7\mathcal{O}_{7\gamma} + \mathcal{C}_9\mathcal{O}_{9V} + \mathcal{C}_{10}\mathcal{O}_{10A}) \tag{2.5}$$

and

$$\begin{aligned}\mathcal{L}_{\text{eff}}^{had} = -\frac{4G_F}{\sqrt{2}} \{ & V_{ub}V_{us}^*[\mathcal{C}_1(\mathcal{O}_1^u - \mathcal{O}_1^c) + \mathcal{C}_2(\mathcal{O}_2^u - \mathcal{O}_2^c)] \\ & - V_{tb}V_{ts}^*[\mathcal{C}_1\mathcal{O}_1^c + \mathcal{C}_2\mathcal{O}_2^c + \sum_{i=3..6} \mathcal{C}_i\mathcal{O}_i + \mathcal{C}_8\mathcal{O}_{8g}]\} \,.\end{aligned} \tag{2.6}$$

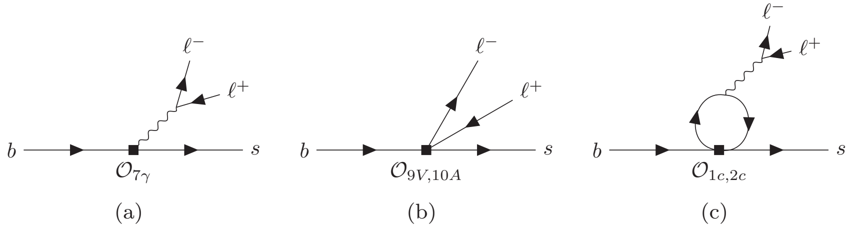
The first term in the Lagrangian, $\mathcal{L}_{\text{eff}}^{s\ell+\gamma}$, contains effective operators that contribute to the transition at tree level, while the second one, $\mathcal{L}_{\text{eff}}^{had}$, describes terms of the type $b \to sq\bar{q}$ that contribute in the decay at loop-level, as shown in Figure 2.1. This Lagrangian can be further simplified by neglecting terms proportional to $V_{ub}V_{us}^*$ that are doubly-Cabibbo suppressed and terms involving the operators $\mathcal{O}_{3..6,8g}$, whose Wilson coefficients at the b-quark scale are small ($\mathcal{O}(10^{-2})$).

The expressions of Eqs. 2.5 and 2.6 simplify to:

$$\mathcal{L}_{b\to s\,\ell^+\ell^-} = \frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^*[\mathcal{C}_1\mathcal{O}_1^c + \mathcal{C}_2\mathcal{O}_2^c] + \frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^*(\mathcal{C}_7\mathcal{O}_{7\gamma} + \mathcal{C}_9\mathcal{O}_{9V} + \mathcal{C}_{10}\mathcal{O}_{10A}) \,, \tag{2.7}$$

where the operators $\mathcal{O}_i$ are defined as:

$$\begin{aligned}\mathcal{O}_{7\gamma} &= \frac{e}{16\pi^2} m_b \bar{b}_R^\alpha \sigma^{\mu\nu} s_L^\alpha F_{\mu\nu} \,, \\ \mathcal{O}_{9V} &= \frac{1}{2}\frac{\alpha_e}{4\pi} \bar{b}_L^\alpha \gamma^\mu s_L^\alpha \bar{\ell}\gamma_\mu \ell \,, \\ \mathcal{O}_{10A} &= \frac{1}{2}\frac{\alpha_e}{4\pi} \bar{b}_L^\alpha \gamma^\mu s_L^\alpha \bar{\ell}\gamma_\mu\gamma_5 \ell \,, \\ \mathcal{O}_1^c &= \bar{b}_L^\alpha \gamma^\mu c_L^\alpha \bar{c}_L^\beta \gamma_\mu s_L^\beta \,, \\ \mathcal{O}_2^c &= \bar{b}_L^\alpha \gamma^\mu c_L^\beta \bar{c}_L^\beta \gamma_\mu s_L^\alpha \,,\end{aligned} \tag{2.8}$$



(a) (b) (c)

Figure 2.1: Feynman diagrams involving the operators $\mathcal{O}_{7\gamma}$, $\mathcal{O}_{9V,10A}$ and $\mathcal{O}_{1c,2c}$, as defined in Eq. 2.8.

and b, s, c and ℓ are the SM fermionic fields, α and β denote the colour indices and $F_{\mu\nu}$ is the electromagnetic tensor, as discussed in Ch. 1. A definition of the operators described in Eq. 2.6 and neglected in Eq. 2.7 can be found in Ref. [21]. It is important to notice that the V-A structure of weak interactions discussed in Ch. 1 remains valid in this effective theory and is enforced by the combination of the *vector* operators such as $\mathcal{O}_{7\gamma}$, $\mathcal{O}_{9V}$ and the *axial-vector* operator $\mathcal{O}_{10A}$. Even if the operators $O_{1,2}$ only enter at the loop-level, their Wilson coefficients at the b-quark scale are sizable [1] and cannot be neglected. The most prominent effects of these contributions are the narrow resonance peaks at the J/ψ and $\psi(2S)$ mass, as highlighted in grey in Figure 2.2 for $B^0 \to K^{*0}\mu^+\mu^-$ decays. It is however important to stress that their contribution is present in the entire dilepton mass region even if less noticeable, where they can interfere with the tree-level diagrams.

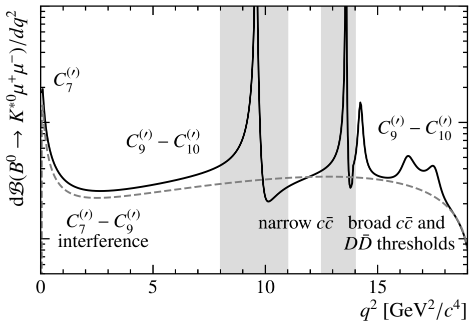


Figure 2.2: Schematic view of the $B^0 \to K^{*0}\mu^+\mu^-$ spectrum as a function of the dilepton invariant mass squared, with the dominant Wilson coefficients displayed and possible $c\bar{c}$ contributions.

---

[1] The values of the WCs for the operators $O_{1,2}$ at $\mu = 4.2\,\mathrm{GeV}/c^2$ correspond to $\mathcal{C}_1 = -0.3$ and $\mathcal{C}_2 = 1$, comparable in size to $\mathcal{C}_7$ =-0.3, $\mathcal{C}_9$ =4.3 and $\mathcal{C}_{10}$ =-4.2 [22, 23].

Additional operators can be included if specific scenarios beyond the SM (BSM) are considered. The most important ones correspond to:

- Chirality-flipped operators, $\mathcal{O}'_{7\gamma,9V,10A}$ that can be easily obtained starting from the unprimed operators defined in Eq. 2.8 by means of the simple substitution $q_{L(R)} \to q_{R(L)}$.
- Scalar and pseudoscalar operators $\mathcal{O}_{S,S'}$ and $\mathcal{O}_{P,P'}$, defined as:

$$\mathcal{O}_{S(S')} = \bar{b}^\alpha_L s^\alpha_{R(L)}\, \bar{\ell}\ell\,, \tag{2.9}$$

$$\mathcal{O}_{P(P')} = \bar{b}^\alpha_L s^\alpha_{R(L)}\, \bar{\ell}\gamma_5\ell\,, \tag{2.10}$$

- Tensor operators $\mathcal{O}_{T,T5}$, defined as:

$$\mathcal{O}_{T} = \bar{b}^\alpha_L \sigma^{\mu\nu} s^\alpha_L\, \bar{\ell}\sigma_{\mu\nu}\ell\,, \tag{2.11}$$

$$\mathcal{O}_{T5} = \bar{b}^\alpha_L \sigma^{\mu\nu} s^\alpha_L\, \bar{\ell}\sigma_{\mu\nu}\gamma_5\ell\,. \tag{2.12}$$

Of these three additional set of operators only the first one will be explicitly considered in the following, since the impact of the remaining two sets of operators is either heavily constrained [24] or expected to be negligible in most of the possible SM extensions. Similarly, the WCs are treated as real parameters, since imaginary contributions to the WCs are expected to be sizable only for CP-violating observables [25], not considered in the following.

2.2 The $\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-$ differential decay rate

Starting from the effective Lagrangian in Eq. 2.7 and including the NP contribution from chirality flipped operators, it is possible to write down the matrix element of the decay $\overline{B}^0 \to \overline{K}^{*0}(\to K^-\pi^+)\ell^+\ell^-$. Under the approximation of lepton factorization[2](LFA) the matrix element $\mathcal{M}$ becomes:

$$\begin{aligned}\mathcal{M} \;=\; & \frac{G_F\alpha_e}{\sqrt{2}\pi} V_{tb}V^*_{ts}\{[\langle K\pi|\bar{s}\gamma^\mu(\mathcal{C}_9 P_L + \mathcal{C}'_9 P_R)b|\bar{B}\rangle \\ & -\frac{2m_b M_B}{q^2}(\langle K\pi|\bar{s}i\sigma^{\mu\nu}q_\nu(\mathcal{C}_7 P_R + \mathcal{C}'_7 P_L)b|\bar{B}\rangle + \frac{16\pi^2}{M_B m_b}\mathcal{H}^\mu)](\bar{\ell}\gamma_\mu\ell) \\ & +\langle K\pi|\bar{s}\gamma^\mu(\mathcal{C}_{10} P_L + \mathcal{C}'_{10} P_R)b|\bar{B}\rangle(\bar{\ell}\gamma_\mu\gamma_5\ell)\}\,. \end{aligned} \tag{2.13}$$

The terms $\langle K\pi|\bar{s}\Gamma b|\bar{B}\rangle$ in the first two rows represent the matrix elements of the flavour changing quark currents between an initial $\overline{B}^0$ meson and the final state $\overline{K}^{*0}$. The third term, $\mathcal{H}^\mu$, describes instead the non-local hadronic effects, which appear when the lepton pair couples with the electromagnetic current, through a penguin contraction of four quark operators $\mathcal{O}_{1,2} \sim \bar{s}b\bar{c}c$.

The $\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-$ channel is studied considering the decay of $\overline{K}^{*0} \to K^-\pi^+$. The additional information provided by the angle between K and π allows having

---

[2] In this approximation the hadronic and leptonic part of the amplitude factorize.

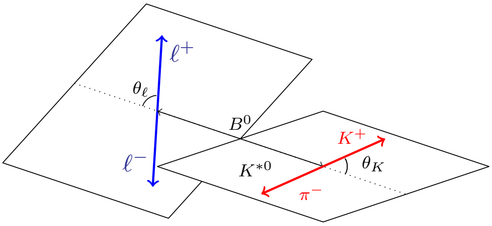
information on the polarization of the $\overline{K}^{*0}$, providing an increased number of observables that can be used to study the effective Lagrangian $\mathcal{L}_{b\to s\,\ell^+\ell^-}$. The decay is described as a function of four variables: q^2, the invariant mass squared of the dilepton system and three decay angles $\vec{\Omega} = (\theta_\ell,\, \theta_K,\, \phi)$, whose definition is sketched in Figure 2.3. Different angular conventions are typically used by experimentalists and theorists: while most theorists follow the definitions in Refs. [26, 27], this analysis uses the convention adopted in previous LHCb publications [28, 29]. The angle between the ℓ^- (ℓ^+) and the direction opposite to that of the $\overline{B}^0$ (B^0) in the rest frame of the dilepton system is denoted θ_ℓ. The angle between the direction of the K^- (K^+) and the direction opposite to that of the $\overline{B}^0$ (B^0) in the rest frame of the $\overline{K}^{*0}$ (K^{*0}) system is denoted θ_K. The angle between the plane defined by the dilepton pair and the plane defined by the kaon and pion in the $\overline{B}^0$ (B^0) rest frame is denoted ϕ.

The differential decay rate can be written, after squaring the matrix element $\mathcal{M}$, summing over the external helicities and averaging over the final state spins, as:

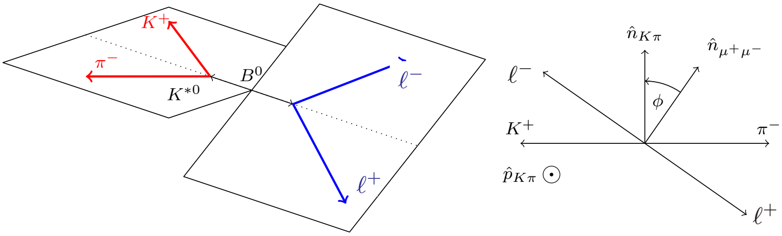
$$\begin{aligned}
\frac{\mathrm{d}^4\Gamma[\bar{B}^0 \to \bar{K}^{*0}\ell^+\ell^-]}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}} &= \frac{9}{32\pi}\sum_i I_i(q^2) f_i(\vec{\Omega}) \\
&= \frac{9}{32\pi}\Big[\, I_{1s}\sin^2\theta_K + I_{1c}\cos^2\theta_K \\
&\qquad + I_{2s}\sin^2\theta_K\cos 2\theta_\ell + I_{2c}\cos^2\theta_K\cos 2\theta_\ell \\
&\qquad + I_3\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi + I_4\sin 2\theta_K\sin 2\theta_\ell\cos\phi \\
&\qquad + I_5\sin 2\theta_K\sin\theta_\ell\cos\phi + I_6\sin^2\theta_K\cos\theta_\ell \\
&\qquad + I_7\sin 2\theta_K\sin\theta_\ell\sin\phi + I_8\sin 2\theta_K\sin 2\theta_\ell\sin\phi \\
&\qquad + I_9\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\,\Big],
\end{aligned} \tag{2.14}$$

where I_i are angular coefficients that implicitly depend on q^2. These coefficients can be conveniently expressed in terms of transversity amplitudes, $\mathcal{A}^{L,R}_{0,\perp,\|}$ and $\mathcal{A}_t$, where $\lambda = 0, \perp, \|, t$ refer to the polarization and L and R to the chirality of the lepton current [30]. The angular coefficients are defined as

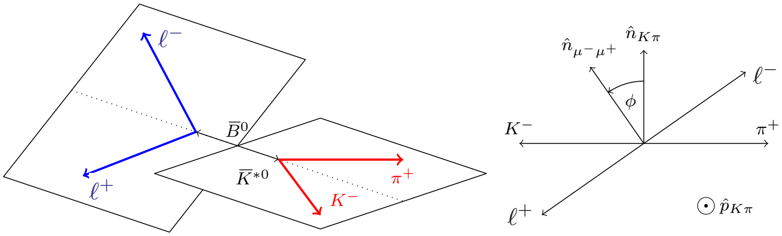
$$\begin{aligned}
I_{1s} &= \frac{2+\beta_l^2}{4}\Big[|\mathcal{A}_\perp^L|^2 + |\mathcal{A}_\|^L|^2 + (L\to R)\Big] \\
&\quad + \frac{4m_l^2}{q^2}\,\mathcal{R}e\Big(\mathcal{A}_\perp^L\mathcal{A}_\perp^{R*} + \mathcal{A}_\|^L\mathcal{A}_\|^{R*}\Big), \\
I_{1c} &= \Big[|\mathcal{A}_0^L|^2 + |\mathcal{A}_0^R|^2\Big] + \frac{4m_l^2}{q^2}\Big[|\mathcal{A}_t|^2 + 2\,\mathcal{R}e(\mathcal{A}_0^L\mathcal{A}_0^{R*})\Big], \\
I_{2s} &= \frac{\beta_l^2}{4}\Big[|\mathcal{A}_\perp^L|^2 + |\mathcal{A}_\|^L|^2 + (L\to R)\Big], \\
I_{2c} &= -\beta_l^2\Big[|\mathcal{A}_0^L|^2 + |\mathcal{A}_0^R|^2\Big],
\end{aligned}$$



(a) θ_K and θ_ℓ definitions for the B^0 decay



(b) ϕ definition for the B^0 decay



(c) ϕ definition for the $\overline{B}^0$ decay

Figure 2.3: Graphical representation of the angular basis used for $B^0 \to K^{*0}\ell^+\ell^-$ and $\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-$ decays in this work. The notation $\hat{n}_{ab}$ is used to represent the normal to the plane containing particles a and b in the B^0 (or $\overline{B}^0$) rest frame. An explicit description of the angular basis is given in the text.

$$\begin{aligned}
I_3 &= \frac{\beta_l^2}{2}\Big[|\mathcal{A}_\perp^L|^2 - |\mathcal{A}_\parallel^L|^2 + (L \to R)\Big], \qquad (2.15)\\
I_4 &= -1 \cdot \frac{\beta_l^2}{\sqrt{2}}\,\mathcal{R}e\Big[\mathcal{A}_0^L \mathcal{A}_\parallel^{L\,*} + (L \to R)\Big],\\
I_5 &= \sqrt{2}\beta_l\,\mathcal{R}e\Big[\mathcal{A}_0^L \mathcal{A}_\perp^{L\,*} - (L \to R)\Big],
\end{aligned}$$

$$
\begin{aligned}
I_{6s} &= -1 \cdot 2\beta_l\, \mathcal{R}e\Big[\mathcal{A}_\parallel^L \mathcal{A}_\perp^{L^*} - (L \to R)\Big], \\
I_7 &= -1 \cdot \sqrt{2}\beta_l\, \mathcal{I}m\Big[\mathcal{A}_0^L \mathcal{A}_\parallel^{L^*} - (L \to R)\Big], \\
I_8 &= \frac{\beta_l^2}{\sqrt{2}}\, \mathcal{I}m\Big[\mathcal{A}_0^L \mathcal{A}_\perp^{L^*} + (L \to R)\Big], \\
I_9 &= -1 \cdot \beta_l^2\, \mathcal{I}m\Big[\mathcal{A}_\perp^L \mathcal{A}_\parallel^{L^*} + (L \to R)\Big],
\end{aligned}
$$

where $\beta_l = \sqrt{1 - 4m_l^2/q^2}$. Notice that the factor -1 in front of the terms $I_{4,6s,7,9}$ in Eq. 2.15 is needed to align the adoption of the LHCb angular notation with the definition of the transversity amplitudes, as obtained in the theory convention [31].

The transversity amplitudes are the physical quantities that describe the decay process and can be expressed as [23, 32]

$$
\begin{aligned}
\mathcal{A}_\perp^{L,R} &= N\Big\{\Big[(\mathcal{C}_9 + \mathcal{C}_9') \mp (\mathcal{C}_{10} + \mathcal{C}_{10}')\Big]\mathcal{F}_\perp \\
&\qquad + \frac{2m_b M_B}{q^2}\Big[(\mathcal{C}_7 + \mathcal{C}_7')\mathcal{F}_\perp^T - 16\pi^2 \frac{M_B}{m_b}\mathcal{H}_\perp\Big]\Big\}, \\
\mathcal{A}_\parallel^{L,R} &= -N\Big\{\Big[(\mathcal{C}_9 - \mathcal{C}_9') \mp (\mathcal{C}_{10} - \mathcal{C}_{10}')\Big]\mathcal{F}_\parallel \\
&\qquad + \frac{2m_b M_B}{q^2}\Big[(\mathcal{C}_7 - \mathcal{C}_7')\mathcal{F}_\parallel^T - 16\pi^2 \frac{M_B}{m_b}\mathcal{H}_\parallel\Big]\Big\}, \\
\mathcal{A}_0^{L,R} &= -N\Big\{\Big[(\mathcal{C}_9 - \mathcal{C}_9') \mp (\mathcal{C}_{10} - \mathcal{C}_{10}')\Big]\mathcal{F}_0 \\
&\qquad + \frac{2m_b M_B}{q^2}\Big[(\mathcal{C}_7 - \mathcal{C}_7')\mathcal{F}_0^T - 16\pi^2 \frac{M_B}{m_b}\mathcal{H}_0\Big]\Big\}, \\
\mathcal{A}_t &= -2N\,(\mathcal{C}_{10} - \mathcal{C}_{10}')\mathcal{F}_t\,,
\end{aligned} \tag{2.16}
$$

where N is a normalization factor given by

$$
N = G_F \alpha_e V_{tb} V_{ts}^* \sqrt{\frac{q^2 \beta_l \sqrt{\lambda}}{3 \cdot 2^{10}\pi^5 M_B}}\,, \tag{2.17}
$$

and λ is the kinematic Källén function defined as

$$
\lambda(M_B^2, M_{K^{*0}}^2, q^2) = M_B^4 + M_{K^{*0}}^4 + (q^2)^2 - 2M_B^2(M_{K^{*0}}^2 + q^2) - 2q^2 M_{K^{*0}}^2\,. \tag{2.18}
$$

Additionally, the expressions depend on the seven *form factors* $\mathcal{F}_i^{(T)}$, defined with respect to $\langle K\pi|\bar{s}\Gamma b|\bar{B}\rangle$ by the following equations [23]:

$$
\begin{aligned}
\langle K\pi(k,\eta)|\bar{s}\gamma^\mu P_{L(R)} b|\bar{B}(q+k)\rangle &= \eta_\alpha^*(\mathcal{F}_\perp S_\perp^{\alpha\mu} \mp (\mathcal{F}_\parallel S_\parallel^{\alpha\mu} + \mathcal{F}_0 S_0^{\alpha\mu} + \mathcal{F}_t S_t^{\alpha\mu}))\,, \\
\langle K\pi(k,\eta)|\bar{s} i\sigma^{\mu\nu} q_\nu P_{R(L)} b|\bar{B}(q+k)\rangle &= iM_B\, \eta_\alpha^*(\mathcal{F}_\perp^T S_\perp^{\alpha\mu} \mp (\mathcal{F}_\parallel^T S_\parallel^{\alpha\mu} + \mathcal{F}_0^T S_t^{\alpha\mu}))\,,
\end{aligned} \tag{2.19}
$$

and on the *charm-loop* contribution $\mathcal{H}^\mu$ [23], whose definition is:

$$\begin{aligned}\eta^*_\alpha \mathcal{H}^{\alpha\mu} &= i\int d^4x e^{iq\cdot x}\langle K\pi(k,\eta)|T\{j^\mu_{em}(x), \mathcal{C}_1\mathcal{O}_1(y) + \mathcal{C}_2\mathcal{O}_2(y)\}|\bar{B}(q+k)\rangle \\ &= \eta^*_\alpha(M^2_B[\mathcal{H}_\perp S^{\alpha\mu}_\perp - \mathcal{H}_\| S^{\alpha\mu}_\| - \mathcal{H}_0 S^{\alpha\mu}_0])\,. \end{aligned} \tag{2.20}$$

The Lorentz structures $S^{\alpha\mu}_\lambda$ can be defined as a function of the momentum of the kaon, k, and of the momentum transferred to the dilepton system, q, and are discussed more in detail in Ref. [23]. A parametrization of the form factors and charm-loop terms as a function of q^2 is difficult to obtain from first principles, since the dynamic that governs the transition from a $\overline{B}^0$ to a $\overline{K}^{*0}$ meson is dominated by QCD in a non-perturbative regime. Multiple attempts to estimate these contributions have been pursued in the last decades, greatly improving the theoretical understanding of these terms (see Appendices A and B for the parametrizations considered in this work), however, as it is discussed in Sec. 2.3, they still correspond to the main sources of uncertainty in the SM prediction of the measured observables in $\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-$ decays.

An expression formally identical to Eq. 2.14 can be obtained for the decay $B^0 \to K^{*0}\ell^+\ell^-$ by substituting the angular coefficients I_i with $\bar{I}_i$. This is possible thanks to the choice of the angular basis described in Figure 2.3, which is able to counterbalance the change of signs that occur on the angular coefficients $\bar{I}_{7,8,9}$ after a CP-transformation. In this way, the joint differential decay rate such as $\mathrm{d}^4(\Gamma \pm \bar{\Gamma})/\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}$ will only contain either CP-symmetric or CP-asymmetric quantities, as discussed in Sec. 2.3.2.

2.3 Conventional observables and experimental results in $B^0 \to K^{*0}\ell^+\ell^-$

The differential decay rate expressed in Eq. 2.14 clearly shows that the angular structure of the decay is solely determined by conservation of linear and angular momentum; the only dependence on the dynamics of the decay is encoded in the angular coefficients $I_i(q^2)$. This provides an ideal venue to test the Standard Model in a model independent way, by comparing its predictions with the measured values of the branching ratio or of the angular coefficients as a function of q^2. In the following, the most recent set of measurements is presented, together with the definition of the corresponding observables.

2.3.1 Branching ratio

A branching ratio simply represents the probability that, of all the channels a $\overline{B}^0$ is allowed to decay to, it decays through a specific channel such as $\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-$. Formally, it can be defined as:

$$\mathcal{B}(\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-) = \frac{\tau_B}{\hbar}\int \mathrm{d}\vec{\Omega}\mathrm{d}q^2 \frac{\mathrm{d}^4\Gamma[\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-]}{\mathrm{d}\vec{\Omega}\mathrm{d}q^2}\,, \tag{2.21}$$

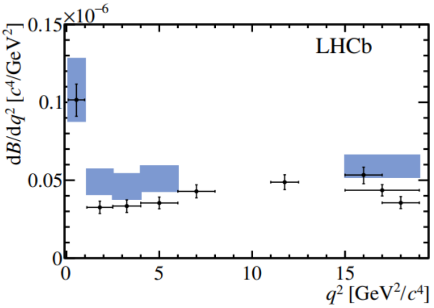


Figure 2.4: Differential branching ratio of $B^0 \to K^{*0}\ell^+\ell^-$ as a function of q^2. In black the data overlaid with the SM prediction from Refs. [34, 35]. The uncertainties shown here are the quadratic sum of the statistical and systematic uncertainties including also the uncertainty on the normalization mode $B^0 \to K^{*0} J/\psi (\to \mu^+\mu^-)$.

where the parameters τ_B and $\hbar$ correspond to the lifetime of the $\overline{B}^0$ meson and the reduced Plank constant, respectively. Alternatively, it is possible to study the differential branching ratio as a function of q^2, computed in the bin Δq^2, by using the expression

$$\mathrm{d}\mathcal{B}(\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-)/\mathrm{d}q^2|_{\Delta q^2} = \frac{\tau_B}{\Delta q^2 \hbar} \int_{\Delta q^2} \mathrm{d}\vec{\Omega}\mathrm{d}q^2 \frac{\mathrm{d}^4\Gamma[\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-]}{\mathrm{d}\vec{\Omega}\mathrm{d}q^2} \,. \tag{2.22}$$

LHCb measured both these quantities [33] averaging between $\overline{B}^0$ and B^0 contributions, obtaining their most precise measurements up-to-date. For the total branching ratio, the measurement corresponds to

$$\mathcal{B}(B^0 \to K^{*0}\ell^+\ell^-) = (0.904^{+0.016}_{-0.015} \pm 0.010 \pm 0.006 \pm 0.061) \cdot 10^{-6} \,, \tag{2.23}$$

where the uncertainties are, from left to right, statistical, systematic, due to the extrapolation to the full q^2 region and due to the branching ratio of the normalization mode. The differential decay rate is shown instead in Figure 2.4. As it can be seen, the observed $\mathrm{d}\mathcal{B}/\mathrm{d}q^2$ lays systematically below the expected SM predictions. The significance of this deviation is currently limited by the theoretical prediction on this measurement, whose main contribution is linked to the current knowledge on the hadronic matrix elements.

2.3.2 Angular observables

Once enough statistics is available, it is possible to exploit the rich angular structure of $\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-$ decays to estimate the eleven angular coefficients I_i as a function of q^2. Conventionally, it has always been preferred to use a set of derived angular observables defined as

$$S_i = (I_i + \bar{I}_i)/(\frac{\mathrm{d}(\Gamma + \bar{\Gamma})}{\mathrm{d}q^2}) \quad \text{and} \quad A_i = (I_i - \bar{I}_i)/(\frac{\mathrm{d}(\Gamma + \bar{\Gamma})}{\mathrm{d}q^2}) \,, \tag{2.24}$$

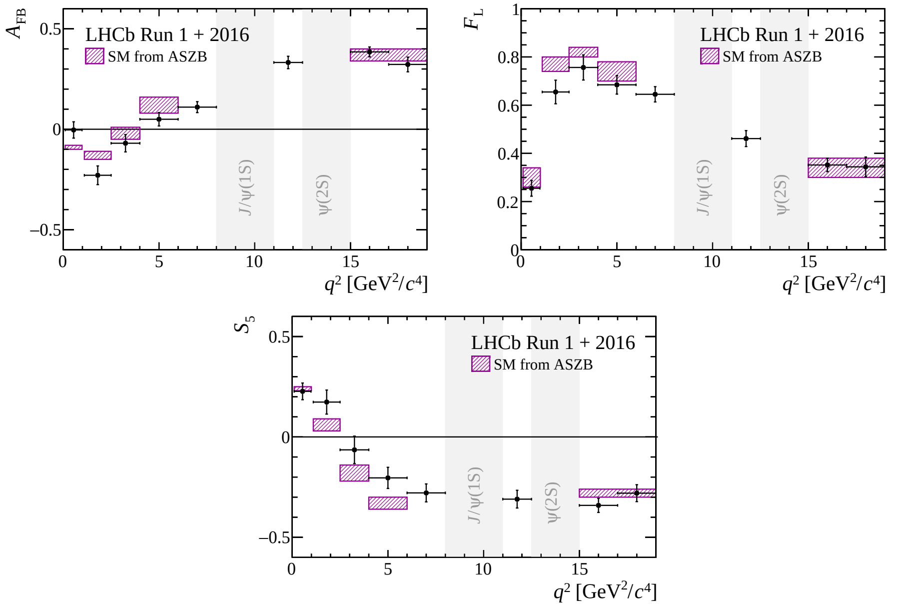


Figure 2.5: Result of the CP-averaged angular observables F_L, A_{FB} and S_5 in bins of q^2 as obtained by the LHCb experiment [29]. The experimental data, in black, is overlaid with the SM predictions based on the prescriptions of Refs.[34, 36]

corresponding to the sum (CP-symmetry) and difference (CP-asymmetry) between the angular coefficients for the decay of $\overline{B}^0$ and B^0, normalized to the sum of the decay rates. Additionally, it is possible to define two other angular observables:

$$F_L = S_{1c} \quad \text{and} \quad A_{FB} = \frac{3}{4} S_{6c}\,, \tag{2.25}$$

corresponding to[3] the fraction of longitudinal polarization of the $\overline{K}^{*0}$ and to the forward-backward asymmetry of $\ell^-(\ell^+)$ with respect to the $\overline{B}^0$ (B^0) in the dilepton rest frame.

The full set of CP-averaged and CP-asymmetric observables S_i and A_i has been measured by LHCb [28, 29], and a discrepancy with respect to the SM predictions has been seen also in this case. A direct interpretation of such discrepancies was however hindered by the sizable uncertainty in theoretical predictions, dominated by the form factor parametrization. Figure 2.5 shows some of the most interesting observables as measured by LHCb [29].

Starting from the S-basis, it is possible to define an alternative set of optimized observables, for which the form factor uncertainty cancels at leading order. The

---

[3] This definition is only valid under the approximation of massless leptons, in which $S_{1c} = -S_{2c}$. The more general definition corresponds to $F_L = (3/4 I_{1c} - 1/4 I_{2c})$.

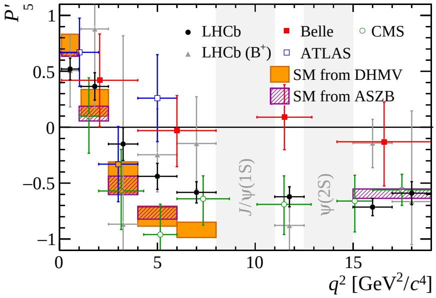


Figure 2.6: A summary of the available measurements of the P'_5 observables in the $B \to K^*\mu^+\mu^-$ channel as a function of q^2, including the contributions from LHCb [29, 37], Belle [38], ATLAS [39] and CMS [40]. The experimental data, is overlaid with the SM predictions based on the prescriptions of Refs. [34, 36, 41, 42].

observable basis, also referred to as P-basis [43, 44], corresponds to:

$$\begin{aligned} P_1 &= \frac{2S_3}{(1-F_L)} = A_T^{(2)} \, , \\ P_2 &= \frac{2}{3}\frac{A_{FB}}{(1-F_L)} \, , \\ P_3 &= \frac{S_9}{(1-F_L)} \, , \\ P'_{4,5,8} &= \frac{S_{4,5,8}}{\sqrt{F_L(1-F_L)}} \, , \\ P'_6 &= \frac{S_7}{\sqrt{F_L(1-F_L)}} \, . \end{aligned} \tag{2.26}$$

Figure 2.6 shows a summary of the measured values of P'_5 in the $B^0 \to K^{*0}\ell^+\ell^-$ and $B^+ \to K^{*+}\mu^+\mu^-$ channels as a function of q^2. Analogous plots can be found in the corresponding references for some or all of the remaining P observables. A coherent discrepancy between the SM predictions and the observed values in LHCb can be noticed for the q^2 bins $\Delta q^2 = \{[4,6],[6,8]\}\,\mathrm{GeV}^2/c^4$, corresponding to a local tension of 2.5σ and 2.9σ respectively. A direct interpretation of these results as NP contributions is however non-trivial, since the vector-like nature of these deviations might also be explained by large hadronic contributions from $b \to s\bar{c}c$ operators, that could mimic a NP contribution if not properly estimated [41, 42, 45–51].

2.3.3 Lepton flavour universality tests

The two sets of observables described in Sec. 2.3.1 and 2.3.2 have clearly pointed out the existence of a series of discrepancies between data and SM predictions. In

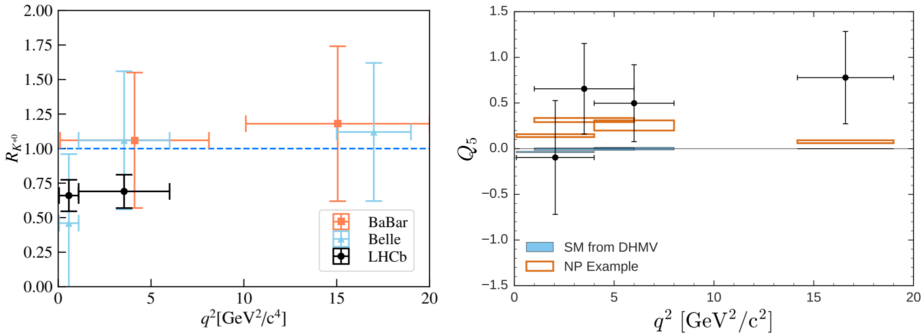


Figure 2.7: On the left, a comparison of the LHCb R_{K^*} measurements [52] with experimental results obtained from the B factories [53, 54]. On the right, a comparison of the measurement of the Q_5 observable by Belle [38] with the SM (blue) and NP scenario (orange) corresponding to a shift in the Wilson coefficients for $B^0 \to K^{*0}\mu^+\mu^-$ of $\mathcal{C}_9^{NP}$=-1.1 [56].

most situations, the significance of the discrepancy is limited by the uncertainty of the theoretical prediction, specifically by our knowledge of local and non-local hadronic matrix elements, which are difficult to compute from first principles. A possible way to overcome these limitations is to consider ratios or differences of observables between $B^0 \to K^{*0}\ell^+\ell^-$ decays, where the lepton ℓ corresponds to either a muon or an electron. Thanks to the fact that hadronic matrix elements are described by QCD, a theory that cannot distinguish between different lepton generations, it is possible to assume that the differential decay rate of Eq. 2.14 has to be identical for all lepton generations in the SM, up to corrections due their different masses. Any unaccounted difference between lepton generations would then be a clear sign of NP. Figure 2.7 shows the experimental results obtained by LHCb [52], Babar [53] and Belle [54] for the observables $R_{K^{*0}}$(left) and by Belle for the angular observable Q_5 [38](right). The first one is defined as the ratio of the branching ratios in the bin Δq^2:

$$R_{K^{*0}}(\Delta q^2) = \frac{\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)|_{\Delta q^2}}{\mathcal{B}(B^0 \to K^{*0}e^+e^-)|_{\Delta q^2}}\,, \tag{2.27}$$

and, in regions of the phase space where the mass difference between electrons and muons can be neglected, is expected to be unity within an uncertainty of $\mathcal{O}(10^{-2})$ [55]. The second one is defined as the difference between the observable P_5' between muons and electrons, $Q_5(\Delta q^2) = P_5^{\mu(\prime)} - P_5^{e(\prime)}$, and is expected to be zero in the SM [56]. All measurements are compatible with their SM predictions, except for the value observed for R_{K^*} by LHCb, where a tension of 2.1 and 2.5σ is observed in the two q^2 bins of $\Delta q^2 \in \{[0.045, 1.1], [1.1, 6]\}$ GeV$^2/c^4$, respectively.

2.4 Hunting for NP and compatibility with SM

The set of deviations presented so far is not peculiar of $B^0 \to K^{*0}\ell^+\ell^-$ decays, but it seems to repeat itself in a coherent way also in other $b \to s\,\ell^+\ell^-$ transitions. Examples of it are the deficits observed for the branching ratio measurements of $B^+ \to K^+\mu^+\mu^-$ [57], $B_s \to \phi\mu^+\mu^-$ [58], the deviations observed in the angular analysis of $B_s \to \phi\mu^+\mu^-$[59] and $B^+ \to K^{*+}\mu^+\mu^-$[37] and in the ratios of branching ratios R_K[60], $R_{K^0_S}$[61], $R_{K^{*+}}$[61] and R_{Kp}[62] corresponding to the decay modes $B^+ \to K^+\ell^+\ell^-$, $B^0 \to K^0_S\ell^+\ell^-$, $B^+ \to K^{*+}\ell^+\ell^-$ and $\Lambda_b \to pK\ell^+\ell^-$, respectively. All together these deviations are referred to as *$b \to s\,\ell^+\ell^-$ anomalies*.[4]

Several attempts have been made in the last years to interpret the impact of these experimental results in a model independent way, using the well established EFT approach described in Sec. 2.1. This has usually been accomplished by selecting a subset of the measurements described above and, after a parametrization of the observables of interest as a function of the Wilson coefficients and some remaining nuisance parameters, performing a *global fit* by minimizing the distance with respect to the published results. Most of these fits are obtained varying a small number of WCs, typically one or two. Figure 2.8(a) shows a comparison between some recent global fit results [69–72] when including most of the observables mentioned before, while Figure 2.8(b) shows the same comparison when considering only the LFU observables and the $B_s \to \mu^+\mu^-$ branching ratio. The parameters of interest of the fit correspond to NP contributions in $\mathcal{C}_9$ and $\mathcal{C}_{10}$ exclusively in muons. Three main points are of particular interest:

- The best fit results are scattered in a diagonal that goes roughly from (-1,0) to (0.5, 0.5) depending on the hypothesis used to perform the global fit.
- All approaches point to a best fit scenario whose compatibility with the SM hypothesis (0,0) is above 5σ.
- The preferred scenarios are mainly of two kinds: a NP contribution in $\mathcal{C}_9^{NP\mu}$ only or a combined shift in $\mathcal{C}_9^{NP\mu}$ and $\mathcal{C}_{10}^{NP\mu}$, such that $\mathcal{C}_9^{NP\mu} = -\mathcal{C}_{10}^{NP\mu}$. This last scenario is particularly appealing since it would preserve the V-A structure of the leptonic current in the SM.
- There is a complete agreement between the different approaches when considering only LFU observables and the branching ratio of $B_s \to \mu^+\mu^-$. This is a consequence of the effective cancellation of the hadronic uncertainties in these observables. The scatter shown in Figure 2.8(a)

---

[4] Another very promising set of anomalies has been observed in LFU tests of charged current transitions of the type $b \to c\ell^-\bar{\nu}$, where ℓ is either a tau or a muon. While some attempts to explain $b \to c\ell^-\bar{\nu}$ and $b \to s\,\ell^+\ell^-$ anomalies in a common framework have been investigated in the literature [63–67], the following work will be focused exclusively on the anomalies observed in $b \to s\,\ell^+\ell^-$ transitions. More information can be found in Ref. [68] and references therein.

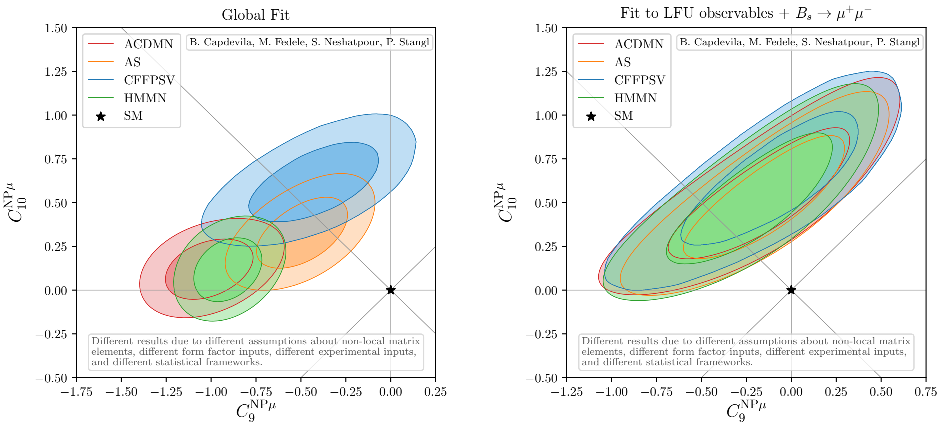


Figure 2.8: (*a*) Global fit results obtained floating simultaneously the NP contributions $\mathcal{C}_9^{NP}$ and $\mathcal{C}_{10}^{NP}$ in muons using the four different approaches described in Refs. [69–72]. (*b*) A comparison between the four approaches when only LFU observables and the branching ratio of $B_s \to \mu^+\mu^-$ are considered in the fit.

is dominated by the choice of parametrization of the hadronic matrix elements and of the measurements considered.

While these results are usually quite interesting and valuable, since they can be used by the theoretical community to build NP models or prioritize the experimental work, in general they should not be used to estimate the global compatibility of these measurements with the SM hypothesis. The main reason is that, by limiting the number of WC used in the fit to a subset (one or two) of those that enter the effective Hamiltonian, a bias might be introduced in the estimation of the significance of the NP hypothesis. A more robust and conservative estimation of the significance has been computed taking into account this issue, equivalent to considering the *look-elsewhere effect* (LEE) in searches of new resonances, showing the global significance with respect to the SM hypothesis is of 4.3σ[25].

Chapter 3

The LHCb detector at LHC

3.1 The large hadron collider at CERN

The Large Hadron Collider (LHC) [73] is the largest and most powerful particle accelerator in the world, situated at the European Organization for Nuclear Research (CERN) on the Franco-Swiss border just outside Geneva. The accelerator is installed in a 26.7 km tunnel, roughly 100 m underground, that originally hosted the Large Electron-Positron collider (LEP). Its main goal is to study proton-proton collisions at high energies to precisely test the SM of particle physics. Additionally, it can also be operated to collide heavy ions such as lead, to study strongly interacting matter in conditions similar to the first moments after the Big Bang. The LHC operation time consisted of two main runs, referred to as RUN 1 and RUN 2 in the following. RUN 1 started in 2010 and ended in 2012, and in the first two years was characterized by a centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV/c^2}$ that was then ramped up to $\sqrt{s} = 8\,\mathrm{TeV/c^2}$ for 2012. RUN 2 started in 2015, after the end of the first long shutdown (LS1), where the machine was prepared to run at the energy of $\sqrt{s} = 13\,\mathrm{TeV/c^2}$. The accelerator ran in this configuration until 2018, when it was shut down to prepare the machine for the RUN 3.

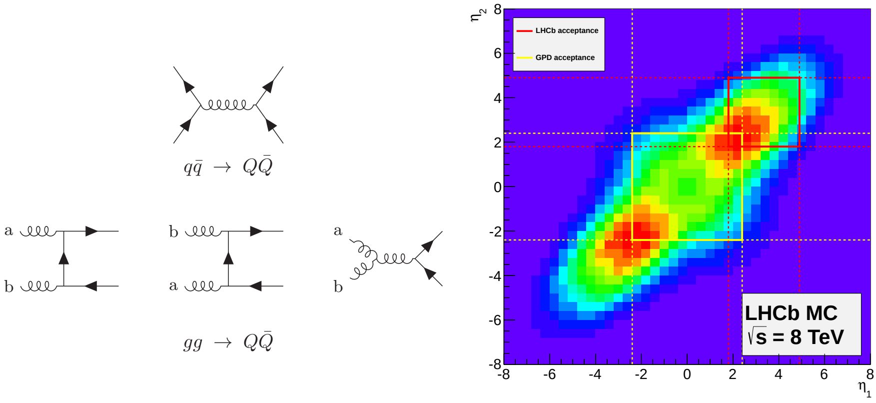
In LHC, protons are extracted from hydrogen gas by means of ionization and are accelerated in four steps using the LINAC, booster, Proton Synchrotron (PS) and Super Proton Synchrotron (SPS), respectively, until they are ready to be injected in the LHC at the energy of $450\,\mathrm{GeV}/c^2$. Protons are then split in two separate beams travelling in opposite directions that are accelerated up to $7\,\mathrm{TeV/c^2}$, reaching almost the speed of light. Each proton beam is composed of 2808 bunches, roughly 10 cm in length, with a spacing of 7 m with respect to each other. Each bunch contains roughly $1.2 \cdot 10^{11}$ protons, completing a full revolution roughly 11000 times per second. The beams are bent in quasi-circular trajectories by superconducting dipole magnets, producing magnetic fields up to 8.3 T, and further focused by quadrupole or higher-order magnets that, before a collision, squeeze the beams down to $\sim 1\,\mu\mathrm{m}$ in the transverse direction. The two proton beams are then collided in four interaction points, each of which hosts one of the large LHC experiments: ALICE [74], ATLAS [75], CMS [76] and LHCb [77]. ATLAS and CMS operate at the peak luminosity of $L = 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and are general-purpose detectors (GDP): their aim is to study collisions that produce high transverse momentum particles to search for on-shell production of NP particles and precise measurements of the properties of the Higgs boson, that they observed in 2012 for the first time [1, 2]. LHCb and ALICE are instead two experiments with a more focused physics program: the first one is optimized for the study of B-decays and will be discussed in more detail in the rest of the chapter, while ALICE is a dedicated ion experiment where lead ions are collided to study the properties of quark-gluon plasma.

3.2 The LHCb detector

LHCb is a dedicated heavy flavour physics experiment, whose main goal is the search for indirect signatures of NP in CP violating and rare decays of beauty and charm hadrons. This is possible thanks to the large production cross-section of b- and c-hadrons [78–80] available at LHC, that allowed collecting a total of $9\,\mathrm{fb}^{-1}$ of integrated luminosity[1] between Run 1 and Run 2, corresponding to roughly 10^{12} $b\bar{b}$ and 10^{13} $c\bar{c}$ pairs. At the energies of the LHC, the heavy hadrons are produced through two main processes [81], $q\bar{q} \rightarrow Q\bar{Q}$ and $gg \rightarrow Q\bar{Q}$, shown diagrammatically in Figure 3.1(a). As a result, $b\bar{b}$ and $c\bar{c}$ pairs are predominantly produced in the forward or backward direction with respect to the proton beam, as shown by Figure 3.1(b) for $b\bar{b}$. This has important consequences for the design of a detector: to efficiently collect large numbers of $b\bar{b}$ pairs it is sufficient to instrument one of these two regions. This efficiency principle has driven the design of the LHCb detector to its current set-up as a single-arm forward spectrometer covering the pseudorapidity region $1.8 < \eta < 4.9$; by covering only 4% of the solid angle, is able to collect 25% of all $b\bar{b}$ pairs produced. For comparison, GDPs experiments covering the pseudorapidity region $\eta < |2.4|$, corresponding to 90% of the solid angle, collect less than twice the fraction of $b\bar{b}$ pairs in the LHCb acceptance. Proton-proton collisions allow the production of all b hadron species; allowing the study of the decays of B_s, Λ_b, etc., not accessible at electron-positron accelerators such as BaBar [83] and Belle [84]. This advantage comes however with a price: the heavy hadrons of interest are produced together with hundreds of particles emerging from the inelastic scattering of the two protons. In such a polluted environment, finding the decay of a b- or c- hadron can become very challenging. To overcome this issue, LHCb was built with and excellent vertex resolution, that enables it to distinguish between primary vertices (PV), where the p-p collisions happened and the heavy hadrons were produced, and secondary vertices (SV), where the heavy hadrons decayed after flying for few millimetres in the detector. The backgrounds that survive can then be further reduced employing a good momentum and invariant mass resolution and an excellent charged particle identification, which allows to identify and eventually suppress exclusive background. These excellent performances are achieved also thanks to a reduction of the delivered instantaneous luminosity of two order of magnitudes with respect to ATLAS and CMS experiments, corresponding to roughly $2 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The luminosity is kept constant by an appropriate offset of the proton beams, allowing to have events dominated by single proton-proton interactions and reducing the occupancy and radiation damage to the detector. This configuration has the additional advantage of reducing systematic

---

[1] The integrated luminosity $\mathcal{L}$ is a measurement of the size of the collected dataset and corresponds to the time integral of the instantaneous luminosity L, *i.e.* $\mathcal{L} = \int L dt$, where L is the number of collisions that are produced in LHCb per cm^2 per second.



(a) (b)

Figure 3.1: (a) Leading-order diagrams for heavy-quark pair production [81]. (b) Distribution of the pseudorapidity of a b and $\bar{b}$ quark produced at the center-of-mass energy of 8 TeV [82]. In red and yellow are highlighted the acceptances covered by LHCb and a GPD respectively.

uncertainties, since it allows keeping the trigger configurations of the detector more constant over time. For these reasons LHCb is now the leading beauty physics experiment.

Figure 3.2 shows a sketch of the detector with its main components. A right-handed coordinate system is adopted with the z-axis along the beam pipe and the y-axis along the vertical axis of the page. Its design, conceptually very similar to the spectrometers of ATLAS and CMS, is dictated by the different ways in which electrons, photons, hadrons and muons interact with ordinary matter. The spectrometer is characterized by a forward angular coverage of approximately [10,300] mrad and [10,250] mrad in the bending (x-z) and non-bending (y-z) plane, respectively. Its subdetectors can be grouped in two main categories: vertex and track reconstruction systems and particle identification systems, discussed in Secs. 3.2.1 and 3.2.2, respectively.

3.2.1 Tracking and vertex reconstruction systems

The tracking detectors are built to record the passage of charged particles originating from p-p collisions before and after they are bent by a magnet. The knowledge of the trajectory of these particles allows reconstructing the vertex where they originated and to estimate their momenta. Tracking detectors are therefore built with a great attention to material budget, reducing as much as possible the interactions with charged and neutral particles to avoid biasing their estimation of momenta or stopping them too early in the detector. In

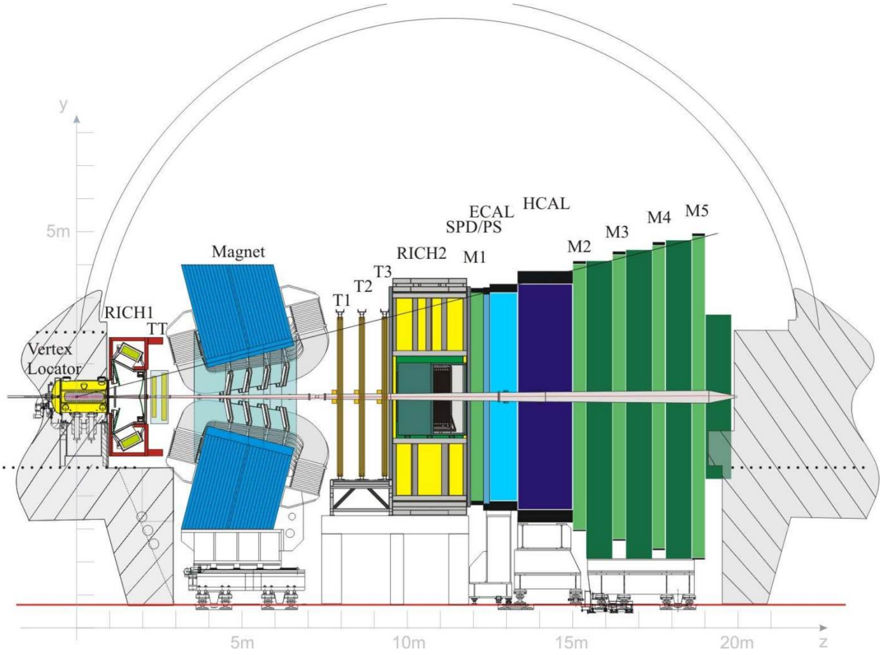
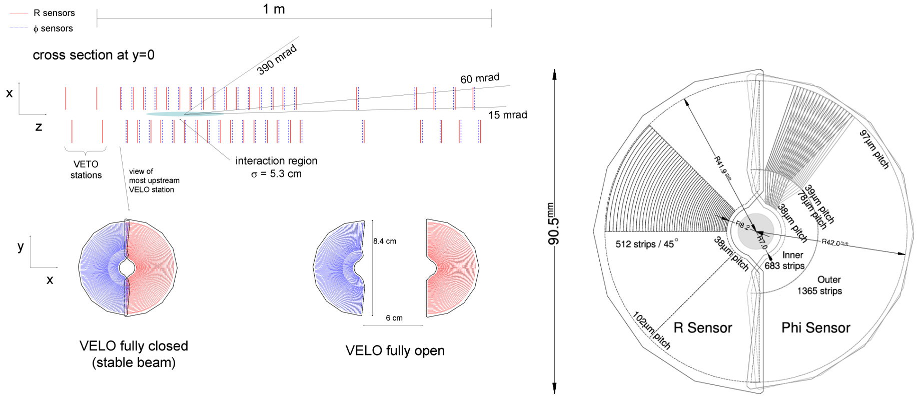


Figure 3.2: Later view of the LHCb detector [85]. For clarity, it is superimposed the right-handed coordinate system described in the text, where the z-axis runs horizontally from left to right along the beam pipe, the y-axis runs vertically from the bottom to the top of the figure and the x-axis is perpendicular to the picture, entering the page.

LHCb, starting from the interaction point, the tracking system is composed of the VErtex Locator (VELO), Tracker Turicensis (TT), dipole magnet and three tracking stations T1, T2 and T3 composed each of an Inner Tracker (IT) and Outer Tracker (OT). In the following is presented a brief discussion of the most important features for each of these elements.

3.2.1.1 Dipole magnet

In LHCb, a warm dipole magnet [86] is used to bend charged particles along the x-z plane and measure their momentum. It is characterized by saddle-shaped coils in a window-frame yoke with sloping poles that match the acceptance of the spectrometer, with an integrated bending power of 4 Tm for tracks of 10 m. The magnet has been designed to combine, on one side, magnetic fields inside the Hybrid Photon Detectors (HPD) of the RICH below 2 mT and, on the other side, a field as high as possible in the region between the VELO and TT. The first requirement ensures the proper operation of the HPDs, very sensible to external magnetic fields, while the second one allows for a rough estimation of the momentum of those particles that are swept outside the spectrometer before



(a) (b)

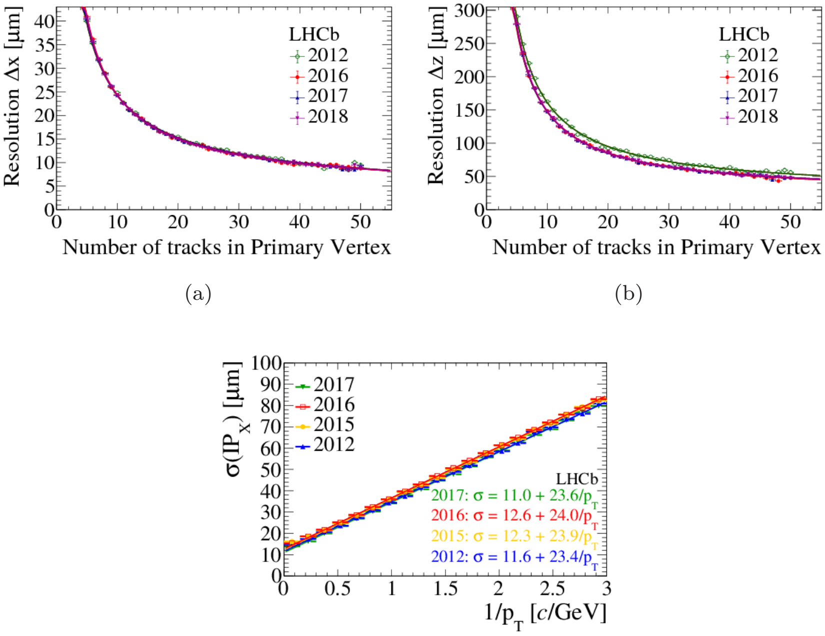
Figure 3.3: (a) A sketch in the (x, z) plane of the positions of the silicon submodules along the beam axis (top) and a front view of the first module in a closed and open configuration (bottom). (b) Sketch illustrating the r-ϕ geometry of a VELO sensor.

T1. The magnetic field, measured in (x, y, z) with precisions of $4 \cdot 10^{-4}$ in the volume of interest, is used to determine the expected trajectory of the charged particles and improve the performance of the tracking algorithms. Its polarity is regularly inverted during data-taking to avoid potential detection asymmetries, that could introduce systematic effects in measurements of CP observables.

3.2.1.2 VELO

The VELO detector [77] is a silicon detector designed for the precise measurement of track coordinates close to the interaction region and the location of primary and secondary vertices, which are fundamental aspects of the identification of heavy hadrons decays. The detector, whose geometry is sketched in Figure 3.3(a), is composed of 21 circular silicon modules, each of which provides a measurement of the radial and azimuthal coordinate, r and ϕ, of a *hit*. The two types of sensors have a thickness of 300 µm, and are characterized by strips with a pitch that increases linearly from 38 µm to 101.6 µm for the r-sensor and by strips that are disposed in a *dog-leg* design with a pitch that goes from 38(39) µm to 78(97) µm for the inner (outer) part of the ϕ sensor. A sketch of this geometry is shown in Figure 3.3(b).

The positioning of the modules along the beam axis has been optimized to ensure that any particle produced within 10.6 cm of the interaction point and inside the LHCb acceptance hits at least three modules, providing at the same time a short track extrapolation distance to the PV position. This last feature is so important that each of the circular modules is composed of two retractable halves that, during injection, are opened up to a radius of 3 cm but can be closed



(a) (b)

(c)

Figure 3.4: (*a*) and (*b*) show the primary vertex resolution in the x and z coordinate as a function of the number of tracks composing the vertex for the Run 1 offline and Run 2 (used both offline and online) PV reconstruction algorithms [87]. (*c*) shows the impact parameters resolution along the x-axis as a function of the inverse transverse momentum of the particle for Run 1 and Run 2. Similar performances are obtained along the y-axis [87].

down to 8 mm during physics runs. This has the additional advantage of greatly reducing the ageing of the silicon strips.

To minimize the amount of material between the primary and secondary vertices, the VELO is encapsulated in a secondary vacuum container, which preserves a void higher than in the rest of the LHC and reduces the effects that the beams could induce on the silicon modules. This keeps the average radiation length of the VELO below 18% of a radiation length.

The excellent performance observed in LHCb is justified by the careful design of the detector, that has allowed to measure positions of PVs composed of 20 tracks with a resolution of ∼100 μm (15 μm) along (perpendicularly to) the beam axis and impact parameters with resolutions better than 40 μm for tracks with transverse momenta of 1 GeV/c, as shown in Figure 3.4.

3.2.1.3 Silicon tracker

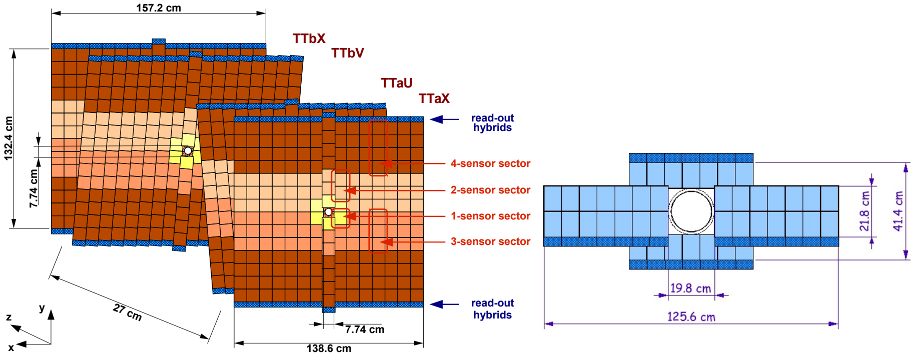
The Silicon Tracker [77] consists of two detectors: the Tracker Turicensis and the Inner Tracker. The first one is located upstream of the dipole magnet, and it has a planar geometry that covers the full LHCb acceptance. Thanks to a fringe field that leaks from the dipole magnet into the detector, it allows having a rough estimate of the momenta for those particles that are deflected outside the detector acceptance after the magnet. The second one constitutes the inner part of the tracking stations T1, T2 and T3, downstream of the magnet. Each of the four ST stations is composed of four detection layers, referred to as x, u, v and x, depending on the orientation of their silicon micro-strips. As shown in Figure 3.5(a) for the TT, in the first and last layer the strips are orientated vertically, while in the second and third layer the strips are rotated of an angle -5° and +5° with respect to the x-layer. This geometry is chosen to allow a faster track reconstruction.

Since the momentum resolution of the spectrometer is dominated by multiple scattering, it is fundamental to keep the material budget of the detector as small as possible. For this reason, the TT was designed to have the front-end readout electronics and mechanical support outside the acceptance of LHCb. Due to its central position, the same was not possible for the IT and a significant effort was made to keep the amount of material needed as small as possible. With the current geometry and the use of silicon micro-strip sensors with a pitch of 200 μm, it has been possible to reach single hit resolutions of 50 μm, needed to fulfil the tracking requirements of the experiment.

The TT layers are housed in the same light-tight, thermally and electrically insulated volume that ensures a constant temperature of 8°C, and are arranged in two pairs, (x-u) and (v-x), with a distance of 27 cm with respect to each other, as shown in Figure 3.5(a). Each layer is constituted of half-modules, covering half of the height of the TT and composed of a column of seven silicon sensors with a width of 9.64 cm, a length of 9.44 cm and a thickness of 500 μm. The sensors are electronically grouped in read-out sectors consisting of one, two, three or four sensors. This is done to keep the maximum strip occupancy within few percent.

The three IT stations are instead composed of four individual boxes, arranged in a cross-like structure around the beam pipe, as shown in Figure 3.5(b). Each box has four (x-u-v-x) layers composed of seven detector modules, consisting of one or two silicon sensors depending on the position of the box. The thickness of the sensors goes from 320 μm for single-sensor modules to 410 μm for two-sensor modules, while they all have the same size of 7.6 cm ×11 cm.

The four ST stations were designed to operate for ten years at nominal luminosity. The evolution of their radiation damage has been monitored using measurements of leakage currents and effective depletion voltages. At the end of RUN 2 the innermost sensors had not yet reached the point of type inversion, fulfilling the goal of their designers.



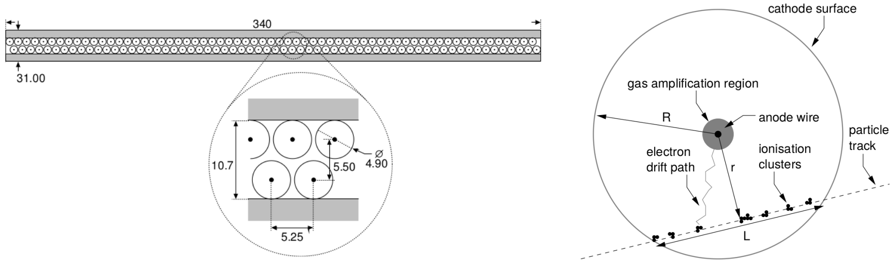
(a) (b)

Figure 3.5: (*a*) A schematic view of the different layers of the Tracker Turicensis, where the diffent orientation between the (x-u-v-x) is clearly represented.The four different types of read-out sectors employed are indicated by different shadings of orange and circled in red. [88]. (*b*) Layout of an x detection layer in the second IT station [89].

3.2.1.4 Outer Tracker

The Outer Tracker (OT) [77] is a gaseous drift-time straw tube detector, built for the tracking of charged particles and the measurement of their momentum over a large acceptance area. The detector is composed of three stations, subdivided in four layers where the orientation of the straw tubes follows the stereographic configuration (x-u-v-x) already discussed for the ST. Together with the IT, the OT forms the tracking stations T1, T2 and T3 shown in Figure 3.2, covering the full LHCb acceptance. The relative size between IT and OT was chosen to ensure OT occupancies below 10% in its inner parts.

Each layer consists of approximately 90 modules, where two staggered rows of gas-tight straw tubes are filled with a gas mixture of $\mathrm{Ar/CO_2/O_2}$, ensuring drift-times below 50 ns. As shown in Figure 3.6, each straw is characterized by a conductive cylinder with an internal diameter of 4.9 mm that acts as a cathode, while at the centre an anode wire with a diameter of 25 μm is set at a voltage of +1550V. A charged particle traversing the detector, will ionize the gas mixture along its trajectory. The electrons and ions produced, accelerated by the electric field in the cylinder, will drift towards the anode and cathode, respectively. By measuring the time it takes between the passage of the particle and the first signal in the straw detector, it is possible to estimate the drift-time and, consequently, the distance of closest approach of the charged particle with respect to the anode wire. This technique allows obtaining a hit resolution on x-coordinates of roughly 200 μm.



(a) (b)

Figure 3.6: (*a*) Sketch of the cross-section of an OT module [90]. (*b*) Schematic description of the interaction of a charged particle with a straw detector [90].

3.2.1.5 Track reconstruction

The hits deposited in VELO, TT, IT and OT stations by a charged particle can then be combined to reconstruct its trajectory in the detector. Depending on their path, it is possible to define different types of track, as sketched in Figure 3.7(*a*):

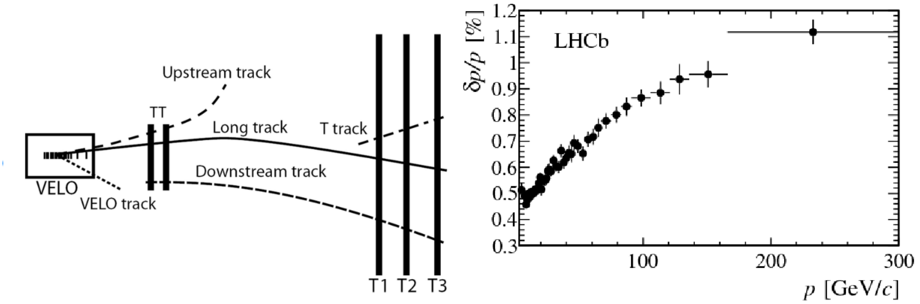
- *VELO tracks*, tracks reconstructed only inside the VELO and usually characterized by large angles that might also point backwards. They are especially useful for PV reconstruction.

- *Upstream tracks*, tracks reconstructed only in the VELO and TT stations and usually characterized by momenta so low that are ejected from detector after the magnet.

- *Downstream tracks*, tracks reconstructed only between the TT and T stations and fundamental for the identification of long-lived particles.

- *Long tracks*, tracks that traverse the full tracking systems leaving hits in the VELO and T stations, optionally in the TT.

Long tracks are the most important type of tracks and, since they have traversed the whole magnetic field and left hits over many different detector planes, they have the most precise estimate of momentum, with resolutions of about 0.5% for particles below 20 GeV/c, rising to about 0.8% for particles around 100 GeV/c (see Figure 3.7(*b*)). They are therefore the tracks usually used for physics analysis. Their reconstruction [91] starts with the search of straight line segments in the VELO, called VELO-*seed*, requiring at least three hits in r- and ϕ-sensors. From this seed, two complementary algorithms are used to build the final track. The first one, referred also as *forward tracking*, associates a seed with a single hit in the T station and uses the detailed knowledge of the magnetic field profile of the magnet to estimate the particle trajectory and look for additional hits in the



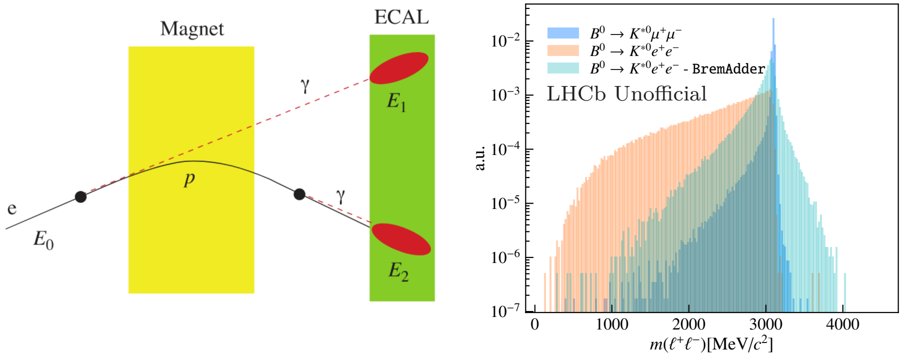
(a) (b)

Figure 3.7: (a) A schematic illustration of the various track types [77]: long, upstream, downstream, VELO and T tracks.(b) Relative momentum resolution as a function of momentum for long tracks reconstructed in data using $J/\psi \to \mu^+\mu^-$ decays [91].

remaining T stations compatible with it. The second one, called *track matching*, searches for straight segments in T stations exploiting their stereo geometry and matches them with the VELO-seeds. In both cases the trajectory is extrapolated in the TT and compatible hits are added to improve the momentum resolution of the track. The final parameters of the tracks are then recomputed using a *Kalman filter*, an algorithm mathematically equivalent to the least square fit and capable of taking into account the effects of multiple scattering and the energy lost ionizing the material encountered by the particle in its path. A χ^2 per degree of freedom, χ^2_{track}/ndf, is then used to determine the quality of the reconstructed track. In LHCb, particles with momenta greater than 10 GeV/c going through all stations have a probability to be reconstructed as long tracks of 94%, while the fraction of fake tracks, misreconstructed tracks that do not correspond to the trajectory of a real particle, can go from few percents up to 20% depending on the multiplicity of the event. This fake rate can be further reduced using a neural network classifier, `GhostProb`, based on the result of the track fit, track kinematics and the number of measured and expected hits in the different tracking stations.

3.2.1.6 Tracking and bremsstrahlung: the `BremAdder` algorithm

The momentum resolution shown in Figure 3.7(b) does not hold for all particle species, but it can degrade based on the specifics of the particle considered. For energies above 100 MeV/c^2, muons are minimum ionizing particles (MIP) and loose a small fraction of their energy in the tracking stations. Electrons, instead, interact with matter mainly through bremsstrahlung, where collinear photons are emitted along the trajectory of the lepton carrying away a sizable fraction of its energy. Depending on the region of the detector where this radiation has



(a)

(b)

Figure 3.8: (a) Schematic representation of the bremsstrahlung emission of an electron before and after the magnet. No representation of the emitted radiation inside the magnet is shown as it can be considered negligible [92]. (b) Comparison between the dilepton invariant mass distributions in simulated decays of $B^0 \to K^{*0} J/\psi (\to \mu^+\mu^-)$, in blue, and $B^0 \to K^{*0} J/\psi (\to e^+e^-)$, in orange and green. The two distributions represent the dilepton invariant mass obtained using, in orange, the momentum estimation directly from the tracks and, in green, the momentum corrected using the `BremAdder` algorithm.

been emitted, the quality of the electron momentum determined can vary. The emission can happen in three main regions: upstream, inside and downstream of the magnet, as sketched in Figure 3.8(a). Bremsstrahlung photons emitted downstream of the magnet usually do not affect the momentum determination of the electron since they leave its trajectory unaltered. Additionally, they usually end up in the same calorimeter cells of their corresponding electron, providing also an unbiased estimation for its energy. The radiation emitted by electrons due to the bending inside the magnet, usually referred to as *synchrotron radiation*, is very small and can be usually safely neglected. Bremsstrahlung photons emitted upstream of the magnet are instead the most dangerous ones, since the energy they carry away lowers the momentum of the electron before the magnet. As a consequence, the momentum estimated from the bending of the electron in the magnetic field after the emission will be biased. This leaves a clear sign when considering invariant mass shapes as the dilepton invariant mass $m(\ell^+\ell^-)$: a long radiative tail that is represented in Figure 3.8(b) by the yellow histogram. As expected, for the muons (in blue) the distribution is more symmetric and with a better resolution.

To mitigate this effect, LHCb has developed an algorithm, called `BremAdder`, that finds clusters in the ECAL that are compatible with a photon hypothesis and assigns them to electron tracks that might have produced them, partially recovering the lost energy. The procedure can be summarized with four main steps:

i. The tangents to the fitted electron track at the origin vertex and at the TT are extrapolated to the ECAL, defining a search window in the calorimeter.

ii. Photon candidates whose position is compatible with the search window defined are then associated with the considered electron track.

iii. The direction of the photon 4-momentum is then approximated using the direction of the 3D barycentre of the photon shower in the ECAL from the PV of the event, while its magnitude corresponds to the energy registered in the calorimeter.

iv. The photon 4-momentum is then added to the electron 4-momentum to correct for the emitted bremsstrahlung.

This correction improves the momentum resolution for electrons, as can be seen in Figure 3.8(b) in green. Here, the reconstructed invariant mass $m(e^+e^-)$ is more symmetric around the true value $M_{J/\psi}$, with a better invariant mass resolution. Additionally, electrons that are corrected with one or more photons are characterized by a better PID performance with respect to the electrons, for which no photon candidate was found.

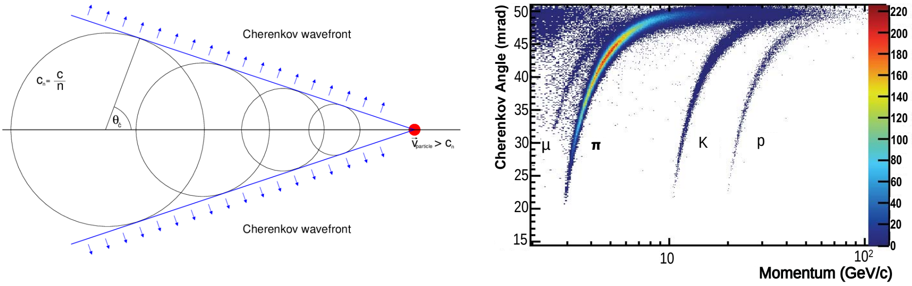
3.2.2 Particle identification systems

The particle identification systems are mainly built to provide additional information on the identity of a reconstructed track or to detect the presence of neutral particles. They are composed of two Ring Imaging CHerenkov detectors (RICH) placed before and after the magnet, a PreShower (PS) and Scintillator Pad Detector (SPD), an electromagnetic calorimeter (ECAL), a hadronic calorimeter (HCAL) and five muon stations (M1-M5). The last three elements of the detector can be thought as a series of successive destructive filters for electrons and photons, hadrons and muons, respectively.

3.2.2.1 RICH detectors

The Ring Imaging CHerenkov (RICH) [77] detectors of LHCb provide information on the identity of the charged particles that traverse them thanks to their emission of Cherenkov radiation. Cherenkov radiation is produced every time a charged particle in a material medium moves faster than the speed of light in that medium. This creates an electromagnetic shockwave, as sketched in Figure 3.9(a), consisting of a coherent wavefront of conical shape that is emitted at a well-defined angle θ_C, with respect to the trajectory of the particle. This angle depends on the speed of the particle as described by

$$\cos\theta_C = \frac{1}{\beta n}\,, \tag{3.1}$$



(a)

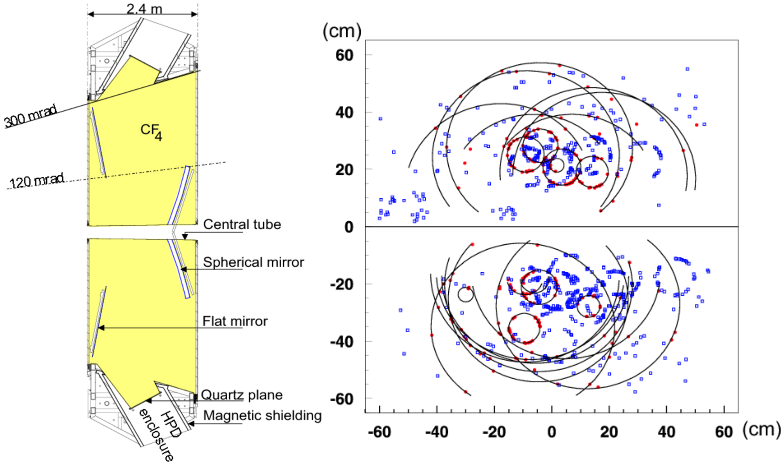
(b)

Figure 3.9: (*a*) Sketch of the geometrical construction of the Cherenkov wavefront [93]. (*b*) Reconstructed Cherenkov angle for isolated tracks, as a function of track momentum in the C_4F_{10} radiator. The Cherenkov bands for muons, pions, kaons and protons are clearly visible [94].

where n is the refractive index of the medium and $\beta = v/c$, with v being the speed of the particle and c the speed of light in vacuum. The identity of the particle can then be inferred by the momentum information obtained from the tracking stations combined with its Cherenkov angle θ_C, as shown in Figure 3.9(*b*).

The main goal of the RICH detectors in LHCb, composed of RICH1 and RICH2, is to distinguish kaons, protons, electrons and muons from pions in a range of momenta that goes from 1 to 100 GeV/c. RICH1 is located upstream of the magnet, between the VELO and the TT, covering the angular region from ±25 mrad to ±300 (250) mrad in the bending (non-bending) plane. Its volume is filled with C_4F_{10}, a gaseous radiator with refractive index of $n = 1.0014$, covering the low momentum range from 1 to 60 GeV/c. RICH2 is instead located downstream of the magnet, between the last tracking station and the first muon station, covering the angular region from ±15 mrad to ± 120(100) mrad in the bending (non-bending) plane. Its volume is filled with a CF_4, a gaseous radiator with refractive index of $n = 1.0005$, covering the high momentum range of 15 to 100 GeV/c.

In both detectors, the Cherenkov photons emitted are reflected and focussed outside the spectrometer acceptance by a set of spherical and flat mirrors, whose geometry is shown for RICH2 in Figure 3.10(*a*). The light is then detected by Hybrid Photon Detectors (HPD), where the characteristic circular pattern shown in Figure 3.10(*b*) can be used to estimate θ_C. Also for the RICH systems, the material budget has been kept at minimum, with a total radiation length of 8% and 15% for RICH1 and RICH2, respectively. This was achieved positioning the HPDs outside the acceptance, since they need to be surrounded by iron shields to permit their proper operation in the fringe field of the LHCb magnet.



(a) (b)

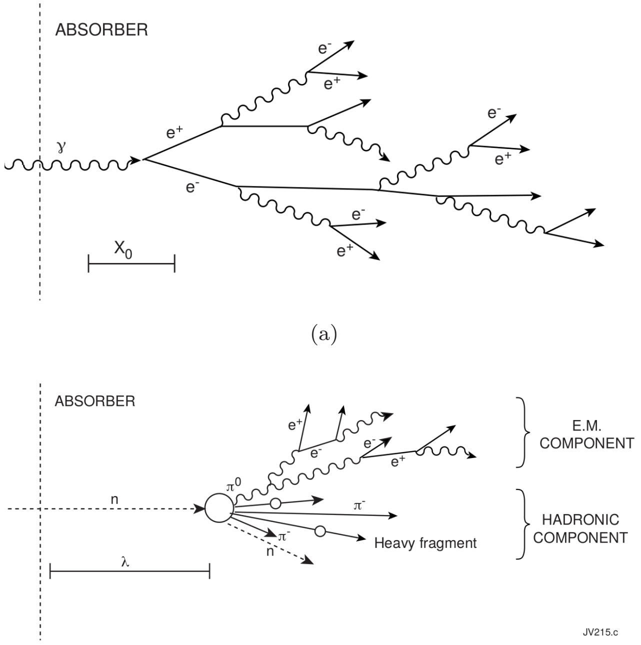
Figure 3.10: (*a*) Schematic of the RICH2 detector [77].(*b*) Display of a typical LHCb event in RICH1 [77].

3.2.2.2 Calorimeters

The calorimeter system allows the identification of electron, photons and hadrons and the measurement of their energy and position. This information is fundamental in the operation of the LHCb experiment, since they are heavily used by the first level trigger (L0) for the selection of candidates with high transverse energy. The calorimeter is composed of three subsystems: the Scintillator Pad Dectector/PreShower (SPD/PS), the Electromagnetic CALorimeter (ECAL) and the Hadronic CALorimeter (HCAL). The ordering of the last two calorimeters follows a classical geometry and is motivated by the different ways electrons, photons and hadrons interact with matter.

For electrons and photons with energies above 1 MeV, the interaction with matter is dominated by bremsstrahlung and pair-production, respectively. Electrons (photons) entering a dense material will therefore interact with its nuclei producing photons (electron-positron pairs), as shown in Figure 3.11(*a*). The energy of the initial particles is therefore split into an *electromagnetic cascade* or *shower* until the energy of its final components is small enough to be dominated by either ionization or Compton scattering (photoelectric effect). The longitudinal size of this shower can be described by the electromagnetic radiation length, $X_0 \approx 180A/Z^2 \mathrm{g\ cm}^{-2}$. This quantity corresponds to (9/7 of) the average distance that an electron (photon beam) needs to travel in material to reduce its energy (intensity) of $1/e$.

Analogously to photons and electrons, hadrons interact with matter producing hadronic showers, whose development is mainly caused by strong interactions. These interactions usually produce many secondary hadrons that in turn interact



(a)

(b)

Figure 3.11: Schematic of the development of electromagnetic (*a*) and hadronic (*b*) showers [96].

with further nuclei, leading to a growth in the number of particles in the cascade. Two types of cascades are possible: electromagnetic or hadronic, as sketched in Figure 3.11(*b*). Also in this case, the shower development can be described by an average mean free path ("hadronic interaction length") between inelastic collisions, $\lambda \approx 35A^{1/3}$g cm^{-2} [95]. By comparing X_0 and λ, it is possible to note that the size of electromagnetic showers is in general much smaller than the hadronic ones. This justifies the standard choice of positioning the electromagnetic calorimeter before the hadronic one.

The SPD/PS detector is positioned in front of the ECAL and consists of two almost identical planes of rectangular scintillator pads divided by a 15 mm thick lead converter, with interactions lengths of $\sim 2.5X_0$ and $\sim 0.06\lambda$. Electrons and photons have a much higher probability to "shower" than hadrons, allowing to distinguish between them by looking at the energy deposited in the PS, as shown in Figure 3.12 on the left. Photons and electrons can then be distinguished based on the absence or presence of an energy deposit in the corresponding SPD pad. The information on the number of hits registered in the SPD (`nSPDHits`) is additionally used in the trigger decision to estimate the multiplicity of the event observed.

The ECAL is a sampling calorimeter built with a "shashlik" geometry, where

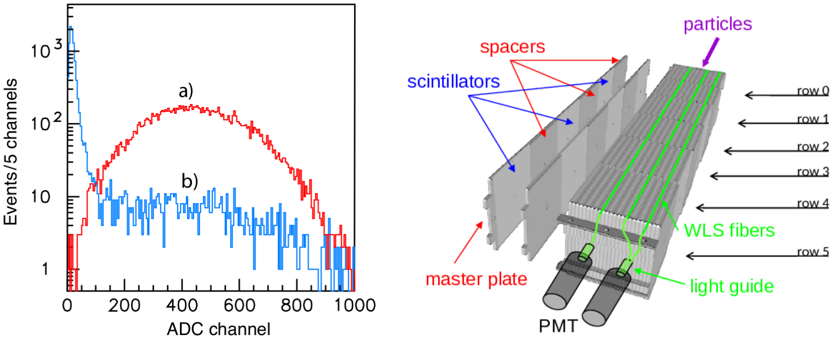
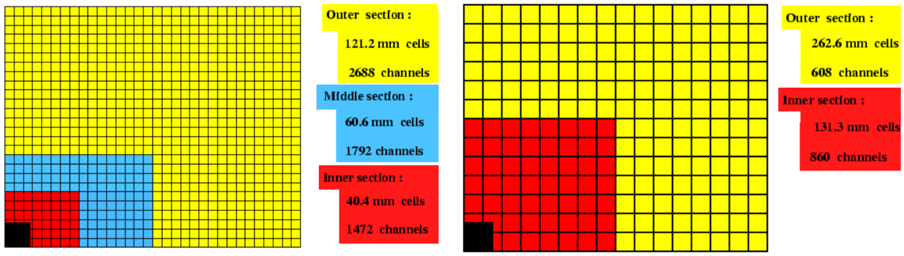


Figure 3.12: On the left, the energy deposition of (a) 50 GeV electrons and (b) pions in the PS [77]. On the right, an exploded view of two scintillator-absorber layers illustrating the periodic structure of a HCAL module [77].

4 mm thick scintillating tiles are alternated to 2 mm thick lead tiles, for a total length of 84 cm. This length, corresponding to $25X_0$ and 1.2λ, ensures a full containment of the electromagnetic showers, allowing a measurement of the energy of electrons and photons with a resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$(E in GeV) [77]. The HCAL is a sampling calorimeter where iron and scintillating tiles of 16 mm and 4 mm are alternated, acting as absorber and active material, respectively. Its geometry is characterized by an orientation of the scintillating tiles parallel to the beam axis rather than perpendicular. In particular, as shown in Figure 3.12(b), the scintillating tiles are interspaced in the lateral direction with 1 cm iron whereas the longitudinal dimension of the iron and scintillating spacers corresponds to the hadron interaction length in steel. Due to space limitations, its length along the beam pipe is 1.65 m, corresponding to roughly 5.6λ, not enough to fully contain the hadronic showers started inside the calorimeter. Since a sizable fraction of the energy of the hadronic shower is lost, the HCAL is used mainly to provide a trigger for charged hadrons with high transverse energy, rather than an unbiased measurement of the energy of the hadronic showers.

In all three calorimeters the signal is collected following the same principle: the light produced by the energy deposited in the scintillating tiles is transmitted to the photomultiplier tubes (PMTs) by wavelenght-shifting fibres. The dimension of the calorimeters are chosen to match projectively those of the tracking systems, with an exception made for the region between 10 and 25 mrad that remains uncovered due to the high radiation level. To keep the occupancy of the detectors more constant across their surfaces, the SPD/PS and ECAL are subdivided into inner, middle and outer sections as shown in Figure 3.13(a), while the HCAL is divided in inner and outer sections as shown in Figure 3.13(b).



(a) (b)

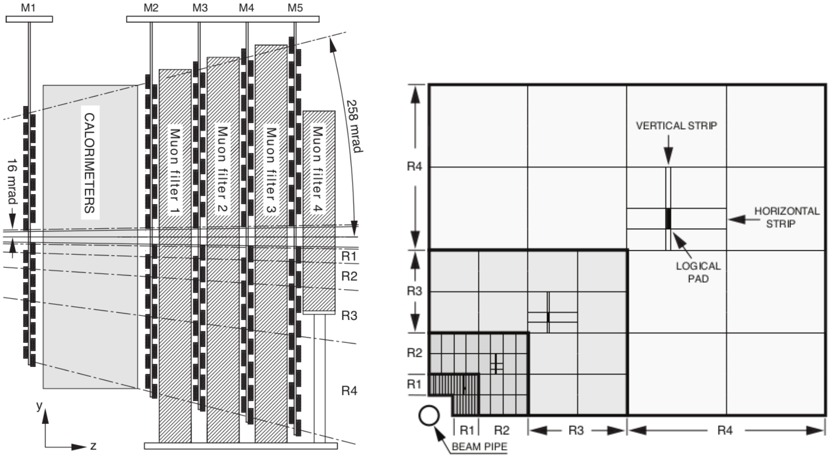
Figure 3.13: Calorimeter cells segmentation of the SPD, PS, and ECAL (a), and the HCAL (b) [77].

3.2.2.3 Muon system

The muon system [77] provides fast information for the high-p_T muon trigger at the earliest level, muon identification for the high-level trigger and additional hits for the track reconstruction in muons. Since muons enter in the final states of many CP-sensitive or FCNC decays, such as the $B^0 \to K^{*0}\mu^+\mu^-$ decays discussed in this work, it is a fundamental part of the LHCb detector. It is composed of five stations of rectangular shape, referred to as M1-M5, positioned in two main regions, as shown in Figure 3.14(a): M1 is placed in front of the calorimeters while the remaining are positioned after and interleaved with 80 cm thick iron absorbers, called *muon filters*. Their position at the end of the spectrometer is justified by the fact that muons in LHCb are MIP and can therefore traverse many meters of material before loosing all their energy. For this reason, they are the most penetrating charged particles in LHCb and they can be selected by simply trying to progressively stop the remaining hadrons after the calorimeters. M1 is an exception on this regard: it is placed in front of the calorimeters to provide the muon stations with a hit minimally affected from multiple scattering, improving the estimation of the transverse momentum of muons used in the trigger.

As shown in Figure 3.14(b), each muon station is divided in four regions, referred to as R1-R4. The linear dimensions and segmentation of the four regions scale as 1:2:4:8 depending on their distance from the beam axis, ensuring comparable particle fluxes and channel occupancies in the four regions. Similarly to the other detectors in LHCb, the transverse dimensions of the muon stations scale with their distance from the interaction point, covering an inner and outer acceptance from ±20(16) mrad to 306(258) mrad in the bending (non-bending) plane and allowing the collection of roughly 20% of the total muons produced from semileptonic decays of b-hadrons.

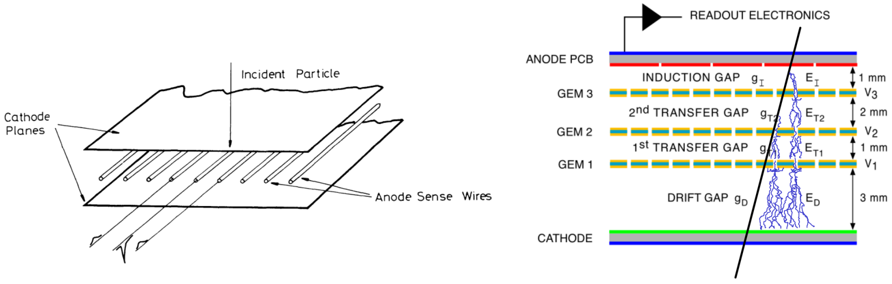
Two main technologies have been used to build these stations: multiwire proportional chambers (MWPC) and triple gas electron multipliers (triple-GEM).



(a) (b)

Figure 3.14: (*a*) Side view of the muon system [77]. (*b*) Front view of one of the quadrants in stations M2 and M3, showing the partitioning into sectors and the definitions of the logical pads [77].

The first one is by far the dominant one, with a total of 1380 chambers used, while the second one has been installed only in the inner region of M1 to better withstand its high particle fluxes. MWPC are gaseous detectors that, in their simplest design, are characterized by a plane of equally spaced anode wires centred between two cathode planes, as shown in Figure 3.15(*a*). In LHCb, the wire plane is characterized by an inter-wire spacing of 2 mm placed symmetrically in a 5 mm gas gap filled with a fast gas mixture of $\mathrm{Ar}/\mathrm{CO}_2/\mathrm{CF}_4$ in the ratios 40/55/5, allowing for a time resolution of 5 ns. The triple-GEMs are also gaseous detectors, but they are composed of three gas electron multipliers placed in between an anode and cathode plane. As it is shown in Figure 3.15(*b*), when a charged particle enters the triple-GEM it ionizes the gas inside it. The electrons produced in the drift gap are then attracted by the electric fields through holes etched in the different GEM foils, where they are multiplied. The electron cascade drifts then to the anode, inducing a current signal on the pads. The gas mixture used in LHCb for the triple-GEM is $\mathrm{Ar}/\mathrm{CO}_2/\mathrm{CF}_4$ in the ratios 45/15/40, allowing to obtain time resolutions smaller than 3 ns. Each station is partitioned into fine rectangular logical pads, whose dimensions scale depending on the region of the detector considered, as shown in Figure 3.14(*b*). When they are traversed by a charged particle, they provide a hit position in the x-y plane. This information is then used by the muon trigger where, after having required five aligned hits coming from all muon stations, a muon track is reconstructed together with its transverse momentum with respect to the interaction point, obtaining a resolution of roughly 20%. Most of the spatial resolution comes from



(a) (b)

Figure 3.15: (*a*) Basic configuration of a multiwire proportional chamber, where every wire acts as an independent proportional counter [97]. (*b*) Schematic cross-section of a triple-GEM detector [77]. Superimposed, a sketch of the passage of a charged particle, together with the drifting trajectory of the electrons produced due the ionization of the gas.

the first three stations, while the last two are mainly used to identify penetrating particles. The characteristics of the muon stations are such that allow LHCb to obtain trigger efficiencies of 95% in a window that is smaller than 25 ns.

3.2.2.4 Combined particle identification

The knowledge of the reconstructed tracks obtained from the track reconstruction algorithm, together with the information extracted from the RICH detectors, calorimeters and muon stations, can be used to determine the identity of the charged particles observed. Only five particle species are considered for this identification: electrons, muons, pions, kaons and protons. As it is schematically shown in Figure 3.16, their interaction with the different subsystems is quite different and can be summarized as:

- Electrons, being so light, when passing through the RICHs emit light with the maximum allowed angle, $\theta_C \sim 1/n$. Additionally, they deposit energy in the SPD and ECAL where they are completely stopped. For this reason no additional response is present in HCAL or muon stations.

- Muons traverse the whole spectrometer loosing small amounts of energy in the detector elements they encounter. They can be hardly distinguished from pions in the RICHs, however their difference becomes clearer when interacting with the HCAL and the muon stations. Protons and pions are the particles that are most often misidentified as muons: it usually happens due to a combination of spurious hits aligned in the muon stations with genuine hadron tracks or due to the existence of a true muon either pointing in their same direction or produced by the decay in flight of one of the hadrons.

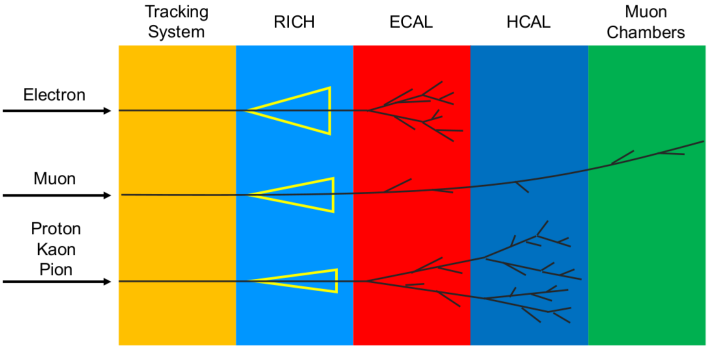


Figure 3.16: Illustration of the responses that different particles have in the LHCb systems [98].

- Hadrons have similar responses in the calorimeters, where a sizable fraction of their energy is deposited, and their separation is mainly achieved by the RICH systems.

From the information collected from these subsystems, two sets of Global PID variables are obtained: `DLL` and `ProbNN`. In the first one, signal and background likelihood distributions are built for each particle species and subdetector. For the RICHs, the ring size and the momentum information of the associated track t are used to build a likelihood function $\mathcal{L}_i^{\text{RICH}}(t)$ for the i-th particle hypothesis. The energy deposits in the calorimeters systems, together with the momentum of the associated track t are used to build the likelihood functions $\mathcal{L}_i^{\text{CAL}}(t)$, where i represents the hypothesis of being or not-being an electron. Likelihoods describing the probability of track being muon or not, $\mathcal{L}_{\mu,not-\mu}^{\text{MUON}}(t)$, are computed based on the average square distance between the associated track and the hits in the muon stations that are closest to the extrapolation points. The number of muon stations used for this calculation is a function of the measured momentum of the track, since only muons with momenta above 6 GeV/c are able to reach the last station M5[2]. The three likelihoods are then combined by simple multiplication in the following way:

$$\mathcal{L}_\mu^{tot}(t) = \mathcal{L}_\mu^{\text{RICH}}(t) \cdot \mathcal{L}_{not-e}^{\text{CAL}}(t) \cdot \mathcal{L}_\mu^{\text{MUON}}(t) \tag{3.2}$$

$$\mathcal{L}_e^{tot}(t) = \mathcal{L}_e^{\text{RICH}}(t) \cdot \mathcal{L}_e^{\text{CAL}}(t) \cdot \mathcal{L}_{not-\mu}^{\text{MUON}}(t) \tag{3.3}$$

$$\mathcal{L}_h^{tot}(h) = \mathcal{L}_h^{\text{RICH}}(t) \cdot \mathcal{L}_{not-e}^{\text{CAL}}(t) \cdot \mathcal{L}_{not-\mu}^{\text{MUON}}(t) \tag{3.4}$$

From these it is possible to build the `DLL` variables, defined as the difference between the log-likelihood of a track t belonging to the particle species j rather

---

[2] As it will be seen in Sec. 5.2.2 an additional variable, called `IsMuon` can be defined for the preselection of muon candidates. It corresponds to a loose binary selection that is exclusively based on the number of muon stations where a hit is found within a certain spatial window around the track extrapolation.

than being a pion:

$$\mathrm{DLL}_j(t) = \log \mathcal{L}_j^{tot}(t) - \log \mathcal{L}_\pi^{tot}(t)\,. \tag{3.5}$$

The second approach aims to improve on the previous work by exploiting the categorization power of six binary one-layer artificial neural networks (NN), one for each of the five particle species plus the introduction of a *ghost* category. The NNs are trained using simulation, providing as input information from the tracking systems, RICHs, calorimeters and muon stations. The `ProbNN` variables obtained from the NN correspond to the probability that the track considered was produced by one of the six particle categories considered. Their performance is overall better than `DLL` variables, especially for what concerns kaon-proton separation [98].

3.2.3 Trigger

Of the 40 million proton-bunch crossings that each second happen in LHCb, only a small fraction produces an event where at least a B meson decays with all its products inside the spectrometer. The role of the LHCb trigger systems is to retain as many charm and beauty hadron decays as possible, rejecting at the same time uninteresting or background decays. This needs to be done taking into account the real-life limitations of computing power, data storage capacity and rates. In LHCb this is achieved with two trigger levels: Level-0 (L0) and High Level trigger (HLT) [77, 99].

The L0 trigger is a hardware-trigger, implemented using custom-made electronics capable of operating synchronously with the bunch crossing frequency of the LHC. Since B mesons have a large mass, their daughters are usually characterized by large transverse momenta and energies. These features are exploited by the L0 trigger to decide which events to keep and reject, effectively reducing the rate to roughly 1 MHz, frequency at which the whole detector can be read out. The L0 trigger is divided into three independent units: L0-Calorimeter, L0-Muon and L0-PileUp trigger. Each component, described briefly in the following, provides information to the L0 Decision Unit (L0 DU), where the final decision on the rejection of the event is taken. The L0-Calorimeter unit combines the information obtained from SPD/PS with the transverse energy E_t [3] measured in ECAL and HCAL, providing different trigger candidates. The most relevant in this work are:

- `L0Hadron`, corresponding to the candidate with the highest E_t deposit in HCAL, to which is also added the energy observed in the matching cells of ECAL;
- `L0Electron`, corresponding to the candidate with the highest E_t deposit in ECAL, with hits in the corresponding SPD/PS cells in front of it.

---

[3] The transverse energy is defined as $E_t = E\sin\theta$, where E is the observed deposited energy in a cluster of 2×2 cells and θ is the angle between the z-axis and the center of the 2×2 cells with respect to the primary vertex.

A candidate passes the `L0Hadron` or `L0Electron` selection if its transverse energy is above a threshold given by the specifics of the run considered. Typical values of these two thresholds in 2012 correspond to 3.6 GeV/c^2 and 2.7 GeV/c^2, respectively [100]. The L0-Muon unit computes the transverse momentum of all candidates leaving hits in the muon stations, retaining only the two candidates with highest p_T for each quadrant of the detector. As seen before, a candidate passes the `L0Muon` selection if its transverse momentum is above a threshold, that for 2012 was typically 1.8 GeV/c [100]. The L0-PileUp trigger, using two sets of overlapping VELO r-sensors positioned upstream of the VELO detector, estimates the number of interactions produced per bunch-crossing and vetoes events where more than a single interaction is observed. To further reduce the contribution of high multiplicity events, a further cut on `nSPDHits` is applied, requiring them to be below 600 and 450 in Run 1 and Run 2, respectively. This has the advantage of removing events characterized by a difficult reconstruction, saving time for the High Level Trigger.

The HLT is composed of two software applications, HLT1 and HLT2, that run asynchronously on a processor farm using commercially available equipment. Their aim is to reduce the rate of events obtained from the L0 trigger to $\approx$ 5 (10)kHz for Run 1 (Run 2), allowing their offline storage. For each of the types of L0 trigger selections available, the HLT1 algorithm performs a partial reconstruction of the tracks in the VELO and T-stations, confirming the decision at the L0, and performing additional selections on the quality of the reconstructed tracks. The events, the tracks and vertices obtained are then passed on to the last step of the trigger, HLT2, where a full event reconstruction can be carried out. Thanks to the fact that the two HLT levels can be run independently, is possible to obtain high-quality alignment and calibrations before this last step is performed, ensuring reconstructed events ready to be used in the offline analysis already after the trigger. The reconstruction is followed by mixture of several exclusive and inclusive selections designed to collect specific types of events. The selections that are most interesting for this work are the topological lines, referred to as `Hlt2Topo[2,3]BodyBBDT` (`Hlt2Topo[2,3]Body`) in Run 1 (Run 2), and designed to cover all b-hadron decays with at least two or three charged particles in the final state and a displaced decay vertex. These lines correspond to a cut- or MVA-based selection on discriminating variables such as the minimum transverse momenta, the minimum χ^2_{IP}, the minimum and maximum track χ^2, the sum of the transverse momenta, the distance-of-closest-approach (DOCA) of all input particles, to name just a few. Additional lines such as `Hlt2TopoE[2,3]Body`, `Hlt2TopoEE[2,3]Body`, `Hlt2TopoMu[2,3]Body` and `Hlt2TopoMuMu[2,3]Body` are also used and built similarly, but include some additional electron and muon identification information. More information can be found in Ref. [101].

Chapter 4

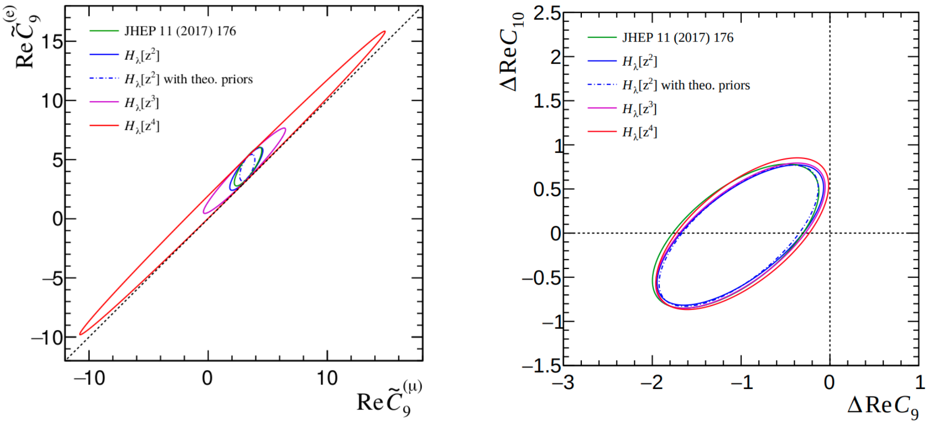
Analysis strategy

The tensions observed in the measurement of the branching fraction and angular observables of $B^0 \to K^{*0}\mu^+\mu^-$ decays, together with the remaining set of B *anomalies*, have caused a great excitement in the physics community, that lead to the publication of several models capable of explaining the deviations observed as hints of physics beyond the SM [102–120]. However, the vectorial nature of the observed pattern could also be explained by an unexpectedly large contribution from $b \to sc\bar{c}$ operators, which could either mimic or camouflage any genuine NP effect [41, 42, 45–51]. Attempts to improve the current understanding of these hadronic contributions by performing an unbinned amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ decays including an explicit parametrization of the charm-loop contribution from data-driven analysis [121–123] or exploiting the analytical properties of its structure [23, 124, 125] have been recently proposed and are currently under study. Their sensitivity to NP is still however intrinsically limited by the assumptions that enter the parametrization of the dilepton invariant mass distribution.

A complementary strategy to investigate the set of B *anomalies* is the study of LFU tests such as R_{K^*} [52–54] and Q_i [38], which correspond to ratios of branching ratios and differences in the angular observables for the $B^0 \to K^{*0}\ell^+\ell^-$ modes that involve muons or electrons. Their precise theoretical prediction results from the fact that local (form factors) and non-local (charm loop) hadronic contributions, being derived from QCD, are expected to be the same for different lepton generations, resulting in a convenient cancellation of all hadronic uncertainties. While the tensions observed in this class of measurements are still inconclusive, when joined with the other deviations observed in $b \to s\,\ell^+\ell^-$ transitions, they give rise to a non-trivial and coherent pattern, as it is shown by global fit analyses.

In the following, an attempt to enhance the sensitivity of previous LFU measurements in $B^0 \to K^{*0}\ell^+\ell^-$ decays is presented. It relies on a simultaneous unbinned amplitude fit to the full decay rate of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays, benefitting from the full description of the two decays and combining the discovery potential of the ratio R_{K^*} and the difference of the angular observables Q_i in a single quantity. The key feature of this approach, presented for the first time in Ref. [126], corresponds to the description of the two differential decay rates using the exact same parametrization, except for the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$, which are parametrized separately between the two lepton generations. The hadronic contributions are identical for muon and electron channels and are treated as nuisance parameters obtained mainly from the large statistics of the former. Their description relies, for the $B^0 \to K\pi$ form factors, on a parametrization that combines the predictions available at low and high q^2 using QCD Light-Cone Sum Rules and Lattice QCD [34], respectively; while the charm-loop is described as a combination of singular terms, which describe the analytical properties of $\mathcal{H}_\lambda(q^2)$ at the mass of the J/ψ and $\psi(2S)$ resonances, and a simple polynomial expansion truncated at the second order [23]. More details on the exact parametrization used for the description of these hadronic matrix elements are provided in Appendix A and B.

The sensitivity of this analysis has been previously studied on a series of signal-only simulations, generated under the NP scenario $\tilde{\mathcal{C}}_9^\mu - \tilde{\mathcal{C}}_9^e = -1$, where the usual WCs $\mathcal{C}_i$ are renamed as $\tilde{\mathcal{C}}_i$, since an accurate disentanglement between the physical meaning of the WCs and the hadronic pollution cannot be achieved at the current stage of the theory. $\mathcal{C}_7$ is also included in the fit but fixed to its SM value, due to its strong constraint from radiative decays [127]. The sensitivity of the analysis to the WC $\tilde{\mathcal{C}}_9^{\mu,e}$ and $\tilde{\mathcal{C}}_{10}^{\mu,e}$ has then been tested considering different parametrizations for the non-local hadronic contributions, as shown in Figure 4.1(a). As it can be seen, the sensitivity to $\tilde{\mathcal{C}}_9^\mu$ and $\tilde{\mathcal{C}}_9^e$ is strongly dependent on the model assumptions used for the non-local matrix element. This should not be of surprise, since it is exactly the reason why a direct determination of the WCs in the muon channel alone is so difficult. However, it is noticeable that the high correlation of $\tilde{\mathcal{C}}_9^\mu$ and $\tilde{\mathcal{C}}_9^e$ is sufficient to preserve the true underlying physics regardless of the hadronic parametrization used: the two-dimensional pull estimator with respect to the LFU hypothesis (dashed diagonal line) is unbiased.



(a) (b)

Figure 4.1: Two-dimensional sensitivity scans for the pair of Wilson coefficients $\tilde{\mathcal{C}}_9^\mu$, $\tilde{\mathcal{C}}_9^e$ (a) and $\Delta\mathcal{C}_9$, $\Delta\mathcal{C}_{10}$ (b) for different non-local hadronic parametrisation models evaluated in the NP scenario $\Delta\mathcal{C}_9 = -1$ and with the expected statistics after LHCb Run 2 [126]. The contours correspond to 3σ statistical-only uncertainty bands and the dotted black line indicates the LFU hypothesis.

This remarkable feature lead to the definition of two new parameters of interest: the difference in WC between muons and electrons

$$\Delta\mathcal{C}_i = \tilde{\mathcal{C}}_i^{\mu} - \tilde{\mathcal{C}}_i^{e}\,. \tag{4.1}$$

In these observables, the cancellation of the hadronic uncertainties becomes manifest: if an erroneous ansatz for the non-local hadronic matrix element is used in the parametrization of the amplitude, it biases the observed value of the Wilson coefficients by the quantity $\mathcal{C}_i^{hadr}$. However, since the parametrization is shared, the same bias is expected for all lepton generations. Therefore, the difference between the WCs becomes

$$\Delta\mathcal{C}_i = \tilde{\mathcal{C}}_i^{\mu} - \tilde{\mathcal{C}}_i^{e} = \mathcal{C}_i^{\mu,true} - \cancel{\mathcal{C}_i^{hadr}} - \mathcal{C}_i^{e,true} + \cancel{\mathcal{C}_i^{hadr}} = \mathcal{C}_i^{\mu,true} - \mathcal{C}_i^{e,true} = \mathcal{C}_i^{\mathrm{NP},\mu} - \mathcal{C}_i^{\mathrm{NP},e}\,, \tag{4.2}$$

where $\mathcal{C}_i^{\ell,true} = \mathcal{C}_i^{\mathrm{SM}} + \mathcal{C}_i^{\mathrm{NP},\ell}$ is the combination of the WC predicted by the SM together with a possible NP contribution. As can be seen, only a LFU-breaking NP contribution could cause a difference between the WCs observed in two different lepton generations. When investigating the sensitivity as a function the observables $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, as shown in Figure 4.1(b), the cancellation becomes evident. The sensitivity is found to be completely independent not only on the parametrization of the non-local hadronic contributions, but also on the form factors uncertainties. Additionally, the inclusion of the observed signal yield as a constraint on the WCs that describe the differential decay rate has been shown to stabilize the fit and increase the sensitivity to NP, as shown in Figure 4.2.

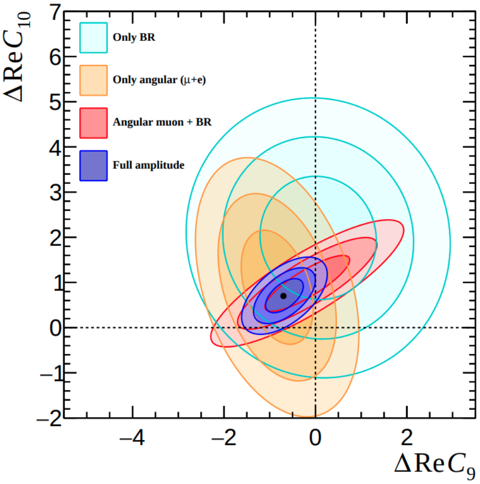


Figure 4.2: Statistical sensitivity projected (signal only) for the current LHCb dataset to the difference in WCs between muons and electrons, under the NP hypothesis of $\Delta\mathcal{C}_9 = -\Delta\mathcal{C}_{10} = -0.7$, taken from Ref. [126]. The different analysis approaches and corresponding sensitivities are overlaid.

Overall, the method discussed allows to exploit the full set of observables available in $B^0 \to K^{*0}\ell^+\ell^-$ decays, providing an unprecedented precision of LFU in a single analysis. Starting from the best fit values of $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, the branching ratios and angular observables can then be derived as function of q^2 or binned in q^2 and $m(K\pi)$ to compare against the values measured in previous analysis and the corresponding theoretical predictions. Such a comparison is shown in Figure 4.3, as an example, for the binned angular observables F_L, P_5' [29], the differential branching ratio [33] for $B^0 \to K^{*0}\mu^+\mu^-$ decays and the LFU test R_{K^*} (in black). The bands are obtained using signal-only toys, where the best fit value considered correspond to the SM and two other NP scenarios (green, blue and red).

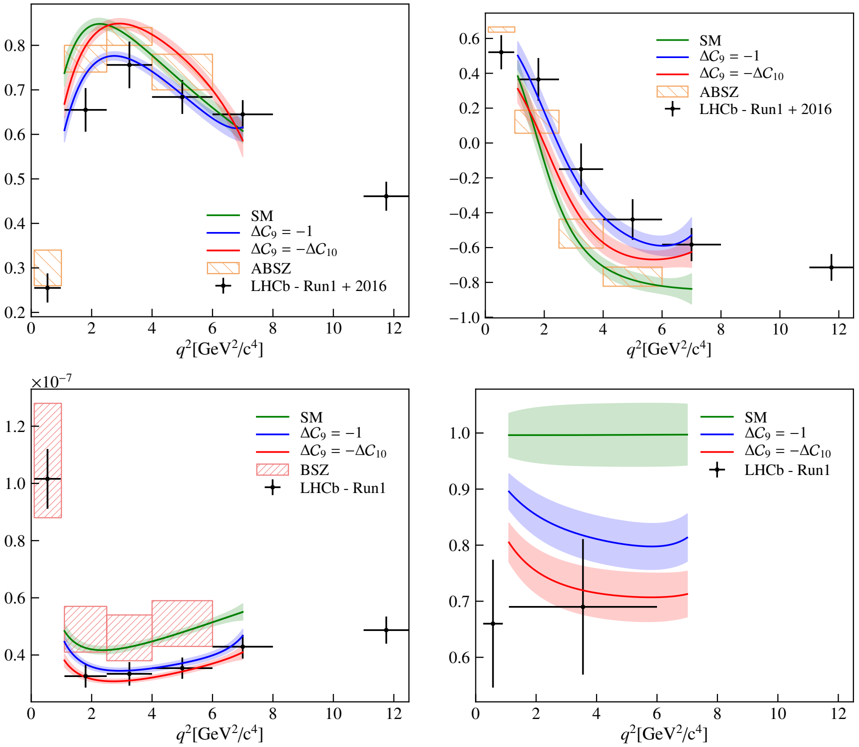


Figure 4.3: Comparison between the values of the measured binned observables F_L, P_5'[29] and the differential branching ratio [33] for $B^0 \to K^{*0}\mu^+\mu^-$ (black) and the values obtained from the signal parametrization under the scenarios of `SM`, $\Delta\mathcal{C}_9 = -1$ and $\Delta\mathcal{C}_9 = -\Delta\mathcal{C}_{10}$ for a set of signal-only toys. The bottom right plots correspond to the $R_{K^{*0}}$ observable. Additionally, the theoretical predictions ABSZ [34, 104] and BSZ [34] are overlaid.

4.1 Analysis overview

The main goal of this work is to determine the observables $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ directly from data, performing an unbinned amplitude analysis of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays, using the datasets collected at the LHCb experiment during Run 1 and Run 2, corresponding to a total of $9\,\mathrm{fb}^{-1}$.

The main experimental challenges of the analysis originate in the different behaviour that electrons and muons have when interacting with the material of the spectrometer, due to their different masses. At the energies of a typical $B^0 \to K^{*0}\ell^+\ell^-$ decay, the momentum of the leptons is in a range between 0.1 and $100\,\mathrm{GeV}/c$. In this regime, a muon is a minimum ionizing particle, going through the detector and loosing a small fraction of its total energy. Being so penetrating, the muon will likely leave a set of aligned hits in the muon stations, allowing the `L0Muon` trigger to easily recognize its signature. For an electron with the same momentum, the interaction with matter is instead dominated by bremsstrahlung emission. If these photons are emitted before the magnet, the momentum estimated from the bending of the electron trajectory will be biased. This effect is mitigated by the use of the `BremAdder` algorithm, described in Sec. 3.2.1.6, that is capable of partially recovering the energy radiated by associating compatible photon clusters in the ECAL back to the corresponding electron from which the radiation originated. Additionally, due to the high number of particles that shower in the ECAL, only those electrons that deposit transverse energies above 2.4-3 GeV/c^2 are able to fire the trigger and saved for offline analysis.

As a consequence, decays involving muons or electrons in the final state differ mainly in two fundamental aspects: the statistics available for the two collected modes, which due to the higher trigger efficiency, is roughly four times larger for muons than for electrons, and the resolution of the measured momenta, that due to the longer trajectories in the detector and negligible bremsstrahlung emission, is much better in muons than in electrons.

In order to reduce some of these differences, an alternative definition of the dilepton invariant mass squared is chosen as default variable used in the electron channel. This variable is computed by constraining the signal candidates to originate from the primary vertex and to have an invariant mass corresponding to the nominal mass of the B^0 meson. For this reason, this alternative dilepton invariant mass is hereafter referred to as constrained q^2 or q^2_c. The use of this variable is expected to have two main effects: mitigate the differences with respect to the muon mode (resolution and migration) and to reduce the size of the leakage of $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays into the signal region. The last one allows the central-q^2 region, usually defined as $[1.1,6]\,\mathrm{GeV}^2/c^4$ [128], to be extended to $[1.1,7]\,\mathrm{GeV}^2/c^4$, increasing the statistics of the rare electrons sample of roughly 20%. Notice that this variable is not used in the muon case as the difference between q^2 and q^2_c in that case is found to be of minimal importance.

Once the choice to enlarge the central-q^2 region for the electron is taken, the remaining dilepton invariant mass regions are chosen accordingly, making sure that the two lepton flavours considered are aligned. In the following are reported the main q^2 (q_c^2) regions used in this analysis:

i. central-$q^2(q_c^2)$ region: $1.1 < q^2(q_c^2) < 7.0\,\mathrm{GeV}^2/c^4$ for muons (electrons),

ii. J/ψ-$q^2(q_c^2)$ region: $7 < q^2(q_c^2) < 11\,\mathrm{GeV}^2/c^4$ for muons (electrons).

To further increase the statistics of the electron mode, signal candidates are selected not only from those events for which an electron in the final state has triggered, but also in those events that were triggered independently of signal. A similar extension is performed in the muon mode for consistency.

The different behaviour of muons and electrons in the detector has an additional drawback: many of the effects linked to the detection, reconstruction and identification of the two modes are uncorrelated between each other, and therefore a separate control of the systematic uncertainties must be ensured to avoid the introduction of spurious LFU-breaking effects. In order to verify that the corrections applied to the simulation can be trusted, a series of cross-checks are performed in the J/ψ-q^2 (J/ψ-q_c^2) region for the $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ ($B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$) channels, hereafter referred to as *resonant* or *control* modes. The first check corresponds to the measurement of the ratio of branching ratios between the normalization modes, expected to be unity in the SM [129]. This measurement requires a direct control of the efficiencies obtained for the electron and muon modes, since none of the possible systematic uncertainties are expected to cancel. A second powerful check is provided by the measurement of the angular observables of $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ and $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ decays, exploiting the amplitude and acceptance parametrization procedure that will then be used for the final fit to the rare mode. The compatibility between the observables measured in the two different lepton generations requires the control of the efficiency across the different variables needed to describe the decay of interest. In both these cases, thanks to the high statistics of the resonant mode, the dominant uncertainty is expected to be systematic. A measurement that is compatible with the expected result is therefore a very stringent test of our understanding of the simulation used for electron and muons.

Once these cross-checks have been passed, an unbinned maximum likelihood fit to the selected $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ candidates can be performed, as a function of the angles $\cos\theta_K$, $\cos\theta_\ell$ and ϕ, q^2, the invariant mass of the $K\pi$ system, $m(K\pi)$, and the invariant mass of the $K\pi\ell\ell$ system, $m(K\pi\ell\ell)$. The likelihood is minimized to obtain the best estimates for the parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, together with their corresponding uncertainty. The last one will then be used, together with the corresponding systematic uncertainties, to estimate the compatibility of the best fit results with the SM hypothesis.

Currently, the measurement presented in this work is still under the collaboration review and the corresponding best fit result is blinded. Additionally, the computation of the systematic uncertainties of the analysis is ongoing and only those contributions that are expected to have a sizable effect on final the measurement are investigated here.

The remainder of the thesis is organized as follows. Chapter 5 presents the data and simulation samples used in the analysis, together with the selections requirements needed to isolate candidates of $B^0 \to K^{*0}\ell^+\ell^-$ decays in data. These criteria result in a sample of approximately 2000 and 500 $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ candidates in central-q^2 and central-q^2_c regions, respectively. Chapter 6 describes the different steps needed to reduce the known discrepancies between data and simulation, fundamental to avoid the introduction of spurious LFU-breaking contributions in the final measurement. Chapter 7 discusses the effect of the selections on the total number of events and their impact on the angular and dilepton distributions by computing the total efficiencies and effective acceptances of the rare and resonant modes. Chapter 8 reports the results of the mass fits performed on $B^0 \to K^{*0}J/\psi\,(\to \ell^+\ell^-)$ candidates in data, whose yields are needed for the calibration of the constraint on the observed number of events of rare mode in the final fit. Chapter 9 tests the quality of the correction chain, fit strategy and background composition performing two LFU test on the control channels $B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$ and $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$. Additional cross-checks involving the $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ are also discussed. In Chapter 10, the signal and background parametrizations, together with the constraint on the observed yield are discussed in detail. Follows a presentation of the blind fit results to the rare mode candidates in data, together with its expected sensitivity to NP and the most important systematics. The conclusions are presented in Chapter 11.

Chapter 5

Selecting $B^0 \to K^{*0}\ell^+\ell^-$ decays

The chapter begins with the introduction of the datasets used in the analysis of $B^0 \to K^{*0}\ell^+\ell^-$ decays in Sec. 5.1. The selections used to isolate the decays of interest and to reject the many possible sources of backgrounds are discussed instead in Sec. 5.2.

5.1 Data and simulation samples

5.1.1 Data samples

The measurement is based on the *p*-*p* collision data collected by the LHCb experiment during the years 2011 and 2012 (Run 1), 2015 and 2016 (Run 2p1), and 2017 and 2018 (Run 2p2). Table 5.1 reports the integrated luminosities and centre-of-mass energies for each of the years considered.

The final states of the decay of interest, *i.e.* $B^0 \to K^{*0}\ell^+\ell^-$ with $\ell = e, \mu$, are reconstructed based on the combination of four charged tracks: two oppositely charged tracks corresponding to either a muon or an electron couple, and two oppositely charged tracks corresponding to a kaon and pion. Since the charge of the kaon unambiguously identifies the b or $\bar{b}$ nature of the decaying B^0, the decay is self-tagging and no external information is needed to distinguish B^0 and $\overline{B}^0$ mesons.

Table 5.1: Summary of the integrated luminosity and centre-of-mass energy corresponding to the data samples used.

Year	$\mathcal{L}$ [fb^{-1}]	$\sqrt{s}$ [TeV]
2011	1.1	7
2012	2.1	8
2015	0.3	13
2016	1.7	13
2017	1.7	13
2018	2.2	13

5.1.2 Simulation samples

The use of simulated samples is fundamental in many aspects of this measurement, from the determination of the efficiencies, mass distributions and angular distributions of the decays of interest, to the study of the contamination of specific backgrounds in our final datasets. These samples are produced starting from p-p collisions generated using PYTHIA [130], where a specific LHCb configuration [131] is used. The decays of unstable particles are subsequently described by EVTGEN [132], in which the emission of final state radiation (FSR)[1] is included using PHOTOS [133]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [134] as described in Ref. [135]. Many aspects of the simulation are afterwards cross-checked in a data-driven way and, whenever some differences are thought to impact the final observables of interest, a corresponding correction is performed. Additional samples for which the reconstruction with GEANT4 has not been performed, referred to as *generator level* samples, are also used in the measurement. They are particularly useful in the parametrization of the acceptance for the decays of $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0}J/\psi\,(\to \ell^+\ell^-)$ or in the correction of specific variables before the reconstruction. In the following, $B^0 \to K^{*0}\ell^+\ell^-$ decays are generated according to the form factor parametrization of Ref. [136], while $B^0 \to K^{*0}J/\psi\,(\to \ell^+\ell^-)$ decays are generated using the helicity amplitudes reported in Ref. [137].

Different selection, or *truth-matching*, criteria are applied to simulated events depending on the purpose:

i. When designing the selection (Sec. 5.2), only reconstructed candidates in which each of the particles generated is correctly assigned to the corresponding tracks, clusters and mass hypothesis are considered. Events that pass this truth-matching condition are referred to as *Signal*.

ii. When performing corrections to simulation (Ch. 6), two other types of reconstructed candidates are kept together with *Signal* candidates:
    - *Quasi-signal* candidates, where the only difference with the *Signal* category is the misidentification of an intermediate resonance;
    - *Low-mass* candidates, corresponding to partially reconstructed decays where the common mother particle is at maximum 100 MeV/c^2 above the mass of the B^0 meson and all reconstructed particles have been identified correctly.

iii. When performing mass fits to simulation (Sec. 8.2), calculation of efficiencies and parametrization of detector effects (Ch. 7), the candidates in which at most one of the reconstructed final state particles is a ghost are kept together with the *Signal*, *Quasi-signal* and *Low-mass* candidates.

The inclusion of *Low-mass* candidates in the truth-matching definition for most applications is due to the fact that these candidates actually peak in the $K\pi\ell^+\ell^-$ invariant mass, since they correspond for the biggest part to signal events where bremsstrahlung photons have been emitted. The inclusion of ghosts in most

---

[1] In the following, with final state radiation we refer to the electromagnetic emitted by final state particles during the decay process.

applications is due to the presence of this background in data, and neglecting this contribution could bias our estimation of mass shapes, efficiencies and acceptances. Reconstructed candidates of background channels are always truth-matched checking particles identities and daughter-mother relationships.

5.2 Selections

Signal candidates of $B^0 \to K^{*0}\ell^+\ell^-$ decays are isolated in data thanks to a chain of selections that, conceptually, is composed of seven main steps:

i. A cut-based selection on the quality of the tracks and vertices used for the reconstruction of the decay.

ii. Cuts on PID variables to reduce misidentified (mis-ID) backgrounds, together with a set of fiducial cuts needed to ensure that data-driven corrections to the PID efficiencies can be computed.

iii. Trigger requirements and corresponding fiducial cuts.

iv. The definition of the ranges in q^2, $m(K\pi\ell\ell)$ and $m(K\pi)$ where the analysis is performed.

v. Cuts developed to reduce the presence of specific exclusive backgrounds.

vi. A cut on a multivariate classifier trained to reduce the contribution from combinatorial backgrounds.

vii. A cut for the removal of multiple candidates per single event.

The following sections describe each of these steps in more detail.

5.2.1 Requirements on the quality of the reconstructed decay

Sets of four reconstructed charged tracks need to pass the following quality requirements:

- Each of the four tracks must be of a good quality and have a high probability to be produced by a real charged particle. This is enforced requiring that the χ^2 per degree of freedom of the track fit, $\chi^2_{track}/\mathrm{ndf}$, and the ghost probability of the track are small.
- Any two tracks in the final state must have an angular distance larger than 0.5 mrad. This ensures that the number of *clones*, defined as tracks that share at least 70% of their total hits, is reduced.
- The four tracks must be compatible with being originated in a single point in space, corresponding to the decay vertex of the B^0 meson, while their compatibility with the PV of the event is small. This is a powerful condition to discriminate B decays, since they are characterized by a flight distance of roughly 1 mm before decaying in the detector. These criteria are enforced requiring that:

  i. the χ^2 per degree of freedom ($\chi^2_{\text{vtx}}/\text{ndf}$) corresponding to a vertex fit to the $K\pi$ system, to the dilepton pair and to all four particles together are small, suggesting that the three vertices considered are of good quality;
  ii. the χ^2 difference between the fits to the PV with and without the inclusion of:
      * each of the four tracks separately ($\chi^2_{\text{IP}}(\text{PV}, \textit{track})$),
      * the dilepton pair ($\chi^2_{\text{FD}}(\text{PV}, \ell\ell)$),
      * all four tracks together ($\chi^2_{\text{FD}}(\text{PV}, B^0)$),

      are expected to be large if the four charged tracks form a detached SV. To a first approximation, these quantities can be understood as the χ^2 distances of the four tracks and the two vertices with respect to the PV vertex.

- The B^0 candidate built from a kinematic fit to the four charged tracks must point back to the PV. This is enforced requiring that $\chi^2_{\text{IP}}(\text{PV}, B^0)$ and that the angle θ_{DIRA}, corresponding to the angle between the reconstructed momentum of the B^0 meson and its direction of flight from the PV to the decay vertex, are small.

- The invariant masses of the $K\pi\ell^+\ell^-$, $K\pi$ and $\ell^+\ell^-$ are restricted within some very loose intervals, out of which no decay of the type $B^0 \to K^{*0}\ell^+\ell^-$ is expected.

- Events with large occupancies are not considered, as the reconstruction performances greatly decreases with large track multiplicities.

A summary of the requirements applied is reported in Table 5.2 for brevity.

5.2.2 Particle identification requirements

The association of a reconstructed track with the particle specie that generated it is fundamental to reduce the contributions of mis-ID backgrounds and multiple candidates per events. The selections reported in the first row of Table 5.3 are designed to increase the purity of real kaons, pions, muons and electrons in the data samples. To avoid relying on the PID response obtained from simulation, the effect of these selections is studied in specific high-statistics calibration samples that contain muons, pions, kaons, protons and electrons identified without any cut on PID variables. The cuts listed under “Fiducial, Calibration” in Table 5.3 correspond to the selections that are needed to align the data of this analysis with the datasets from which the PID response is calibrated. The cuts listed are of two main types: momentum cuts that ensure the same phase space coverage of the calibration samples, and acceptance cuts that require that the final state particles are in the acceptance of the PID detectors needed for their identification and that they have interacted with them. Of the last type are cuts such as:

Table 5.2: Summary of decay quality requirements, as described in the text.

Type	**Requirement**
Multiplicity cut	`nSPDHits` $< 600(450)$ Run 1 (Run 2)
B^0	$\lvert m - m_B^{\mathrm{PDG}}\rvert < 1500\,\mathrm{MeV}/c^2$ $\cos\theta_{\mathrm{DIRA}} > 0.9995$ $\chi^2_{\mathrm{IP}}(\mathrm{PV}) < 25$ $\chi^2/\mathrm{ndf}(K\pi\ell\ell) < 9$ $\chi^2_{\mathrm{FD}}(\mathrm{PV}) > 100$
K^{*0}	$\lvert m - m_{K^{*0}}^{PDG}\rvert < 300\,\mathrm{MeV}/c^2$ origin vertex $\chi^2/\mathrm{ndf} < 25$
$\ell\ell$	$m < 5500\,\mathrm{MeV}/c^2$ $\chi^2/\mathrm{ndf}(\ell\ell) < 9$ $\chi^2_{\mathrm{FD}}(\mathrm{PV}) > 16$
K, π, μ, e	$\chi^2_{track}/ndf < 3$ $\chi^2_{\mathrm{IP}}(\mathrm{PV}) > 9$ `GhostProb` < 0.4
Clones	$\theta(\ell_{1,2}, h) > 0.5\,\mathrm{mrad}$ $\theta(\ell_1, \ell_2) > 0.5\,\mathrm{mrad}$

- `hasRICH==1`, that requires that the RICH has registered some photons emitted by the track considered;
- `InAccMuon==1`, that requires that the track considered is within the acceptance of the muon stations;
- `IsMuon==1`, a loose binary selection that is exclusively based on the number of muon stations in which a hit is found within a certain spatial window around the track extrapolation;
- `InAccEcal==1` an `hasCalo==1`, that require respectively that the track is within the acceptance of the ECAL and that an energy cluster is associated with it;
- `not(`$\lvert$`xProjection`$^{L0CaloTool}_{\mathrm{ECAL}}\rvert < 363.6\,\mathrm{mm}$ `and` $\lvert$`yProjection`$^{\mathrm{L0CaloTool}}_{\mathrm{ECAL}}\rvert < 282.6\,\mathrm{mm}$), where the tracks that intersect the ECAL in its innermost region are not considered since the corresponding cells are not read out during data-taking;
- $(ee)-\mathrm{ECAL}_{\mathrm{Distance}} > 100\,\mathrm{mm}$, where the distance between the two electron tracks extrapolated at the position of the ECAL is required to be larger than 100 mm to reduce potential non-factorization effects both in PID and trigger corrections.

Table 5.3: Summary of the PID selection requirements with their corresponding fiducial cuts.

Type		**Requirement**
PID	K	$\mathrm{DLL}_{K\pi} > 0$ $\texttt{ProbNNk} \cdot (1 - \texttt{ProbNNp}) > 0.05$
	π	$\texttt{ProbNNpi} \cdot (1 - \texttt{ProbNNk}) \cdot (1 - \texttt{ProbNNp}) > 0.1$
	μ	$\texttt{ProbNNmu} > 0.2$
	e	$\mathrm{DLL}_{e\pi} > 2$ $\texttt{ProbNNe} > 0.2$
Fiducial, Calibration	Multiplicity	$\texttt{nSPDHits} < 600(450)$ Run 1 (Run 2)
	all tracks	`hasRICH==1` $\chi^2_{track}/ndf < 3$
	K, π	$p_{\mathrm{T}} > 250\,\mathrm{MeV}/c$ $p > 2000\,\mathrm{MeV}/c$ `InAccMuon==1`

	e	$p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$
		$p > 3000\,\mathrm{MeV}/c$ && $p < 200\,\mathrm{GeV}$
		$\sqrt{p^2_{x,track} + p^2_{y,track}} > 200\,\mathrm{MeV}/c$
		`hasCalo==1`
		`InAccEcal==1`
		$(ee) - \mathrm{ECAL}_{\mathrm{Distance}} > 100\,\mathrm{mm}$
		$\texttt{not}(\lvert\texttt{xProjection}^{L0CaloTool}_{\mathrm{ECAL}}\rvert < 363.6\,\mathrm{mm}$
		$\texttt{and}\ \lvert\texttt{yProjection}^{L0CaloTool}_{\mathrm{ECAL}}\rvert < 282.6\,\mathrm{mm})$

	ee	$p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$

	μ	$p_{\mathrm{T}} > 800\,\mathrm{MeV}/c$
		$p > 3000\,\mathrm{MeV}/c$ && $p < 200\,\mathrm{GeV}$
		`InAccMuon==1`
		`IsMuon==1`

	K^{*0}	$p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$

5.2.3 Trigger requirements

As already discussed in Sec. 3.2.3, the trigger is the first filtering step in the collection of events produced at p-p collisions in LHCb. Only the events that pass the selections applied for L0, HLT1 and HLT2 trigger are saved for successive offline analysis. In the following, a description of the selections applied at each of these subsequent steps is given, together with the additional fiducial cuts needed to simplify the alignment with simulation.

5.2.3.1 Basic requirements

After the reconstruction of the event, it is possible to distinguish between two main classes of triggered events: events in which the reconstructed candidate caused the trigger to fire, referred to as Trigger-On-Signal (TOS), and events for which the triggering was caused by the underlying decay of the other b-hadron produced in the p-p collision and referred to as Trigger-Indipendent-of-Signal (TIS). Only events that pass the following L0 selections, hereafter referred to as *trigger categories*, are considered:

- **L0I**: `L0Hadron_TIS` (B) `or` `L0Muon_TIS` (B) `or` `L0Electron_TIS` (B), available in electrons and muons in the final state;
- **L0M!**: (`L0Muon_TOS` (μ_1) `or` `L0Muon_TOS` (μ_2)) `and not` L0I (B), available only for samples with muons in the final state;
- **L0E!**: (`L0Electron_TOS` (e_1) `or` `L0Electron_TOS` (e_2)) `and not` L0I (B), available only for samples with electrons in the final state;

where the definition of the `L0Hadron`, `L0Electron` and `L0Muon` trigger decisions can be found in Sec. 3.2.3. The three trigger categories defined are mutually exclusive. The main trigger category of the analysis is L0I, in which the differences between muons and electrons are expected to be minimized. The remaining categories are defined by excluding the main category L0I, as it is denoted by the negation sign “!”. Two additional trigger categories are considered for some specific applications, *e.g.* corrections to signal simulation:

- **L0M**: (`L0Muon_TOS` (μ_1) `or` `L0Muon_TOS` (μ_2)), available only for samples with muons in the final state,
- **L0E**: (`L0Electron_TOS` (e_1) `or` `L0Electron_TOS` (e_2)), available only for samples with electrons in the final state,

which correspond to the trigger categories L0M! and L0E!, without the removal of the events in L0I. For this reason, these categories are also referred to as *inclusive*.

After passing one of the three set of selections for the L0, the events are reconstructed and selected by the HLT, always requiring to have triggered on-signal. Table 5.4 summarizes the selections applied for the different years and leptons involved in the final state. The trigger lines `Hlt1TrackAllL0` and `Hlt1TrackMVA` select events in which at least a single detached high momentum track is found, applying a cut based selection and an MVA, respectively. The lines used in HLT2 select events mainly based on their topology and are designed to select two or more charged particles coming from a displaced vertex, as discussed in more detail in Sec. 3.2.3.

Table 5.4: Summary of the HLT lines used in the measurement.

	Muon modes	**Electron modes**
2011, 2012		
HLT1	`Hlt1TrackAllL0`	
HLT2	`Hlt2Topo[2,3]BodyBBDT`	
	`Hlt2TopoMu[2,3]BodyBBDT`	`Hlt2TopoE[2,3]BodyBBDT`
2015		
HLT1	`Hlt1TrackMVA`	
HLT2	`Hlt2Topo[2,3]Body`	
	`Hlt2TopoMu[2,3]Body`	
2016, 2017, 2018		
HLT1	`Hlt1TrackMVA`	
HLT2	`Hlt2Topo[2,3]Body`	
	`Hlt2TopoMu[2,3]Body`	`Hlt2TopoE[2,3]Body`
	`Hlt2TopoMuMu[2,3]Body`	`Hlt2TopoEE[2,3]Body`

5.2.3.2 Alignment of data and simulation

The trigger configurations for L0 and HLT can change with time. For this reason, to each trigger configuration used in data is assigned a unique *Trigger Configuration Key* (`TCK`). The precise evolution of the trigger configuration in data is therefore always available. When producing the simulation however, it is common to reconstruct the events using a single TCK per year and magnet polarity, usually the one that has been used more frequently in that period. In the following, three main approaches will be pursued to limit the impact of potential misalignments due to this approximation:

- In the case of simulation requirements that correspond to tighter cuts than those used in data, two possible approaches are possible: either selecting in data those `TCKs` that perfectly match or are tighter than the simulation, or apply an additional cut on both data and simulation that is comparable to the tightest cut used in simulation. The first approach is used in the determination of the *prior* trigger corrections discussed in Sec. 6.3, while the second one is used to improve the agreement between data and simulation for the L0E! trigger category.
- In the case of simulation requirements that are looser than those used in data, it is possible to improve the agreement between the two by appropriately tightening the selections in simulation. This is the approach

used to improve the agreement of the HLT1 responses for the years 2012 and 2016.

A comparison between the variation of the `TCK` configurations for the L0 trigger in data and simulation has been done for each year and magnet polarity. Table 5.5 shows an example of such comparison for the year 2012. As it can be seen, while less than 5% of the data was collected with a `L0Muon_TOS` trigger threshold that differs from the simulated one, a large variation is observed for `L0Electron_TOS`. Data and simulation can however be realigned, at least threshold-wise, if the choice of a common and tighter selection is chosen for both. It is important to highlight that the transverse energy as registered by the calorimeter for a given final state particle is not accessible, and a direct cut on this quantity is therefore not possible. It is however possible to use an offline proxy of this variable, corresponding to the projected transverse energy E_T^{L0} at the position in which the particle track intersects the calorimeter. The definition of the L0E! category can then be modified to accommodate this additional fiducial cut, by simply substituting:

$$\texttt{L0Electron\_TOS}(e_i) \;\rightarrow\; \texttt{L0Electron\_TOS}(e_i) \;\texttt{and}\; E_T^{\mathrm{L0}}(e_i) > E_{threshold}^{\mathrm{L0}}. \tag{5.1}$$

A redefinition of the thresholds for `L0Electron_TIS` (B) should also technically be done, but is ignored due to its expected limited impact. A summary of the thresholds used for each year is shown in Table 5.6. As can be seen, a threshold of $3000\,\mathrm{MeV}/c^2$ is chosen for 2012, higher than the highest used in data and simulation for that year.

Table 5.5: `L0-TCK` thresholds for 2012 data taking period.

2012 (MagUp)				2012 (MagDown)			
`L0-TCK` (hex)	% data	$p_T^{L0ADC}(\mu)[\mathrm{MeV}/c]$	$E_T^{L0ADC}(e)[\mathrm{MeV}/c^2]$	`L0-TCK` (hex)	% data	$p_T^{L0}(\mu)[\mathrm{MeV}/c]$	$E_T^{L0}(e)[\mathrm{MeV}/c^2]$
`MC`	100%	>1760	>2960	`MC`	100%	>1760	>2960
`0x990042`	37.53	>1760	>2720	`0x990042`	19.81	>1760	>2720
`0xa30044`	16.77	>1760	>2960	`0xa30044`	14.22	>1760	>2960
`0x94003d`	15.45	>1760	>2720	`0x97003d`	13.71	>1760	>2720
`0x97003d`	11.34	>1760	>2720	`0x990044`	11.1	>1760	>2960
`0xa10045`	7.36	>1760	>2960	`0xac0046`	9.58	>1760	>2860
`0x990044`	3.03	>1760	>2960	`0x8c0040`	9.56	>1480	>2500
`0xa10044`	2.74	>1760	>2960	`0x94003d`	9.29	>1760	>2720
`0x9f0045`	2.73	>1760	>2960	`0xa90046`	5.97	>1760	>2860
`0xa30046`	1.83	>1760	>2860	`0xab0046`	4.85	>1760	>2860
`0xac0046`	0.79	>1760	>2860	`0xa30046`	1.04	>1760	>2860
`0x990043`	0.27	>1760	>2720	`0x860040`	0.67	>1480	>2500
`0xa20044`	0.15	>1760	>2960	`0x95003d`	0.09	>1760	>2720
`0x95003d`	0.01	>1760	>2720	`0x8e0040`	0.07	>1480	>2500
`0xad0046`	<0.01	>1760	>2860	`0x7f0040`	0.02	>1480	>2500
`0x9a0042`	<0.01	>1760	>2720	`0x7e003a`	0.01	>1480	>2500
				`0xad0046`	0.0	>1760	>2860

Similarly to the L0 trigger, a comparison between the HLT1 trigger lines in data and simulation is performed for each year and magnet polarity. The main differences are observed in the years 2012 and 2016, where the `TCK` in simulation is always looser or equal to the one used in data. To reduce the differences between data and simulation, a tightening of the selections applied in simulation to match the ones applied in data is performed. In doing so, the different fractions of data associated with each cut are reproduced in simulation.

Table 5.6: L0E! thresholds used in the offline alignment of the `L0Electron_TOS`(e_i) trigger decision in data and simulation.

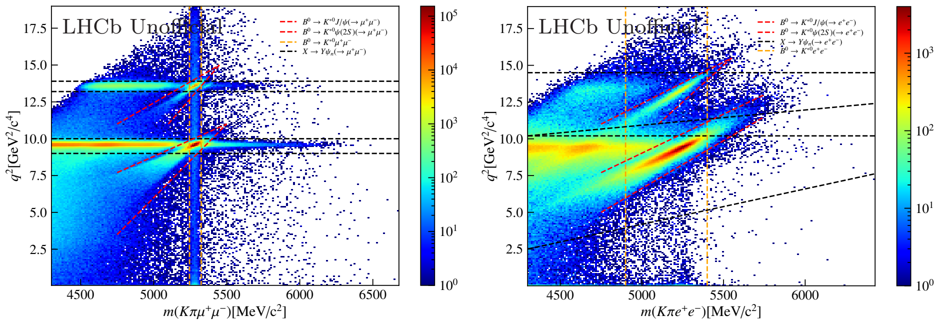
Year	$E^{L0}_{threshold}[\text{MeV}/c^2]$
2011	>2500.
2012	>3000.
2015	>3000.
2016	>2700.
2017	>2700.
2018	>2400.

5.2.4 Fit domains in q^2, $m(K\pi\ell\ell)$, $m(K\pi)$

5.2.4.1 Fit domains in q^2

The decays $B^0 \to K^{*0}\ell^+\ell^-$, $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ originate from the same $b \to s\,\ell^+\ell^-$ transition, but correspond to two different regimes of the $c\bar{c}$ loop contribution described in Sec. 2.2: at the dilepton invariant mass corresponding to mass of the J/ψ and $\psi(2S)$ resonances, $M_{J/\psi}$ and $M_{\psi(2S)}$, the energy is exact to produce these charmonium states on-shell. Since they originate from the same amplitudes, a strict separation between these different regimes is not possible and an interference between these modes permeates the whole phase space. It is however possible to conventionally divide the dilepton invariant mass spectrum in regions, where one of the three contributions is expected to dominate with respect to the others.

Figure 5.1 shows, on the left and on the right, the distributions of reconstructed $K\pi\ell^+\ell^-$ data candidates after the full selection for muons and electrons, respectively. Three main features can be noticed: in red, two diagonal bands that peak around the mass of the B^0 meson in $m(K\pi\ell\ell)$ and around the J/ψ and $\psi(2S)$ resonances in q^2; in black, two horizontal bands at the q^2 values of $M_{J/\psi}$ and $M_{\psi(2S)}$, which seem to decrease monotonically from left to right; in yellow, a vertical band at the mass of B^0 meson in $m(K\pi\ell\ell)$. The events that remain have almost no structure, and seem to decrease with higher $m(K\pi ee)$ and q^2 values. The diagonal bands correspond to the control mode decays of $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$. The diagonal structure comes from the correlation between $m(K\pi\ell\ell)$ and q^2, since an under- or over-estimation of the momenta of the two leptons will bias coherently the two reconstructed masses. This feature, that is rather mild for the muon case, becomes greatly accentuated for the electron mode due to higher energy loss caused by FSR, bremsstrahlung and the stochastic effect of the bremsstrahlung recovery algorithm. The horizontal bands correspond to a mixture of backgrounds that contain real $J/\psi, \psi(2S) \to \ell^+\ell^-$ decays, where the hadronic part is either partially reconstructed or combinatorial, showing therefore no strong correlation with the reconstructed $m(K\pi\ell\ell)$ mass. These two bands meet at the masses of the B^0 meson and J/ψ ($\psi(2S)$) resonance. The vertical bands correspond to the rare $B^0 \to K^{*0}\ell^+\ell^-$ decays. Even though is



(a) (b)

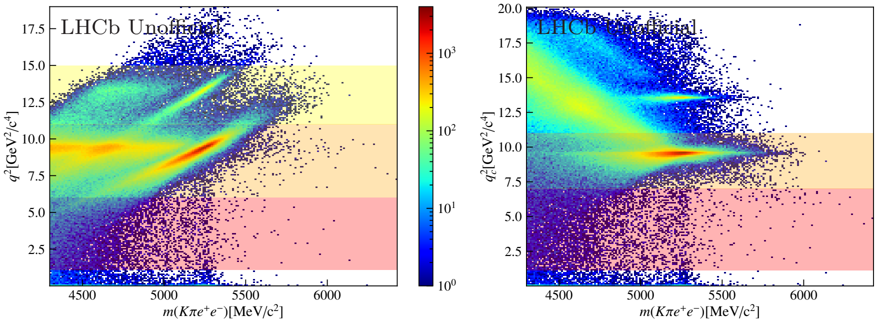
Figure 5.1: Figure 5.1(a) and (b) show the distribution of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ candidates after the full selection as a function of the invariant masses $m(K\pi\ell\ell)$ and q^2. In red are highlighted the contributions corresponding to the resonant modes $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$, in black backgrounds containing a real J/ψ or $\psi(2S)$ and in yellow the candidates for the rare $B^0 \to K^{*0}\ell^+\ell^-$ decay.

not as clear as for the $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decay modes, a correlation between $m(K\pi\ell\ell)$ and q^2 is present also in this case. The remaining events do not have a specific structure and are constituted by the different backgrounds discussed in Sec. 5.2.5 and 5.2.6.

As Figure 5.1 clearly shows, due to the correlation between $m(K\pi\ell\ell)$ and q^2, the regions where $B^0 \to K^{*0}\ell^+\ell^-$, $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays each respectively dominate must be chosen with a two-dimensional cut in these variables. Since the final aim of the analysis is to measure the properties of the $B^0 \to K^{*0}\ell^+\ell^-$ decays and the sensitivity is limited by the mode with the lowest yield, *i.e.* the electron mode, it is fundamental to choose the $B^0 \to K^{*0}e^+e^-$ decay region to have as many candidates as possible. So far, the region[2] preferred in the literature has been the one corresponding to $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ [52], highlighted in red in Figure 5.2(a). As an alternative, it is also possible to consider a selection in q^2_c rather than in q^2; this variable has already been used in Ref. [138] and is chosen in the following as default dilepton mass squared variable for the electron modes. The use of q^2_c causes almost a rotation in the distribution shown in Figure 5.2(b), for which the control mode events correspond to the horizontal bands while the J/ψ and $\psi(2S)$ charmonium backgrounds are now diagonal. This simplifies the cut-based selection in q^2 by allowing the inclusion of more events of signal with an extension of the q^2 range: central-q^2_c is defined as $1.1 < q^2_c < 7.0\,\mathrm{GeV}^2/c^4$ and, as Figure 5.2(b) shows in red, it is a region dominated by $B^0 \to K^{*0}e^+e^-$ decays.

---

[2] We are referring here only to the regions above the ϕ resonance.



(a)

(b)

Figure 5.2: (*a*) Representation of the q^2 regions defined in Ref. [52], corresponding to central-q^2 ($R_{K^{*0}}$), J/ψ-q^2 ($R_{K^{*0}}$) and $\psi(2S)$-q^2 ($R_{K^{*0}}$) for the red, orange and yellow band, respectively. (*b*) Representation of the nominal q^2 regions used in this analysis: central-q_c^2 and J/ψ-q_c^2 in red and orange, respectively.

This allows an increase of 20% in the statistics available for the rare electron channel. Moreover, the use of the q_c^2 variable reduces the migration of events whilst imposing a good rejection power on background sources for which the assumption of invariant mass equal to the nominal mass of the B^0 meson does not hold. Such improvement in the q^2 resolution is a crucial aspect in the unbinned approach used in this analysis. Once this decision is made, all the other q^2 regions are readapted to be mutually exclusive, while the muons are aligned accordingly. Notice that, since the difference between q^2 and q_c^2 is small in muons with respect to electrons, q^2 is used in their selection. In the following, the q^2 regions used hereafter (and already defined in Sec. 4.1) are reported for completeness:

- **Electron mode regions**
    i. central-q_c^2: $1.1 < q_c^2 < 7.0\,\mathrm{GeV}^2/c^4$,
    ii. J/ψ-q_c^2: $7 < q_c^2 < 11\,\mathrm{GeV}^2/c^4$,
- **Muon mode regions**
    i. central-q^2: $1.1 < q^2 < 7.0\,\mathrm{GeV}^2/c^4$,
    ii. J/ψ-q^2: $8 < q^2 < 11\,\mathrm{GeV}^2/c^4$,

Additionally, an alternative set of q^2 regions that can partially overlap with the previous, is used for specific cross-checks:

- **Electron mode regions**
    i. central-q^2 ($R_{K^{*0}}$): $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$,

   ii. J/ψ-q^2 ($R_{K^{*0}}$): $6 < q^2 < 11\,\mathrm{GeV}^2/c^4$,
   iii. $\psi(2S)$-q^2 ($R_{K^{*0}}$): $11 < q^2 < 15\,\mathrm{GeV}^2/c^4$,

- **Muon mode regions**
   i. J/ψ-q^2 ($R_{K^{*0}}$) : $|m(\ell^+\ell^-) - M^{PDG}_{J/\psi}| < 100\,\mathrm{MeV}/c^2$,
   ii. $\psi(2S)$-q^2 ($R_{K^{*0}}$) : $|m(\ell^+\ell^-) - M^{PDG}_{\psi(2S)}| < 100\,\mathrm{MeV}/c^2$,
   iii. interRes-q^2: $11 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$,

The extension of the central-q^2 region is not a trivial task. It can only be achieved by using the variable q^2_c and showing that the contributions leaking from the resonant mode into the central bin, called J/ψ-leakage, can be kept under control. In Sec. 5.2.5.1 the performance of different definitions of the central bin in q^2 is presented and shows that, with this definition of central-q^2_c, it is possible to neglect the J/ψ-leakage contribution in the final fit and assign a systematics due its small remaining contribution.

5.2.4.2 Fit domains in $m(K\pi)$ and $m(K\pi\ell\ell)$

Candidates of $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays are selected in $m(K\pi)$ by requiring a mass window of $200\,\mathrm{MeV}/c^2$ centred around the nominal mass of the K^{*0}, corresponding to roughly twice its expected width. This selection has a high efficiency in rare and resonant modes both for electrons and muons, corresponding to $(90 \pm 1)\%$. This range is also large enough to allow the discrimination between the decays where the $K\pi$ system is not produced by the decay of the $K^*(892)^0$ spin-1 resonance, but rather non-resonantly or from higher scalar resonances.Background events from decays that do not originate from a real K^{*0} are rejected with high efficiency by this cut, as discussed in Sec. 5.2.5.

The choice of the $K\pi\ell^+\ell^-$ invariant mass ranges is fundamental to remove, suppress or include contributions from partially reconstructed decays or leakages from resonant modes in the different regions of q^2 considered. Two types of $K\pi\ell^+\ell^-$ invariant masses are considered:

- $m(K\pi\ell\ell)$, corresponding to the invariant mass obtained after refitting the decay chain constraining the final state particles to be originated from the same primary vertex [139];
- $m^{J/\psi}_{DTF}$ and $m^{\psi(2S)}_{DTF}$, where together with the primary vertex constraint, the decay is refitted constraining also the dilepton pair mass to the nominal mass of the J/ψ and $\psi(2S)$, respectively [139].

The former are used in the amplitude and mass fits of $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ candidates needed *e.g.* to impose a constraint on the observed rare mode yield in the final fit, while the latter are used in the mass fits to $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays needed to

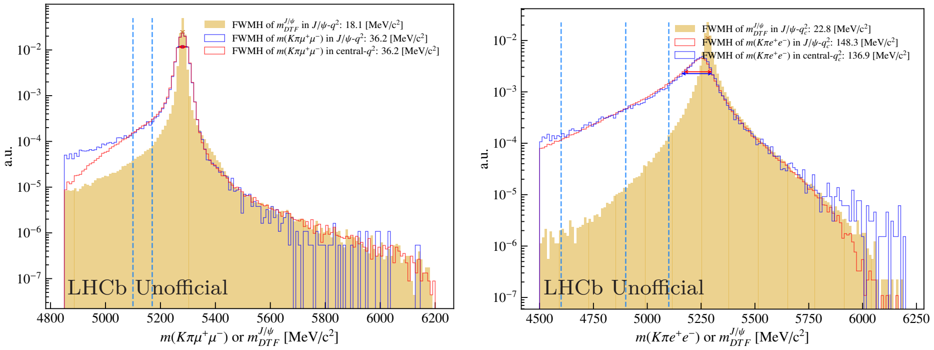
Table 5.7: Summary of the $m(K\pi\ell\ell)$ and $m_{DTF}^{\psi_n}(K\pi\ell\ell)$ ranges, where $\psi_n = J/\psi\,, \psi(2S)$, used for the different amplitude and mass fits discussed in this work.

	$m(K\pi\ell\ell)$	$m_{DTF}^{\psi_n}(K\pi\ell\ell)$
$B^0 \to K^{*0}e^+e^-$	[4900, 5700] MeV/c^2	-
$B^0 \to K^{*0}\mu^+\mu^-$	[5170, 5700] MeV/c^2	-
$B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$	[4600, 5700] MeV/c^2	[4900, 6200] MeV/c^2
$B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$	[5100, 5800] MeV/c^2	[5100, 6200] MeV/c^2
$B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$	-	[5100, 5950] MeV/c^2
$B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-)$	-	[5100, 5750] MeV/c^2

perform the LFU cross-checks of Secs. 9.1.2 and 9.1.3. Table 5.7 summarizes the mass ranges used throughout this work.

In muon fits, the mass resolution is so narrow that 95% of the signal candidates lay within 200 MeV/c^2 of the nominal mass of the B^0 meson, regardless of the use of an additional constraint on the dilepton resonance. In fits to the rare mode, this allows to choose a mass range between 5170 and 5700 MeV/c^2 to reject partially reconstructed contributions that peak at low mass, and to include a sizable fraction of combinatorial background needed to constrain this component under the signal. In fits to the resonant modes, the mass range is enlarged to include a small fraction of partially reconstructed background in the fit, since it improves the constraint of the right-hand tail of this component under the signal and reduces the differences between electron and muon modes.

In electron fits, a noticeable difference can be seen between invariant masses computed with and without the use of an additional constraint on the dilepton resonances. While for $m_{DTF}^{J/\psi}$ the resolution is similar between muon and electrons thanks to the improved resolution that the J/ψ ($\psi(2S)$) mass constraint provides, $m(K\pi ee)$ is characterized by a *full width* at *half maximum* (FWHM) that is roughly six times larger than what is obtained for $m_{DTF}^{J/\psi}$ (see Figure 5.3). Additionally, it is characterized by a long left tail that needs to be included in the fit to ensure high efficiency in the selection of the signal. Between the mass ranges considered, the one that removes the highest fraction of signal is [4900, 5700] MeV/c^2, used in the fit of $B^0 \to K^{*0}e^+e^-$ decay candidates, and corresponds to a cut efficiency that is around 90%. This choice, as it is discussed in more detail in Sec. 5.2.5.1, aims to suppress to a negligible level the J/ψ-leakage component from the final amplitude fit. More details on the parametrization of the invariant masses $m(K\pi\ell\ell)$, $m_{DTF}^{J/\psi}$ and $m_{DTF}^{\psi(2S)}$ and their fits are provided in Ch. 8.



(a) (b)

Figure 5.3: (a) Distribution of the $m(K\pi\mu\mu)$ in central-q^2 (red) and J/ψ-q^2 (blue) regions superimposed to the distribution of $m_{DTF}^{J/\psi}$ in the central-q^2 (yellow) region. (b) Distribution of the $m(K\pi ee)$ in central-q_c^2 (red) and J/ψ-q_c^2 (blue) regions superimposed to the distribution of $m_{DTF}^{J/\psi}$ in the central-q_c^2 (yellow) region. Dashed in blue are represented the lower limits in the ranges used for the mass and amplitude fits, as defined in Table 5.7. The FWHM of each distribution are reported in the legend.

5.2.5 Exclusive background rejection

A good understanding of the different processes that might fully or partially overlap with the signal and control modes is fundamental to have a reliable estimation of their properties. It is possible to divide these different contributions, referred to as *backgrounds*, in three main categories:

- Decays of b-hadrons that share with the signal the same final state particles. These decays can be further subdivided in two: *partially reconstructed backgrounds*, where the four tracks of interest are just a subsample of the final state particles produced in the decay *e.g.* $B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$, and decays with the exact same final state *e.g.* $B_s^0 \to \overline{K}^{*0}\ell^+\ell^-$.

- Decays of b-hadrons in four charged final state particles where, due to the misidentification of one or more of the tracks, the decay can be reconstructed as the signal. Example of these decays are $\Lambda_b^0 \to pK\ell^+\ell^-$, where a proton is misidentified as a pion, and $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays, where the kaon and pion are reconstructed as pion and kaon, respectively. The last type of joint mis-ID is also referred to as *swap*.

- Backgrounds of total or partial combinatorial nature. These backgrounds are built by joining four tracks that are either fully or partially unrelated. Due to the sheer amount of particles produced in each event, it is possible that a (partially) random combination of tracks might pass all selections. An example of such backgrounds are the combinatorial background discussed in Sec 5.2.6 or the events in the black band highlighted in Figure 5.1, that contain a real J/ψ or $\psi(2S)$ resonance.

These categories are not mutually exclusive and, in general, backgrounds deriving from a mixed contribution of the three are expected. In the following, it is provided a discussion of the main sources of background considered for each of the regions of interest in this analysis, together with the additional selections that have been developed to reduce their contributions.

5.2.5.1 Leakage from J/ψ and $\psi(2S)$ resonances

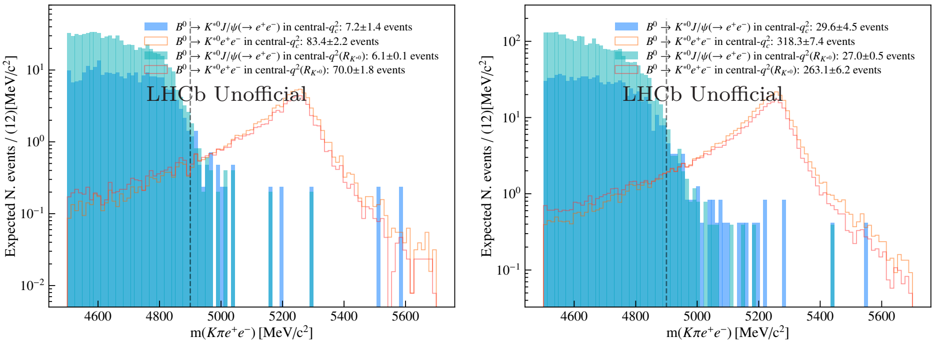
The regions of q^2, $m(K\pi\ell\ell)$ and $m(K\pi)$ defined in Sec 5.2.4 have been chosen to simplify the isolation of the contributions of $B^0 \to K^{*0}\ell^+\ell^-$, $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays in data. Each of these three modes can have a non-negligible component outside its q^2 bin of reference and needs to be included in the analysis as a source of background. Backgrounds of this type are referred to as *leakage*s. As shown in Figure 5.2, the size of these background components heavily depends on the q^2 and $m(K\pi\ell\ell)$ regions used in the analysis. Once again, can be instructive to compare the different q^2 strategies discussed in Sec. 5.2.4 to further justify our nominal choice:

- When considering the q^2 regions central-q^2, J/ψ-q^2 and $\psi(2S)$-q^2 as in Figure 5.2(a), it is clear that the J/ψ-leakage component leaks both in the central-q^2 and $\psi(2S)$-q^2 regions due to its radiative tail.
- When considering the q^2 regions central-q^2_c and J/ψ-q^2_c as in Figure 5.2(b), it can be seen that a large fraction of the radiative tail is now contained inside the corresponding q^2 bin and the leakage reduced for the same limits of q^2.

This can be clearly understood when considering the sizes of the expected J/ψ-leakage and $B^0 \to K^{*0}e^+e^-$ components obtained from simulation in the central-q^2 and central-q^2_c regions, as shown in Figure 5.4. While the size of the J/ψ-leakage component above $4900\,\mathrm{MeV}/c^2$ between the two approaches is compatible, the amount of signal in the same range is increased of roughly 19% when the q^2 range is extended. The only additional complication is the introduction of combinatorial background containing true J/ψ in the upper mass side band of the central-q^2_c region. This contribution is however heavily suppressed by the combinatorial multivariate classifier (see Sec. 5.2.6).

5.2.5.2 Other backgrounds with the same final state

In addition to the possible leakages from different q^2 bins, two other exclusive backgrounds with same final state particles as the signal and control modes need to be considered:



(a) (b)

Figure 5.4: Comparison between the distributions of the $B^0 \to K^{*0}e^+e^-$ decay (solid) and J/ψ-leakage (filled) normalized to expected number of events in Run 1 (a) and Run 2 (b), for the two regions of central-q^2 ($R_{K^{*0}}$) and central-q^2_c. While the J/ψ-leakage contribution is comparable between the two above 4900 MeV/c^2, indicated by the black dashed line, an increase of roughly 19% is observed for the expected yield of $B^0 \to K^{*0}e^+e^-$ decays.

- $B^0_s \to \overline{K}^{*0}\ell^+\ell^-$ decays and the corresponding charmonium channels: These backgrounds cannot be vetoed due to the small mass difference between B^0 and B^0_s mesons. They are therefore either parametrized, as in $B^0_s \to \overline{K}^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0_s \to \overline{K}^{*0}\psi(2S)(\to \ell^+\ell^-)$ mass fits, or neglected, as in $B^0 \to K^{*0}\ell^+\ell^-$ amplitude fits. Their contributions are expected to be around few percent with respect to the B^0 modes due to a CKM and f_s/f_d suppression.

- $B^0 \to [K^+\pi^-]_{J=0}\,\ell^+\ell^-$ decays and corresponding charmomium channels: Decays of this type originate from the same $b \to s\,\ell^+\ell^-$ transition of the signal, but differ for the configuration of the hadronic system in the final state: the K and π are produced either directly in a non-resonant configuration or via the decay of a scalar resonance, such as the $K^*(1430)^0$, and are in both cases characterized by a total angular momentum of $J = 0$. These decays cannot be vetoed as they are indistinguishable from the signal, with which they can interfere. To differentiate between these two decays we will refer to $B^0 \to [K^+\pi^-]_{J=0}\,\ell^+\ell^-$ as *S-wave* and to $B^0 \to K^{*0}\ell^+\ell^-$ as *P-wave*. These two processes can only be disentangled thanks to their different angular and $m(K\pi)$ distribution, and they must be either included as an additional component in the fit or neglected. Being non-resonant around the mass of the K^{*0}, they are suppressed by the choice of the $m(K\pi)$ mass discussed in Sec. 5.2.4.

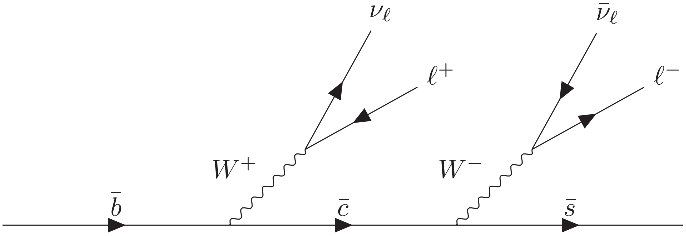


Figure 5.5: Feynman diagram of the generic double-semileptonic transition $H_b \to H_c\ell^+\nu_\ell$, where the H_c further decays semileptonically as $H_c \to H_s\ell^-\bar{\nu}_\ell$. The corresponding spectator quark(s) are not represented.

5.2.5.3 Partially reconstructed decays

Partially reconstructed backgrounds originate from decays that have five or more particles in their final state, four of which can be combined and reconstructed as the decays of interest. Since part of their energy is carried away by the particles that have not been reconstructed, their $m(K\pi\ell\ell)$ mass populates a lower mass sideband with respect to the B^0 meson mass. In the following, the two most important contributions to these class of decays is considered.

$B \to K^+\pi^-\pi\,\ell^+\ell^-$ decays and corresponding charmonium channels:

The decays $B \to K^+\pi^-\pi\,\ell^+\ell^-$, where the leptons have been produced directly or from a charmonium resonance, have the same final state particles of the signal and can be reconstructed as such when the π is lost. Due to the missing energy of the pion, they are characterized by a broad reconstructed mass that peaks roughly $200\,\text{MeV}/c^2$ below the nominal mass of the B^0 meson. Whenever the $m(K\pi\ell\ell)$ mass resolution is small compared to this gap, these decays can be vetoed by a simple choice of the mass window: this is the strategy used in the analysis of muon modes and of $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ and $B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$ decays, if the resolution on $m(K\pi ee)$ mass is improved by constraining the dilepton system to the mass of the resonance. This approach cannot be used in the decays of $B^0 \to K^{*0}e^+e^-$, where no additional constraint is possible and the overlap with the partially reconstructed background does not allow for an efficient veto. The same is true for $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ decays if no additional constraint is used. In both situations, these backgrounds need to be parametrized (see Secs. 8.3.1 and 10.2.3). When performing fits to $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ candidates in data, an additional selection is introduced: $\left|m(B)^{\psi(2S)}_{DTF} - m(B)^{\text{PDG}}\right| > 200\,\text{MeV}/c^2$. The aim of this selection is to reduce the contribution of $B \to \psi(\to J/\psi X)Y$ decays, which are known to be poorly simulated in the inclusive samples of B meson decays available.

Double-semileptonic cascade decays

Another class of partially reconstructed decays are the so-called double-semileptonic cascade (DSL) decays, represented in Figure 5.5 by the corresponding tree-level Feynman diagram. They are characterized by the semileptonic decay of a beauty hadron in a charm hadron, that subsequently decays semileptonically into a strange hadron. When the two semileptonic decays produce leptons of the same generation, these decays have the same final state particles of the decays of interest, with the addition of two neutrinos, that travel undetected in our spectrometer. Due to the fraction of the total energy that they carry, the reconstructed $m(K\pi\ell\ell)$ peaks at lower masses with respect to the fully reconstructed signal.

These are naturally suppressed by the requirements applied in Sec 5.2.1, since the two subsequent semileptonic decays happen in different points in space rather than in a single decay vertex. However, being characterized by branching ratios roughly three orders of magnitude larger than the signal, these can still constitute a sizable source of background for the rare mode even if greatly suppressed.

The most important contribution is expected to come from $B^0 \to D^-\ell^+\nu_\ell$ decays, with the D^- that decays as $D^- \to (K^{*0} \to K^+\pi^-)\ell^-\bar{\nu_\ell}$. Even though the particle identification or $m(K\pi)$ selections have no effect on these decays, it is possible to partially veto them by requiring that the invariant mass of the $K\pi\ell^-$ system is larger than the mass of the D^- meson, the effect of which is shown in Figure 5.6. This effectively corresponds to removing all configurations in which ℓ^- is much less energetic than ℓ^+. This heavily distorts the angular distributions of the signal in $\cos\theta_\ell$, since it can be approximated as the energy asymmetry of the two electrons:

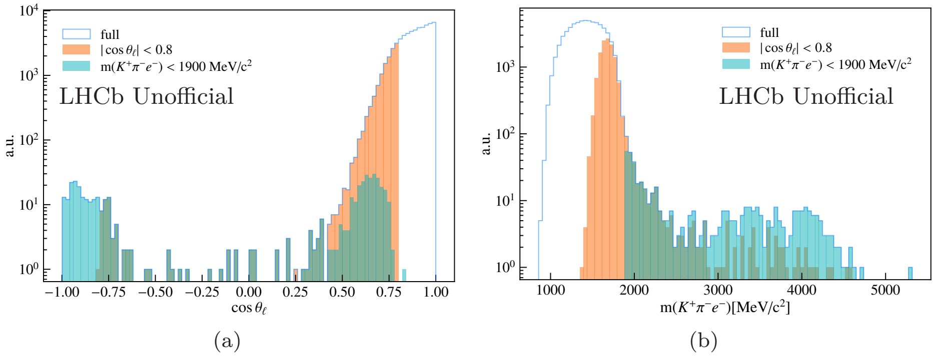
$$\cos\theta_\ell \approx \frac{E_{\ell^+} - E_{\ell^-}}{E_{\ell^+} + E_{\ell^-}} \,. \tag{5.2}$$

This distortion can be avoided by performing a cut above 0.8 or lower in the absolute value of $\cos\theta_\ell$, at the cost of a less performant veto, as shown in Figure 5.6.

An alternative strategy has been designed for this thesis. In fits to the electron rare mode, this contribution is not vetoed, but rather included as an additional background component. In fits to the muon mode, this component is ignored, since its contribution is expected to be negligible in the mass range considered. More details on the parametrization of this background component can be found in Sec. 10.2.2.

5.2.5.4 Backgrounds from misidentification

Since particle identification is a stochastic process, each of the PID cuts discussed in Sec 5.2.2 has associated a corresponding misidentification probability that depends on the particles considered. This probability can go up to a few percent depending on the particle characteristics. Decays that have final states similar to the decay of interest, except for one or two particles that can be misidentified to



(a) (b)

Figure 5.6: Impact of two possible vetoes for the simulated decays $B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)e^-\bar{\nu_e})e^+\nu_e$ as a function of $\cos\theta_\ell$ (a) and $\mathrm{m}(K\pi e^-)$ (b).

match the required particle identification, can therefore be wrongly reconstructed as signal candidates. In the following, the most relevant sources of mis-ID backgrounds are considered. Depending on the size of the expected background, a maximum of two misidentified particles are considered. An additional reduction factor can be achieved if the misidentified decay happens through intermediate resonances, since it is possible to isolate them by an appropriate swap of the mass hypothesis and a tightening of the PID selection.

$K \to \pi$ misidentification

The most relevant sources of $K \to \pi$ misidentification come from $B_s^0 \to \phi(\to K^+K^-)\ell^+\ell^-$ and the corresponding charmonium channels. They are particularly dangerous because they are expected to peak right below the signal, potentially biasing the observed yield. Their contribution is reduced by vetoing all events for which the invariant mass $m(KK_{\to\pi})$ is equal or lower than the ϕ mass. To ensure a higher signal efficiency, a tighter PID selection is used in this region rather than veto. This contribution is included in the fit to the resonant modes, while a systematic is assigned based on the expected number of events for the rare mode.

$p \to h$ misidentification

Decays that produce a proton in the final state must originate from a b-baryon. The dominant source of such decays is constituted by $\Lambda_b^0 \to pK\ell^+\ell^-$ decays and the corresponding charmonium channels, since the decays involving a pion are Cabibbo suppressed. These backgrounds can be reconstructed as signal when a proton is misidentified as a pion ($p_{\to\pi}K$, single mis-ID) or when a proton is misidentified as a kaon and the kaon as a pion ($p_{\to K}K_{\to\pi}$, double mis-ID). In both configurations, the reconstructed decays peak broadly under the signal in $m(K\pi\ell\ell)$, potentially biasing the observed yield. These contributions are reduced by the choice of the $m(K\pi)$ window and PID requirements, but no efficient veto can be constructed due to the lack of an intermediate resonance in the hadronic system. This contribution is included in the fit to the resonant modes, while a systematic is assigned based on the expected number of events for the rare mode.

$K \leftrightarrow \pi$ and $h \leftrightarrow \ell$ swaps

The decays of $B^0 \to K^{*0}\ell^+\ell^-$ and the corresponding charmonium channels can be a source of background themselves, when two particle hypotheses in the final state are swapped. Two main configurations are possible: $K \leftrightarrow \pi$ swaps and $h \leftrightarrow \ell$ swaps. In both cases these contributions can be particularly dangerous since they still peak below the nominal B^0 meson mass and are characterized by a distortion of the angular distributions. While for the $K \leftrightarrow \pi$ swaps no other efficient veto can be designed, two different strategies are followed in $h \leftrightarrow \ell$ swaps depending on the lepton generation considered. For the muon modes, an additional veto can be applied by reconstructing the invariant masses $m(K_{\to\mu}\mu)$ and $m(\mu\pi_{\to\mu})$ and veto the regions corresponding to the J/ψ and $\psi(2S)$ resonances. The same veto is not as efficient in the electron modes due to the worse invariant mass resolution, and cut on the invariant masses constrained to either the J/ψ or $\psi(2S)$ resonance is preferred. It is important to notice that $h \leftrightarrow \ell$ swaps are particularly dangerous for the rare mode, since these decays do not peak at the nominal masses of the J/ψ and $\psi(2S)$ resonances.

$h \to \ell$ misidentification

Semileptonic cascade decays, where only one between the B and the D meson decays semileptonically, constitute a source of background if one of the final state hadrons is misidentified as a lepton. As already seen for DSL, these decays are characterized by large branching ratios with respect to the rare signal modes and can contribute significantly to the signal region even after the full chain of selections, including the PID requirements. Since not all the particles in the final state are reconstructed (at least one neutrino from the semileptonic decay), these decays populate a region of $m(K\pi\ell\ell)$ that is below the nominal B^0 meson mass.

These backgrounds can be divided in two main categories:

- $H_b \to H_c(\to K\pi X)\ell^+\nu_\ell$, where the b-hadron H_b decays semileptonically and the c-hadron H_c hadronically. These include decays of $B^0 \to (D^{*-} \to (\overline{D}^0 \to K^+\pi^-)\pi^-)\ell^+\nu_\ell$, $B^0 \to (\overline{D}^0 \to K^+\pi^-)\pi^-\ell^+\nu_\ell$ and $B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)\pi^-)\ell^+\nu_\ell$.
- $H_b \to H_c(\to K\pi\ell^-\nu_\ell)\pi$, where the b-hadron decays hadronically and c-hadron semileptonically. These include decays such as $B^+ \to (\overline{D}^0 \to K^{*0}(\to K^+\pi^-)\ell^+\nu_\ell)\pi^+$ and $B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)\ell^-\nu_\ell)\pi^+$.

Decays of the first type can be effectively vetoed by reconstructing the D^0 and D^- invariant masses as $m(K^+\pi^-{}_{\to\ell})$ and $m(K^+\pi^-\pi^-{}_{\to\ell})$, and applying a strong PID requirement within a $30\,\mathrm{MeV}/c^2$ window from their nominal masses. When swapping the hypothesis from an electron to a different particle, the use of the corresponding track momentum is used to avoid biases from bremsstrahlung corrections. Decays of the second type cannot be vetoed in this way, but are in general suppressed by the same requirements discussed in the double-semileptonic cascade decays.

Purely hadronic decays are strongly suppressed due to the necessity of a double misidentification of hadrons into leptons. However, due to the broad spectrum of selection efficiencies and branching ratios that these background could have, an exhaustive estimation of the expected contribution is not available. Attempts to estimate the impact of these contributions directly from data has been performed in the ongoing measurements of the R_{K^*} branching fraction ratio [128] and angular analysis of $B^0 \to K^{*0}e^+e^-$ decays [140]. In these analyses, contributions from such decays corresponded to the largest single source of systematic uncertainty in the final measurement. More details on the potential impact that these misidentified backgrounds could have in this analysis can be found in Sec. 10.6.

5.2.5.5 Over-reconstructed backgrounds

Over-reconstructed backgrounds are built from decays for which the particles in the final state are a subset of the ones of the decay of interest, while the remaining particles come from the underlying event. The main source of over-reconstructed background comes from $B^+ \to K^+\ell^+\ell^-$ decays, when a slow pion coming from a different interaction points to the decay vertex of the B^+, allowing this decay to be reconstructed as signal. Due to the additional pion considered, the reconstructed $m(K\pi\ell\ell)$ peaks above the B^0 meson mass and the angular distribution still retains much of the original kinematic of the decay. This contribution can be suppressed by a veto on the masses $m(K\ell^+\ell^-)$ and $m(\pi_{\to K}\ell^+\ell^-)$, requiring both of them to be below $5100\,\mathrm{MeV}/c^2$, corresponding roughly to $m_{B^0} - m_\pi$. Since part of these decays is combinatorial in nature, these contributions are further suppressed by the combinatorial multivariate classifier described in Section 5.2.6.

5.2.5.6 Summary of backgrounds considered in fits

The set of additional selections applied to reduce the specific backgrounds discussed in this section is summarized in Table 5.8. The signal efficiency of all applied vetoes is very high and corresponds to $(97.62 \pm 0.14)\%$, $(94.86 \pm 0.67)\%$, $(97.34 \pm 0.11)\%$, $(94.24 \pm 0.74)\%$, for $B^0 \to K^{*0}\mu^+\mu^-$, $B^0 \to K^{*0}e^+e^-$, $B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$ and $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ decays, respectively.

When possible, the expected size of the backgrounds discussed is estimated in different regions of q^2 through the use of simulated events. The expected number

Table 5.8: Summary of the exclusive background selection requirements.

Background	**Requirement**	**Applied to**
$B_s^0 \to \phi\ell^+\ell^-$	`not`$\big(m(K(\pi \to K)) < 1040\,\text{MeV}/c^2$ `and Pi_ProbNNpi` $< 0.8\big)$	all
$B^0 \to (D^0 \to K\pi)\pi^-\ell^+\nu$	`not`$\big(\lvert m(K^+\ell^-{}_{\to\pi}) - m^{PDG}(D^0)\rvert < 30\,\text{MeV}/c^2)$ `andL_ProbNNl` $< 0.8\big)$	all
$B^0 \to (D^- \to K\pi\pi)\ell^+\nu$	`not`$\big(\lvert m(K^+\pi^+\ell^-{}_{\to\pi}) - m^{PDG}(D^-)\rvert < 30\,\text{MeV}/c^2$ `and L_ProbNNl` $< 0.8\big)$	all
$h \leftrightarrow \ell$ swap	`not`$\big(\lvert m(\mu_{\to h}\mu) - m^{PDG}_{J/\psi,(\psi(2S))}\rvert < 60\,\text{MeV}/c^2$ `and M_ProbNNmu` $< 0.8\big)$	all $\mu\mu$
	`not`$\big(\lvert m_{J/\psi,(\psi(2S))constr.}(h_{\to e}he_{\to h}e) - m^{PDG}_{B^0}\rvert < 60\,\text{MeV}/c^2$ `and E_ProbNNe` $< 0.8\big)$	all ee
$B^+ \to K^+\ell^+\ell^-$	$max\big(m(K\ell\ell), m((\pi \to K)\ell\ell)\big) < 5100\,\text{MeV}/c^2$	all

of background events, N_{BKG}, is estimated using the following expression:

$$N_{\text{BKG}} = \frac{f_{\text{BKG}}}{f_{\text{control}}} \cdot \frac{\mathcal{B}_{\text{BKG}}}{\mathcal{B}_{\text{control}}} \cdot \frac{\epsilon_{\text{BKG}}}{\epsilon_{\text{control}}} \cdot N_{\text{control}}\,, \tag{5.3}$$

where f_{H_b} represents the probability that a b quark hadronizes in the hadron species H_b, $\mathcal{B}_X$ is the branching ratio of the process $H_b \to X$, ϵ_X is the reconstruction efficiency of the decay as obtained from the corrected simulated samples and N_{control} is the observed number of events of the control modes $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ or $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$, depending on the q^2 region considered. Due to the limited statistics available for some of the simulated samples after the full selection, the estimated background is provided as an upper 90% confidence interval rather than a central value with its uncertainty. Table 5.9 (top) reports the estimated efficiency and expected number of events separately for RUN 1 and RUN 2 for the regions of central-q^2_c and J/ψ-q^2_c. Tables for the corresponding q^2 regions in the muon channels can be found in Table 5.9 (bottom).

Based on these tables, it is clear that not all backgrounds can be reduced to a negligible level and must therefore be included in the final fits. Table 5.10 summarizes the decay modes that are considered for the different regions of q^2 under study.

5.2.6 Combinatorial background rejection

Random combinations of particles coming from the fragments of the same p-p collision can become an important source of background when they have invariant masses close to the B^0 meson mass. Even though the requirements discussed in Sec. 5.2.1 greatly reduced the size of this background, a sizable fraction still survives. A multivariate analysis (MVA) has been designed to further reduce this contribution. The algorithm takes, as input, samples that have been labelled as *signal* and *background*, together with a list of variables, usually referred to as *features*, which are used for the discrimination of the two samples. The algorithm is then trained to exploit the discriminating power of the different features, including their multidimensional correlations. The output is a single

Table 5.9: Expected background yields after the full selection in the central-q_c^2 (q^2) and J/ψ-q_c^2 (q^2) regions for the electron (muon) channels. If the number of simulated events after full selection is < 20, a limit at 90% CL is given, else the expected yield is calculated.

Channel	central-q_c^2		J/ψ-q_c^2	
	bkg/limit Run 1	bkg/limit Run 2	bkg/limit Run 1	bkg/limit Run 2
$B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)\ell^-\bar{\nu}_\ell)\ell^+\nu_\ell$	8.28 ± 0.75	40.16 ± 3.10	118.93 ± 7.45	613.99 ± 37.23
$B^0 \to K^{*0}e^+e^-$	83.65 ± 2.19	319.23 ± 7.43	105.62 ± 2.71	398.12 ± 9.23v
$B^0 \to K^{*0}J/\psi(\to e^+e^-)$	7.18 ± 1.38	29.70 ± 4.50	54920.39 ± 538.95	207009.0 ± 1127.8
$B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$	< 0.32	< 0.73	4.30 ± 0.50	14.65 ± 1.17
$B^0 \to K^{*0}J/\psi(\to e^+e^-)(h \leftrightarrow e)$	< 1.57	< 6.61	< 1.57	< 6.60
$B^0 \to K^{*0}J/\psi(\to e^+e^-)(h \leftrightarrow h)$	0.94 ± 0.50	4.32 ± 1.79	238.12 ± 8.06	317.90 ± 14.90
$B_s^0 \to \phi e^+e^-$	0.94 ± 0.11	1.45 ± 0.17	0.80 ± 0.10	1.19 ± 0.14
$B_s^0 \to \phi J/\psi(\to e^+e^-)$	< 1.36	< 0.99	189.16 ± 22.70	294.08 ± 29.00
$B^+ \to K^+\pi^-\pi^+J/\psi(\to e^+e^-)$	< 5.40	< 27.97	1783.69 ± 316.25	4861.85 ± 863.61
$\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$	< 1.10	1.04 ± 0.43	217.84 ± 30.31	244.24 ± 33.84
$B^0 \to K^{*0}\psi(2S)(\to e^+e^-)(h \leftrightarrow e)$	< 0.30	< 0.62	< 0.30	< 0.62
$B^0 \to K^{*0}\psi(2S)(\to e^+e^-)(h \leftrightarrow h)$	< 0.30	< 0.62	0.20 ± 0.09	0.44 ± 0.16
$B_s^0 \to \phi\psi(2S)(\to e^+e^-)$	< 0.16	< 0.20	< 0.19	< 0.23
$\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-)$	< 0.03	< 0.08	< 0.03	< 0.09
$B^+ \to K^+e^+e^-$	< 0.27	< 0.35	< 0.27	< 0.36
$B^+ \to K^+J/\psi(\to e^+e^-)$	< 6.80	< 8.53	< 8.15	< 17.72

Channel	central-q_c^2		J/ψ-q_c^2	
	bkg/limit Run 1	bkg/limit Run 2	bkg/limit Run 1	bkg/limit Run 2
$B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$	< 2.06	< 8.38	245655.02 ± 788.05	745650.95 ± 2017.51
$B^0 \to K^{*0}\mu^+\mu^-$	558.07 ± 12.61	1732.46 ± 38.49	646.04 ± 14.55	1965.90 ± 43.57
$B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-)$	< 0.37	< 0.76	2.97 ± 0.43	4.89 ± 0.64
$B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)(h \leftrightarrow h)$	12.92 ± 2.00	18.07 ± 3.85	1423.85 ± 21.24	1823.28 ± 39.06
$B_s^0 \to \phi J/\psi(\to \mu^+\mu^-)$	< 2.63	1.70 ± 0.56	821.18 ± 95.30	1139.23 ± 111.54
$\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$	< 12.17	< 2.71	971.87 ± 139.59	784.79 ± 107.66
$B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-)(h \leftrightarrow h)$	1.09 ± 0.25	1.58 ± 0.35	0.47 ± 0.16	0.71 ± 0.23
$B_s^0 \to \phi\psi(2S)(\to \mu^+\mu^-)$	< 0.26	0.29 ± 0.11	< 0.36	< 0.37
$\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-)$	< 0.13	< 0.13	< 0.13	< 0.13

Table 5.10: Sources of background which must be taken into account in data fits to $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ decays, along with the q^2 region of importance.

Decay mode	q^2_c **(q^2) region**	Importance in fit
$B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$	central	in ee mode
$B^0 \to (D^- \to K^{*0}\ell\nu)\ell\nu$	central	in ee mode
$B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$	central	in ee mode
$B^0_s \to \overline{K}^{*0} J/\psi\,(\to \ell^+\ell^-)$	J/ψ	yes
$\Lambda^0_b \to pK J/\psi\,(\to \ell^+\ell^-)$	J/ψ	yes
$B^0 \to K^{*0} J/\psi(\to \ell^+\ell^-)(h \leftrightarrow h)$	J/ψ	yes
$B \to X J/\psi(\to \ell^+\ell^-)$	J/ψ	yes
$B^0_s \to \phi J/\psi\,(\to \ell^+\ell^-)$	J/ψ	yes
$B^0 \to K^{*0}\ell^+\ell^-$	J/ψ	merged with signal component
$B^0 \to (D^- \to K^{*0}\ell\nu)\ell\nu$	J/ψ	merged with combinatorial

function, called multivariate classifier, that optimizes the separation of the signal and background categories in the training samples. Given a set of features, the multivariate classifier used in this work will provide a value between zero and one, where values close to zero (one) indicate a large probability of the set of features to come from a signal (background) candidate. A selection on this classifier will effectively correspond to a multidimensional cut in the features provided for the training. An optimization procedure allows determining the threshold on the classifier response that guarantees the best compromise between *signal efficiency* and *background rejection*. In this thesis, the MVA used is a gradient-boosted decision tree [141], as implemented in the `XGBoost` python library [142]. The algorithm makes use of an ensemble of decision trees, *i.e.* a set of sequential binary cuts in the features provided, which are applied one after the other to improve the description of data with respect to the previous prediction step (*gradient-boosting*). Even if taken singularly the decision trees have a poor performance, they become very efficient classifiers when considered collectively.

5.2.6.1 Training and testing the multivariate classifier

The MVA is trained to distinguish between combinatorial background data and signal simulated decays. It is trained separately for electrons and muons and for the different runs Run 1, Run 2p1 and Run 2p2, while no distinction in trigger categories is considered.

The background training samples are obtained from data that correspond to the regions of $m(K\pi\ell\ell)$ above 5450 MeV/c^2 and 5600 MeV/c^2 for muons and electrons, respectively. The choice of the higher invariant mass threshold for the electrons is justified by the worse invariant mass resolution with respect to muons. Multivariate techniques greatly increase in discriminating power as the size of the samples used grows. For this reason, some requirements of the nominal

Table 5.11: Summary of the selection requirements that are modified from the nominal choice for training samples.

Type		**Requirement**
PID	K	$\texttt{DLL}_{K\pi} > -5$ $\texttt{ProbNNk} \cdot (1 - \texttt{ProbNNp}) > 0.05$
	π	$\texttt{ProbNNpi} \cdot (1 - \texttt{ProbNNk}) \cdot (1 - \texttt{ProbNNp}) > 0.1$
	μ	$\texttt{ProbNNmu} > 0.05$
	e	$\texttt{DLL}_{e\pi} > 0$ $\texttt{ProbNNe} > 0.05$
K^{*0}	K^{*0}	$\lvert m(K\pi) - m^{PDG}_{K^{*0}}(895.81)\rvert < 200\,\mathrm{MeV/c^2}$
Kinematic	$\mu^+\mu^-$	$0.1 < q^2 < 8.0\,\mathrm{MeV^2/c^4}$
	e^+e^-	$0.1 < q^2_c < 7.0\,\mathrm{MeV^2/c^4}$
	$m_{B^0\to K^{*0}\mu^+\mu^-}$ (bkg)	$> 5450\,\mathrm{MeV/c^2}$
	$m_{B^0\to K^{*0}e^+e^-}$ (bkg)	$> 5600\,\mathrm{MeV/c^2}$
Removed	$B^0 \to (D^0 \to K\pi)\pi^-\ell^+\nu$	$\texttt{not}(\lvert m(K^+\ell^-{}_{\to\pi}) - m^{PDG}(D^0)\rvert < 30\,\mathrm{MeV}$ $\texttt{and L\_ProbNNl} < 0.8)$
	$B^0 \to (D^- \to K\pi\pi)\ell^+\nu$	$\texttt{not}(\lvert m(K^+\pi^+\ell^-{}_{\to\pi}) - m^{PDG}(D^-)\rvert < 30\,\mathrm{MeV}$ $\texttt{and L\_ProbNNl} < 0.8)$
	$h \leftrightarrow \ell$ swap	$\texttt{not}(\lvert m(\mu_{\to h}\mu) - m^{PDG}_{J/\psi,(\psi(2S))}\rvert < 60\,\mathrm{MeV}$ $\texttt{and M\_ProbNNmu} < 0.8)$ $\texttt{not}(\lvert m_{J/\psi,(\psi(2S))constr.}(h_{\to e}he_{\to h}e) - m^{PDG}_{B^0}\rvert < 60\,\mathrm{MeV}$ $\texttt{and E\_ProbNNe} < 0.8)$

selection are either removed or loosened to increase the statistics available. The changes correspond to:

- the removal of the vetoes against $B^0 \to (D^0 \to K\pi)\pi^-\ell^+\nu$ and $B^0 \to (D^- \to K\pi\pi)\ell^+\nu$ decays;
- the removal of the cuts against $h \leftrightarrow \ell$ swap and clones, that are expected to have a negligible impact in this region of the $m(K\pi\ell\ell)$ spectrum;
- a wider $m(K\pi)$ mass window around the K^{*0} mass;
- a wider central-q^2 and central-q^2_c definition for muons and electrons, where the lower limits are shifted to $0.1\,\mathrm{GeV}^2/c^4$ while in muons the higher limit is shifted to $8\,\mathrm{GeV}^2/c^4$;
- looser requirements on identification of all particle species, as summarized in Table 5.11.

This allows to increase the statistics available for the training of roughly twice for the muon sample and three times for the electron.

The signal samples are obtained from $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ simulated samples after the full chain of corrections described in Ch. 6. The selections applied on these simulated samples are aligned to the ones described for the background samples with two main exceptions: the additional requirement that the selected events are truth-matched as *Signal* and the inclusion of events from the whole $m(K\pi\ell\ell)$ spectrum, not just the sideband. The set of features used in the training of the MVA is established by means of a search procedure between a list of variables that are expected to provide a good separation between

Table 5.12: Summary of the input features used in the MVA training that provided the most discriminating perfomance.

Particle	**Variables**
B^0	p_{T}, $\chi^2_{\mathrm{IP\_OWNPV}}$, $\chi^2_{\mathrm{FD\_OWNPV}}$, χ^2_{vtx}/ndf, χ^2_{DTF}/ndf, DIRA
K^{*0}, J/ψ	χ^2_{vtx}/ndf
h	$\min(p_{\mathrm{T},K}, p_{\mathrm{T},\pi})$, $\min(\chi^2_{\mathrm{IP\_OWNPV},K}, \chi^2_{\mathrm{IP\_OWNPV},\pi})$
ℓ	min,max($p_{\mathrm{T},e^+}, p_{\mathrm{T},e^-}$), min,max($\chi^2_{\mathrm{IP\_OWNPV},e^+}, \chi^2_{\mathrm{IP\_OWNPV},e^+}$)

signal and candidates. The features considered are mainly related to the quality of tracks and vertices of the reconstructed decay, the transverse momentum of the final state particles and of the B^0 meson. Multiple classifiers are then trained with different combinations of these features and the configuration that gives the best discriminating performance is chosen as default (see Table 5.12). To make the best use of the statistics available, the training and testing of the MVA is done using a *k-folding* approach. This corresponds to randomly split the signal and background samples in k sub-samples of roughly the same size and use $k-1$ sub-samples to train the algorithm, while the MVA performance in tested on the sample that was not used for the training. In the following, a *k-folding* approach with ten folds is chosen. Whenever an MVA score is assigned to an event, the output of one of the classifiers that never encountered that event is chosen at random.

The outcome of the training for the MVA considered is summarized in Figures 5.7, which show the ten ROC curves obtained in Run 2p2 for muons and electrons, respectively. Similar performances have been obtained for Run 1 and Run 2p1. The ROC (Receiver Operating Characteristic) curve is a standard tool to describe the performance of classifier at all possible thresholds. It represents in a two-dimensional plane the *true positive rate*, defined as the fraction of signal events that have been classified as signal, against the false positive rate, defined as the fraction of backgrounds events that have been classified as signal. The area under the curve (AUC) provides a way to evaluate the performance of the classifier: the closer its AUC is to unity, the better a classifier is considered to be. This is also the criterion used to select the set of features in Table 5.12. All ten folds of each classifier behave similarly and are characterized by large values of the AUC, suggesting a classifier with good signal-background separation.

The performance of the classifier on training and test samples is studied by comparing the MVA response for the two sets of samples. Figure 5.8 shows this comparison for muons and electrons, respectively. Some differences can be noticed for MVA scores close to one, suggesting that the classifier might be slightly overtrained[3]. This behaviour is expected to be irrelevant and will not

---

[3] With the term *overtrain* we refer to the fact that a classifier performs better in the sample that was trained with respect to the test sample, suggesting that some of its discriminating power came by learning a specific feature of the training sample that is not present in the test

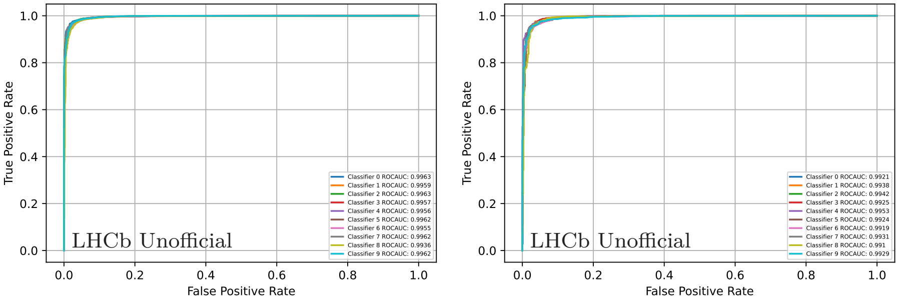


Figure 5.7: ROC curves corresponding to the muon (left) and electron (right) mode classifiers for Run 2p2.

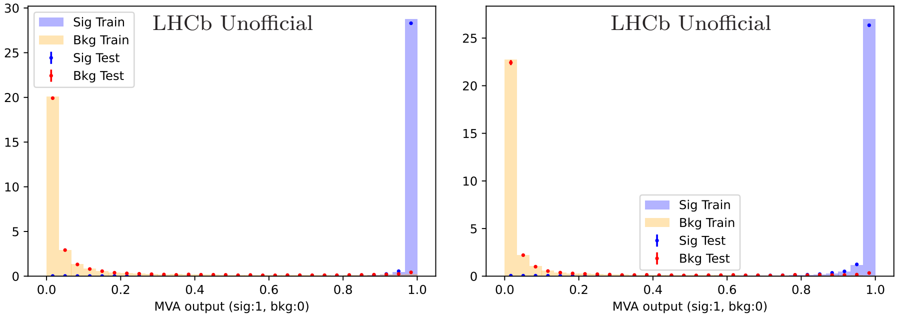


Figure 5.8: Comparison between the performance of the muon (left) and electron (right) mode classifiers on the signal and background training and testing samples for Run 2p2.

be further discussed, since it only implies a small reduction of the performances of the algorithm on events that it has never seen.

The uniformity of the classifier response is studied by computing the fraction of signal and background events that pass a set of different MVA thresholds as a function of the angles $\cos\theta_K$, $\cos\theta_\ell$ and ϕ, the dilepton invariant mass squared q^2 (q^2_c) and the masses $m(K\pi\ell\ell)$ and $m(K\pi)$. The thresholds correspond to the 20th, 40th, 60th and 80th percentiles of the signal and background[4] samples considered, while the bin edges have been chosen to ensure roughly the same number of events in each bin. Figure 5.9 and 5.10 show the result of these uniformity checks for Run 2p2 in muons and electrons, respectively. While non-trivial features are observed in most of the distributions, there are a few patterns that can be highlighted:

---

one. As a consequence, the algorithm will perform worse on real data than on the training sample.

[4] Events with MVA values very close to zero are not included on this study, since they do not represent well the features of the combinatorial background observed in the final sample.

- a small linear trend coherent between all different thresholds can be noticed in the q_c^2 and $m(K\pi ee)$ distribution for the electron background sample;
- a strong trend in the $m(K\pi ee)$ distribution of the electron signal sample, where the cuts result more efficient for events close to the nominal B^0 meson mass;
- a small trend in the $\cos\theta_\ell$ distribution of the electron signal sample;
- a small trend in the $m(K\pi\mu\mu)$ distribution of the muon background sample, similar in trend to the one observed in the corresponding electron sample;
- a trend in the $\cos\theta_K$ distribution of the signal muon sample, linked to the differences between the transverse momentum distribution of B^0 meson in background and signal distributions as explained in Sec. 6.7, and that might also be present in the electron sample, even if less visible.

While a completely flat response on the variables of interest is desirable, the non-uniformities discussed here are not impactful. Most of the trends are common across different MVA cuts, suggesting that signal and background distributions will not change much for different choices of the MVA threshold. Additionally, in the signal case, these variations are taken into account when producing the acceptance functions in Sec. 7.2, therefore no bias is introduced in the final fit if the simulation models these effects correctly. Distortions of the $m(K\pi\ell\ell)$ invariant mass for the combinatorial background could in principle be a problem, since its contribution is extracted from data assuming an exponentially decreasing distribution. However, due to the linear trend observed in both electrons and muons, the exponential hypothesis is expected to be still valid after a cut on the MVA classifier.

5.2.6.2 Cut optimization

The choice of the thresholds used for the MVA selection in electrons and muons is performed by optimizing the statistical significance of the signal with respect to the combinatorial background in the region populated by the signal, separately for each run. This is approximated by the figure of merit (FoM), defined as

$$\frac{S}{\sqrt{S+B}}\,, \tag{5.4}$$

where S and B correspond to the signal and combinatorial yield as obtained in the region of $m(K\pi\ell\ell)$ of [5100, 5400] MeV/c^2. This region has been chosen because it is where most of the signal peaks. Values of the MVA cut for which the FoMs are maximized correspond to the nominal choice of the cut, unless otherwise stated. A direct estimate of the signal yield from the data fits is avoided, since it could introduce a bias in our optimization procedure. The expected signal yield after the cut MVA $> x$ is estimated instead as

$$N_{sig}^{exp}(\text{MVA} > x) = \frac{\epsilon_{sig}^{MC}(\text{MVA} > x)}{\epsilon_{sig}^{MC}(\text{MVA} > x_{ref})} \cdot N_{sig}^{fit}(\text{MVA} > x_{ref})\,, \tag{5.5}$$

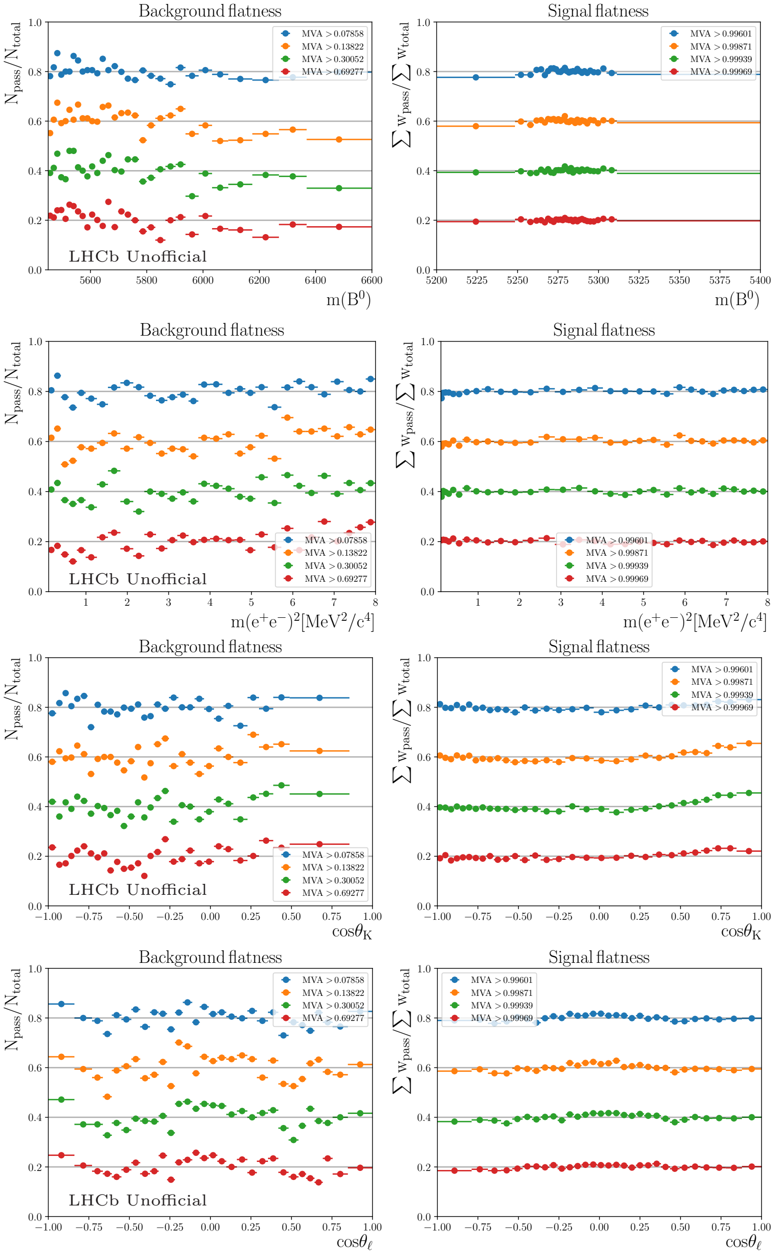


Figure 5.9: Efficiency as a function of the B^0 invariant mass, q^2 and angles for the Run 2p2 classifier of the muon mode. The efficiency as a function of ϕ is not reported here due to its flat behavior in both signal and background.

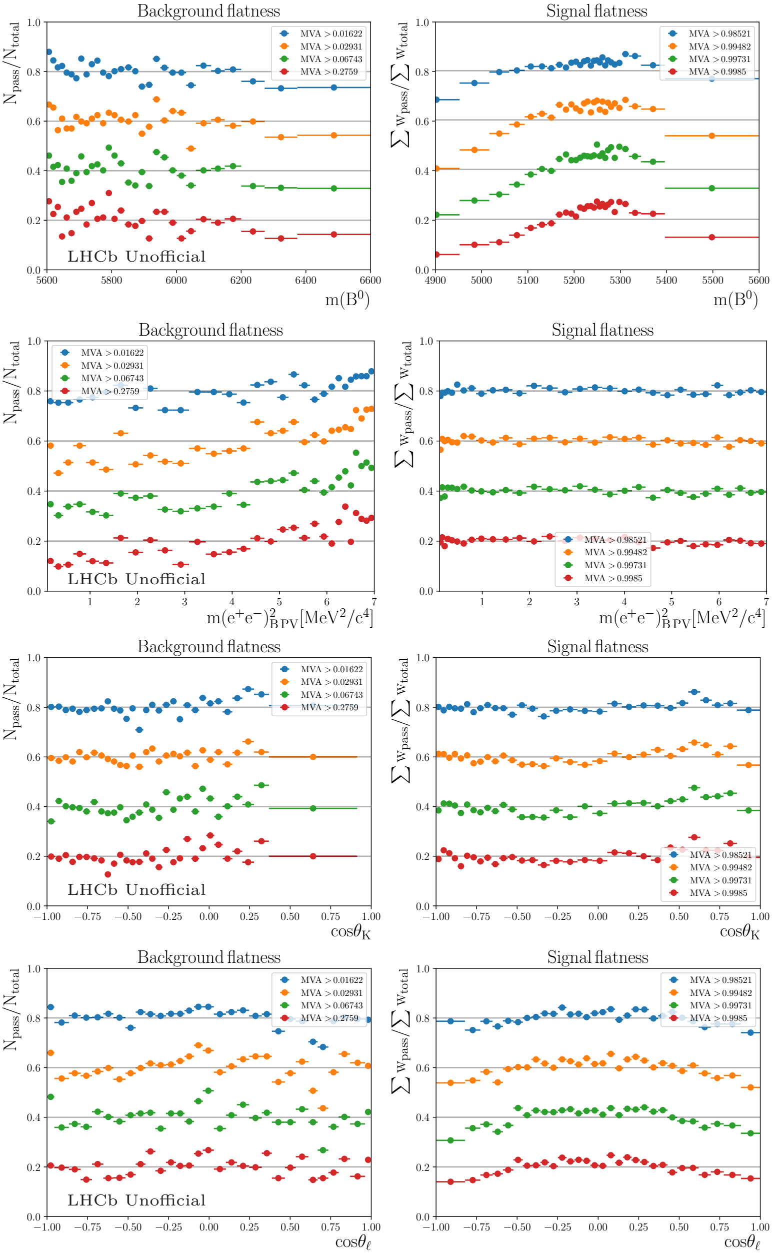


Figure 5.10: Efficiency as a function of the B^0 mass, q^2_c and angles for the Run 2p2 classifier of the electron mode. The efficiency as a function of ϕ is not reported here due to its flat behavior in both signal and background.

where $N_{sig}^{fit}(\text{MVA} > x_{ref})$ corresponds to the yield observed in a mass fit at the reference cut value of MVA $> x_{ref}$, and $\frac{\epsilon^{MC}(\text{MVA}>x)}{\epsilon^{MC}(\text{MVA}>x_{ref})}$ corresponds to the expected change in the observed yield due to the different efficiency of the cut as obtained from simulation, ϵ^{MC}. A similar strategy is also used for the estimation of the partially reconstructed background in the optimization for the electron channel. The combinatorial background yield is instead estimated from a direct fit to data at different thresholds. In all cases, a signal component is also present in such fits, but the value obtained for its yield is kept blinded and not considered.

5.2.6.3 Electron mode

The optimization of the FoM in the electron mode is performed separately for Run 1, Run 2p1 and Run 2p2, without distinguishing between the different trigger categories. Simultaneous fits to reconstructed candidates in the regions of q^2 corresponding to central-q_c^2 and J/ψ-q_c^2 are obtained for a value of the MVA threshold x. The invariant mass shapes used for the signal are obtained from fully corrected simulated samples of $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ decays after the set of selections described in Sec. 5.2.6.1. Their parametrization follows the same strategy described in Sec. 8.2. When fitting data, a common shift and scale of the width of the two distributions is allowed to float freely in data. An additional cut on $m_{J/\psi}^{DTF} > 5150\,\text{MeV}/c^2$ is applied to remove the partially reconstructed background in the J/ψ-q_c^2 bin. The combinatorial background is parametrized by a decreasing exponential with slope λ_{exp}, separately for the different q^2 regions and runs considered. The nominal fit in the central-q_c^2 region (see Ch. 10) would additionally consider two more background components: the partially reconstructed and double-semileptonic contribution. The latter cannot be distinguished from the combinatorial without including the angular distributions in the fit and is, for simplicity, partially vetoed by applying the selection $|\cos\theta_\ell| < 0.8$, whose effect has already been discussed in Sec. 5.2.5.3. The partially reconstructed background is included in the fit and its shape is parametrized from simulation using a *kernel density estimator* (KDE) lineshape [143].

The optimization procedure is then performed in three steps:

1. Simultaneous mass fits are performed for data events in the regions of central-q_c^2 and J/ψ-q_c^2 that pass the reference cut MVA $> x_{ref}$, where x_{ref} corresponds to the values of 0.995, 0.98 and 0.995 for Run 1, Run 2p1 and Run 2p2, respectively. In these reference fits, the partially reconstructed yield is allowed to float freely. In the specific case of Run 2p2, due to the strong suppression performed by the MVA cut, the fit is found to be unstable when both the combinatorial and partially reconstructed component are considered. For this reason, only the latter is used. Figure 5.11 shows the result of these reference fits, while Table 5.13 reports the corresponding fit parameters for the central-q_c^2 region.

Table 5.13: Results obtained from the reference point fits. An extended maximum likelihood fit is made to directly obtain yields and their uncertainties.

Parameter	Run 1	Run 2p1	Run 2p2
N_{Sig}	77±14	113±18	175±15
N_{Comb}	30±43	61±60	–
$N_{\mathrm{Part.\,reco.}}$	19±36	39±55	43±10
$\lambda_{\mathrm{exp.}}$	-0.003±0.003	-0.005±0.002	–

2. Almost the same procedure is then repeated for cut values x ranging from 0.50 to 0.999. The scan is made in steps of size 0.1 from 0.50 to 0.99, and in steps of size 0.001 from 0.99 to 0.999. The region of 0.962 to 0.988 is additionally scanned using step size of 0.002. The only difference is that the yield of the partially reconstructed background is not allowed to float freely, but instead fixed to the efficiency scaled expected value, as obtained from Eq. 5.5 substituting “sig” with “partially reconstructed”.

3. For each of the values x used in the previous step, it is possible to estimate the signal yield S scaling the signal yield observed for x_{ref} and using simulation, as described by Eq. 5.5. The background yield, B, can be estimated by the integrated number of the combinatorial events in the range between [5100,5400] MeV/c^2, as obtained from the data fit. These two quantities can be combined, as shown in Eq. 5.4, to compute the signal significance for each of the x MVA thresholds discussed. The optimization procedure consists then in finding the value of x that maximizes this FoM.

The FoM values obtained for Run 1, Run 2p1 and Run 2p2 are shown in Figure 5.12, together with their corresponding optimal working points of 0.99, 0.98 and 0.97, respectively. Due to the vicinity of these optimal values, a simplified strategy where a single MVA value is chosen for all runs is preferred. Since the variation of the MVA response is smaller for looser cuts, the common value of 0.97 is chosen.

5.2.6.4 Muon mode

A strategy similar to the one discussed for the electrons can be also used in the muon mode with further simplifications: no partially reconstructed and double-semileptonic components need to be considered in these mass fits. Additionally, since the shift and scale of the width of the distribution can be obtained directly from the fit to the data in central-q^2 region, there is no need for a corresponding fit in the J/ψ-q^2 region. The reference fits for the three runs are all performed at a loose MVA value of $x_{ref} = 0.5$, and the results are gathered in Figure 5.13 and Table 5.14. The MVA cut is then scanned for values ranging from 0.0 to 0.99, with a fit made for every increase of 0.01. An additional scan is performed for values ranging from 0.962 to 0.998, with step sizes of 0.002. The variation of the FoM as a function of the MVA cut is shown in Figure 5.14. The optimal values corresponding the cut at which the FoM is maximized correspond to 0.962, 0.96 and 0.92 for Run 1, Run 2p1 and Run 2p2, respectively.

A common lower cut value of 0.8 is chosen. This choice is made based on the consideration of three main factors. Firstly, the limiting factor in this analysis is the statistics of the electronic mode, and thus, choosing a threshold that is less optimal according to the standard FoM is expected to have marginal impact on the sensitivities to the observables of interest. Secondly, the optimization is

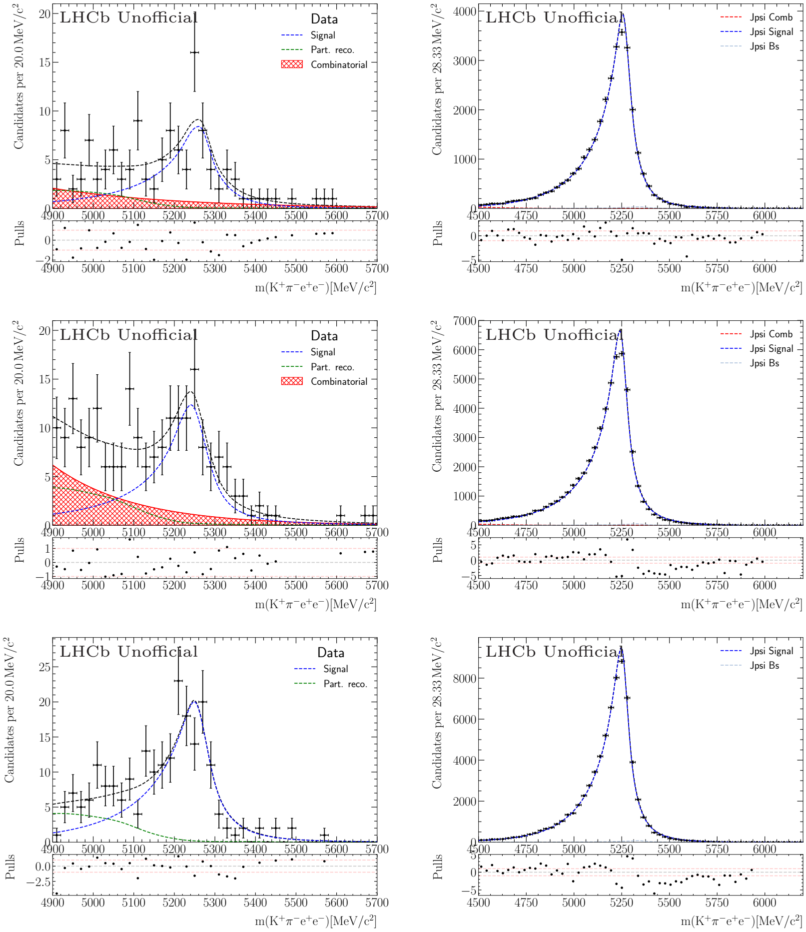


Figure 5.11: Fits to the mass distributions of the electron mode signal and control mode candidates with MVA references thresholds (x_{ref}) of 0.995, 0.98, and 0.995 for Run 1 (top row), Run 2p1 (middle row) and Run 2p2 (bottom row), respectively. For Run 2p2 the inclusion of the combinatorial component resulted in problematic fit result as it was strongly disfavoured, which led to the use of the partially reconstructed component only.

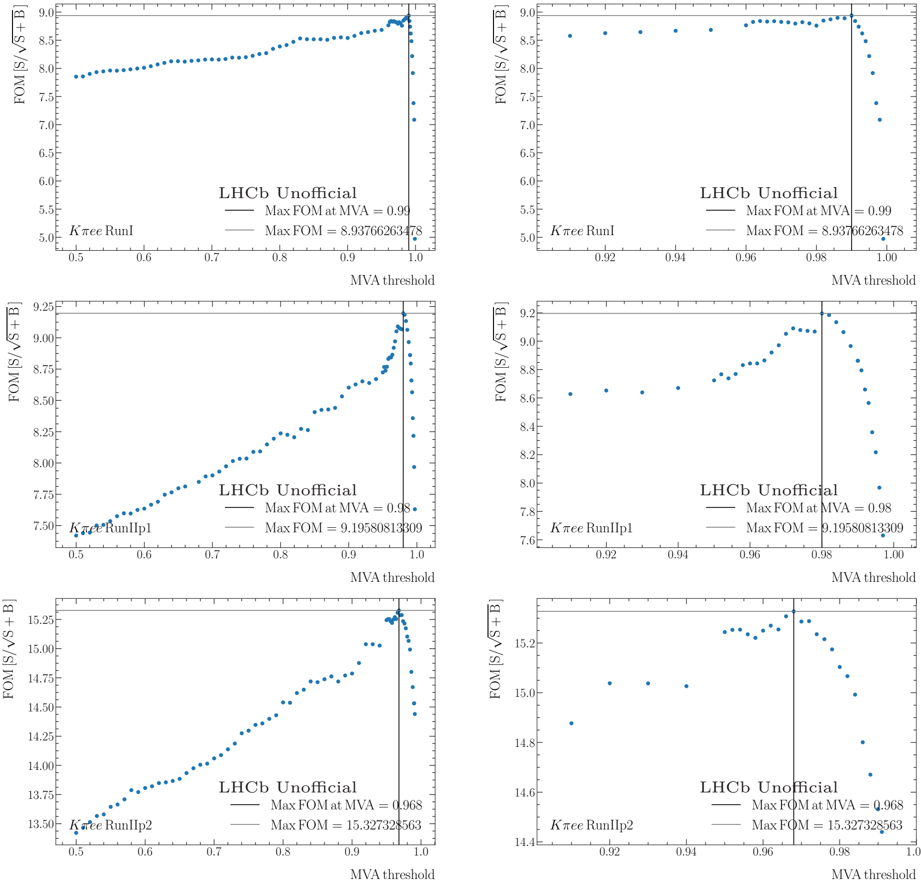


Figure 5.12: Calculated values of the standard FoM $S/\sqrt{S+B}$ for Run 1 (top), Run 2p1 (middle), and Run 2p2 (bottom) samples. The full set of results in the investigated range of $0.5 < \text{MVA} < 0.999$ is shown on the left hand side while results in the region of interest above MVA>0.9 are shown on the right hand side. The optimal value is found to range from 0.97 to 0.99.

Table 5.14: Results obtained from the muon mode reference point fits.

Parameter	Run 1	Run 2p1	Run 2p2
N_{Sig}	666±33	704±34	1444±47
N_{Comb}	711±34	592±33	990±42
$\lambda_{\mathrm{exp.}}$	-0.0032±0.0003	-0.0030±0.0003	-0.0029±0.0003

carried out considering the signal significance alone. In the case of the amplitude fit, very low levels of combinatorial background often lead to fit instabilities when the background angular shape is allowed to vary, hence a looser cut might be preferred. Finally, it is desirable to cross-check the final result of this analysis with the work of Ref. [144], which also considers a region between the two charmonium contributions, *i.e.* $11.0 \leq q^2 \leq 12.5\ \mathrm{GeV}^2/c^4$. The value of the MVA threshold chosen ensures enough combinatorial background in this additional region to still retain a high rate of convergence for each fit attempted in toys.

5.2.7 Multiple candidates

After the full selection, it is possible that more than one signal candidate per p-p collision enters the final sample. When these multiple candidates are present, only one is chosen randomly, while the others are discarded. Table 5.15 reports the fraction of removed candidates over the total.

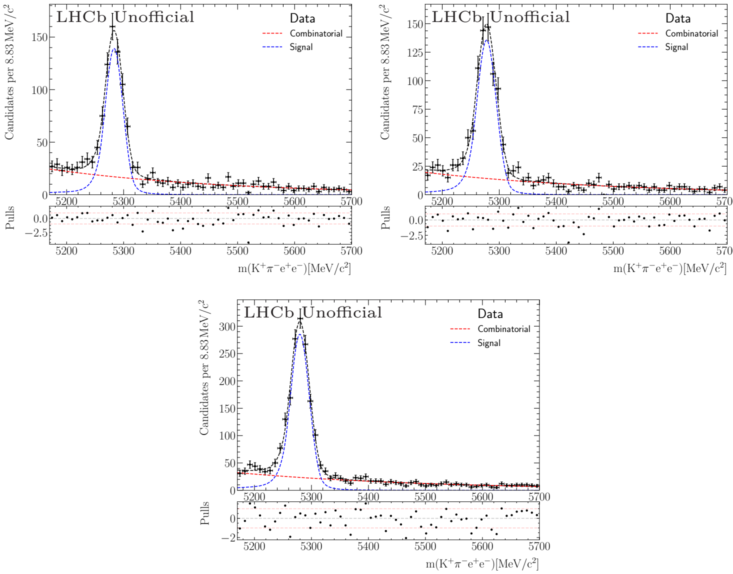


Figure 5.13: Fits to the mass distributions of the muon mode signal candidates selected with the reference threshold of MVA> 0.50.

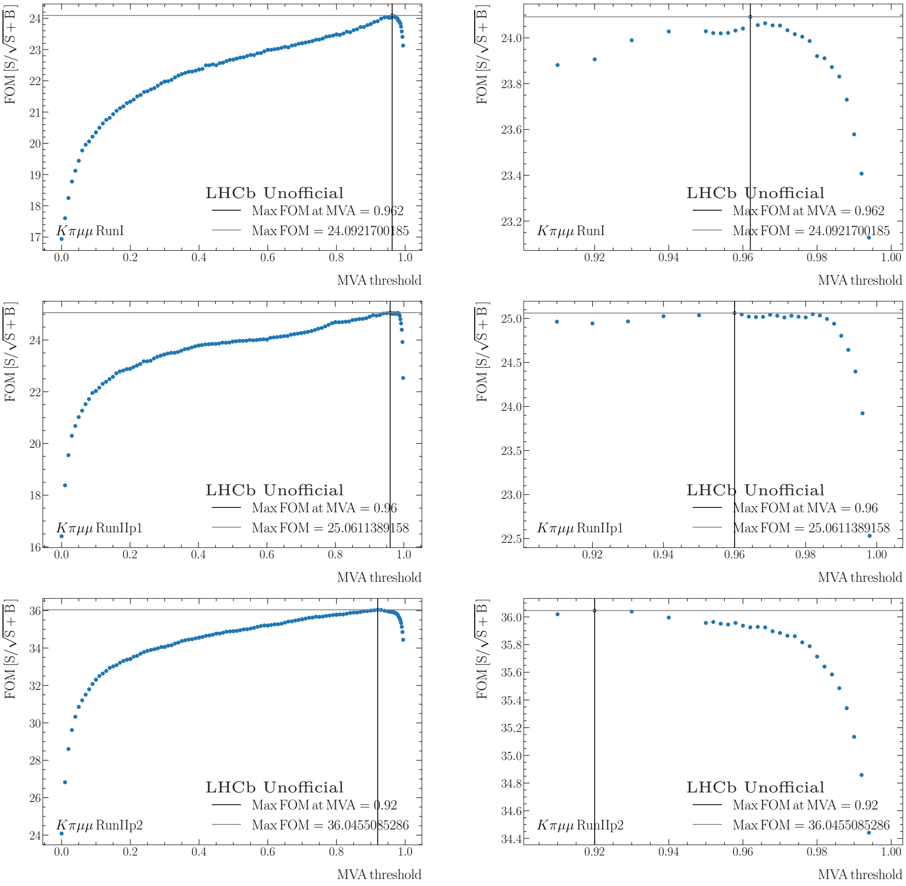


Figure 5.14: Calculated values of the standard FoM $S/\sqrt{S+B}$ for muon mode Run 1 (top), Run 2p1 (middle), and Run 2p2 (bottom) samples. The full set of results in the investigated range of $0.0 < \text{MVA} < 0.994$ is shown on the left hand side while results in the region of interest above MVA> 0.7 are shown on the right hand side. The optimal value is found to range from 0.92 to 0.96.

Inspecting the multiple candidates discarded in data, it has been observed that the far majority ($\approx 98\%$) has exactly two candidates per event. In Run 1, the majority is constituted of candidates where the same tracks are used but some of the particle hypothesis are swapped, while the rest is composed of candidates for which at least a track is different. In general, the additional track considered has a higher probability to be reconstructed as a pion. An opposite behaviour is observed instead in Run 2, where multiple candidates corresponding to a permutation of the particle hypothesis for the same tracks are the minority.

Table 5.15: Fraction of removed multiple candidates for simulation and data.

	Simulation		Data	
Channel	Run 1	Run 2	Run 1	Run 2
$B^0 \to K^{*0}\mu^+\mu^-$	$(0.17 \pm 0.03)\%$	$(0.16 \pm 0.01)\%$	$(0.8 \pm 0.3)\%$	$(0.6 \pm 0.1)\%$
$B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$	$(0.242 \pm 0.005)\%$	$(0.257 \pm 0.004)\%$	$(1.16 \pm 0.02)\%$	$(0.64 \pm 0.01)\%$
$B^0 \to K^{*0}e^+e^-$	$(0.188 \pm 0.06)\%$	$(0.30 \pm 0.04)\%$	$(1.3 \pm 0.6)\%$	$(0.8 \pm 0.3)\%$
$B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$	$(0.566 \pm 0.014)\%$	$(0.564 \pm 0.009)\%$	$(1.2 \pm 0.04)\%$	$(0.84 \pm 0.02)\%$

Chapter 6

Corrections to simulation

Many aspects of the analysis depend on the ability of our simulation to closely represent the behaviour observed in data. For this reason, a series of data-driven corrections are performed on simulated samples to further reduce any relevant difference. In the following, a brief introduction on the structure of the different corrections is provided in Sec 6.1, while the Secs. from 6.2 to 6.6 detail the procedure used on each separate step. Sec. 6.7 ends the chapter with a discussion on the impact of the corrections in the variables used in the final amplitude fits.

6.1 Correction strategy

The corrections discussed in the following have two main goals: improve the agreement between the distributions of the multiplicity of the events and the reconstructed properties of the B^0 meson; improve the agreement between the efficiency observed in data and simulation for PID, tracking and trigger selections. All these aspects have the potential to bias the observables of interest if they are not properly corrected for.

The corrections are organized in a chain, composed of seven subsequent steps, sketched in blue in Figure 6.1. This correction chain is referred to as *nominal* in the following. The steps correspond to:

1. the correction of the PID efficiency for all the particles involved in the final state, encoded in the weight w_{PID} (Sec. 6.2);
2. the correction of the track reconstruction efficiency for electrons, encoded in the weight w_{TRK} (Sec. 6.3);
3. the correction of the generator level distribution of the event multiplicity and kinematics of the B^0 meson, encoded in the weight $w_{\mathrm{Mult\&Kin}}$ (Sec. 6.5);
4. the correction of the L0 trigger efficiency, encoded in the weight w_{L0} (Sec. 6.4.1);
5. the correction of the HLT trigger efficiency, encoded in the weight w_{HLT} (Sec. 6.4.2);
6. the correction of the reconstructed properties of the B^0 meson such as χ^2_{IP} and vertex χ^2, encoded in the weight w_{Reco} (Sec. 6.5).
7. the smearing of the dielectron and $K\pi e^+e^-$ invariant mass (Sec. 6.6).

The effect of these corrections can be combined by means of a simple multiplication, obtaining the total weight $w_{tot} = w_{\mathrm{PID}} \cdot w_{\mathrm{TRK}} \cdot w_{\mathrm{Mult\&Kin}} \cdot w_{\mathrm{L0}} \cdot w_{\mathrm{HLT}} \cdot w_{\mathrm{Reco}}$.

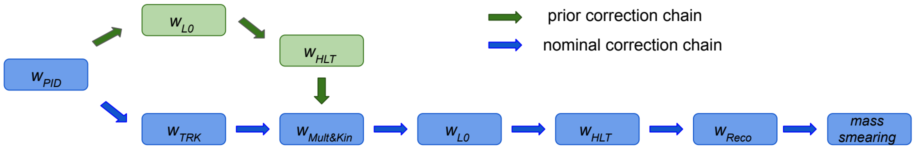


Figure 6.1: The *nominal* (blue) and *prior* (green) correction chains. The *prior* chain corrects PID, L0 and HLT efficiencies before accessing the B meson kinematics and underlying event multiplicity distributions.

All these corrections are obtained separately for different years in electrons and muons. When necessary, an additional separation in trigger categories is performed. Most of the corrections are derived from high statistics control samples of $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ and $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ decays, selected as similarly as possible to the resonant modes of $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ and $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ decays. To ensure the portability of the corrections from one channel to another, a parallel set of corrections is also obtained using the $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ channel and the compatibility of the two is verified. The main advantage of using different channels to correct the simulation is that this minimizes the correlation between the corrections and the simulated samples used in all other parts of the analysis. The only exception to this rule is the last correction step, obtained directly from the $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ decay channel.

Of all the steps involved in the nominal correction chain, $w_{\mathrm{Mult\&Kin}}$ is special, since it relies on a dedicated correction chain, referred to as *prior* correction chain (see Figure 6.1). The *prior* correction chain is computed from $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ decay candidates selected in the inclusive L0M trigger category, where the only data events considered are those for which the `TCK`s of the `L0Muon` trigger line are fully aligned with the simulation.[1] This greatly increases the agreement between data and simulation, reducing the size of the w_{L0} and w_{HLT} corrections in the *prior* correction chain. Samples of data and simulation are then compared after PID and trigger corrections: any difference observed in the kinematics of the B^+ meson and the event multiplicity is interpreted as a misalignment between the distributions generated by the simulation and the properties of the reconstructed events in data. Even if these correction weights are obtained from muons in a specific trigger category, since they represent a correction to the generated kinematic distribution of the decaying B^0 meson, they can be ported to different trigger categories and even different channels, as in the case of $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$, $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$, $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays.

---

[1] Data in 2011, 2012 and 2018 is almost fully aligned with the available simulation, while for 2015, 2016 and 2017 the size of the aligned sample corresponds to roughly 25%, 70% and 40%, respectively.

6.2 Particle identification corrections

The simulation of the detectors that are responsible for the particle identification is non-trivial, since a precise modelling of their response requires a very good understanding of the experimental conditions under which they are operated. For this reason, PID variables obtained from simulation are known to be unreliable and data calibration samples are used to determine the efficiency of specific PID selections. In the following, two different approaches are discussed.

The first approach is used for muons, kaons, protons and pions, and relies on the assumption that PID efficiencies for these different particle species are independent of one another, if not for the kinematic correlation related to their common mother-particle. Under these assumptions, the efficiency of different PID selections factorizes between the particles considered and each of these efficiencies can be determined using different calibration samples.

Pure samples of these different particle species can be obtained, without the use of particle identification on the probe track of interest, in the decays of:

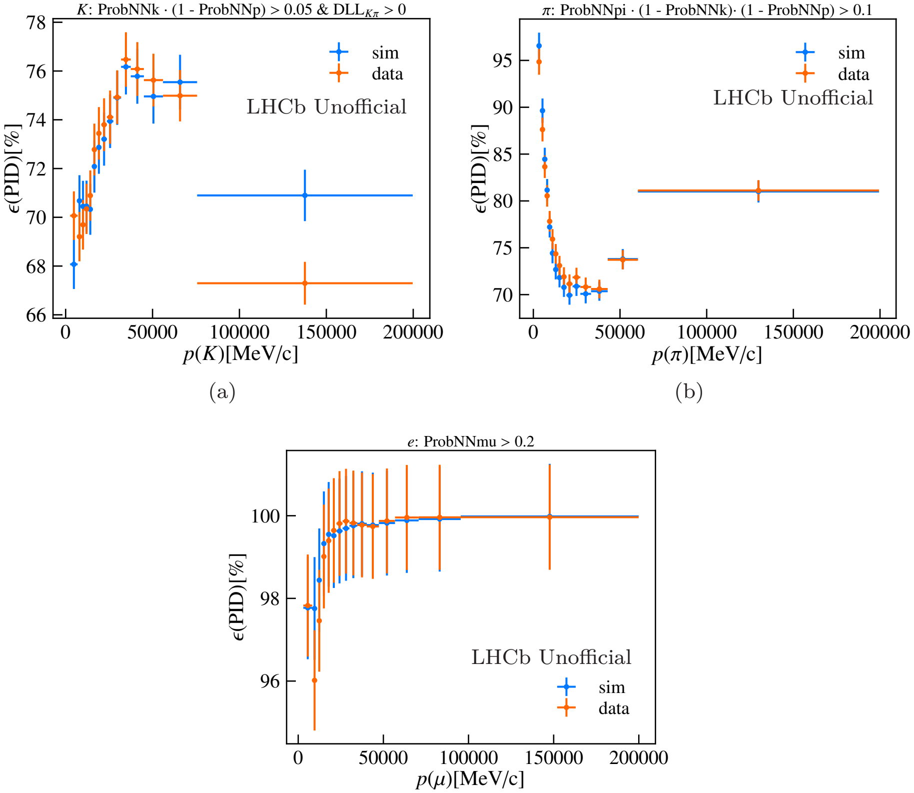
- $J/\psi \to \mu^+\mu^-$, where the J/ψ can be either prompt or coming from a B meson, to obtain pure samples of muons;
- $D^{*+} \to D^0\pi^+$, where pure samples of kaons and pions can be distinguished from the subsequent decay $D^0 \to K^-\pi^+$;
- $\Lambda^0 \to p\pi^-$ and $\Lambda_c^+ \to pK^-\pi^+$, where pure samples of protons are obtained for regions of low and high proton momentum, respectively.

All fiducial requirements applied to the calibration samples that may affect the PID efficiency measured must be also applied in signal and control modes, as already discussed in Sec. 5.2.2 and summarized in Table 5.3.

Two types of efficiency maps are computed: ID and mis-ID maps. The first ones correspond to the probability that a particle passes its corresponding PID selection, *e.g.* a muon that passes `ProbNNmu`>2. The second one corresponds to the probability that a particle passes a PID selection corresponding to a different particle hypothesis, *e.g.* a muon that passes $\texttt{DLL}_{K\pi} > 0$. Efficiency maps for the different particle species are then calculated as a function of the particle momentum, pseudorapidity and number of tracks in the event (`nTracks`). For each bin of event multiplicity considered, a fit to the invariant mass of the J/ψ, D^0 or $\Lambda_{(c)}$ hadron is performed to estimate the amount of signal and background in data. The background can then be statistically subtracted using the $_s\mathcal{P}lot$ method [145], and the 2D-distributions in particle momentum and pseudorapidity for the total and passed signal events can be obtained. The ratio of the passed over the total distribution corresponds to the efficiency maps needed for the PID correction. Each reconstructed track t in simulation is therefore associated with a weight $w^t_{\text{PID}}(p^t, \eta^t, \texttt{nTracks})_{data}$ obtained from the corresponding ID or mis-ID efficiency map in data, depending on the true identity of the particle considered. The total correction weight, w_{PID}, corresponds to the multiplication of all the weights obtained for the particles in the final state, *e.g.* for $B^0 \to K^{*0}\mu^+\mu^-$ decays:

$$w_{\mathrm{PID}}^{K^-\pi^+\mu^+\mu^-} = \prod_{i \in \{K^-,\pi^+,\mu^+,\mu^-\}} w_{\mathrm{PID}}^{i}(p^i, \eta^i, \texttt{nTracks})_{data} \tag{6.1}$$

It is important to point out that w_{PID} corresponds to the efficiency that a certain event passes the set of PID cuts described in Table 5.3, thus replacing the need to apply the corresponding PID requirements in simulation. A comparison of the selection efficiency obtained from simulation and data for the PID selections applied to kaons, pions and muons is shown in Figure 6.2.



(a)

(b)

(c)

Figure 6.2: Comparison between the efficiency in data and simulation for the PID selection of kaons (a), pions (b) and muons (c) as a function of the particle momentum for simulated events of $B^0 \to K^{*0}\mu^+\mu^-$ decays corresponding to the data-taking conditions of Run 2p1.

A different approach is followed to obtain the correction on the PID of the electrons; efficiency maps are obtained from data and simulation of

$B^+ \to K^+ J/\psi\,(\to e^+e^-)$ decays, and the ratio of the two is used to correct the efficiency observed in simulation. This approach has the advantage of obtaining a simulated sample with the efficiencies of the PID cuts as observed in data while still retaining the correlation between the PID response and the other features of the event, in particular the correlation with the other electron in the decay. The efficiency is corrected using weights that correspond to the ratio of the single-electron PID efficiencies in data and simulation. The total correction weight $w_{\rm PID}$ can be written, *e.g.* for $B^0 \to K^{*0}e^+e^-$ decays, as:

$$
\begin{aligned}
w_{\rm PID}^{K^-\pi^+e^+e^-} &= \prod_{i\,\in\{K^-,\pi^+\}} w_{\rm PID}^{i}(p^i, \eta^i, \texttt{nTracks})_{data} \\
&\times \prod_{i\,\in\{e^+,e^-\}} \frac{w_{\rm PID}^{i}(p_T^i, \eta^i, \texttt{nTracks})_{data}}{w_{\rm PID}^{i}(p_T^i, \eta^i, \texttt{nTracks})_{sim}}\,,
\end{aligned}
\tag{6.2}
$$

where $w_{\rm PID}^{i}(p_{(T)}^i, \eta^i, \texttt{nTracks})_{data}$ and $w_{\rm PID}^{i}(p_{(T)}^i, \eta^i, \texttt{nTracks})_{sim}$ correspond to the efficiencies obtained in data and simulation as a function of the variables used in the parametrization. To obtain the PID efficiency in data the ${}_s\mathcal{P}lot$ method cannot be employed reliably, due to the low statistics of the calibration samples and a combinatorial background shape that is correlated to the PID variables. A dedicated fit-and-count approach is developed where, for each bin of `nTracks`, p_T and η of the electron candidate, a fit to the events that pass or fail the electron PID selection is performed. An additional distinction between events that have or do not have an associated photon cluster in the ECAL is needed, due to the different behaviour that the PID response has between these two classes of events. [2] The efficiency in simulation is instead computed using a cut-and-count approach after the same set of selections applied in data. Figure 6.3 shows the efficiency of the electron PID cut in simulated $B^0 \to K^{*0}e^+e^-$ decays before and after the correction described, for events for which a photon compatible with a bremsstrahlung emission was not found (a) or was found (b).

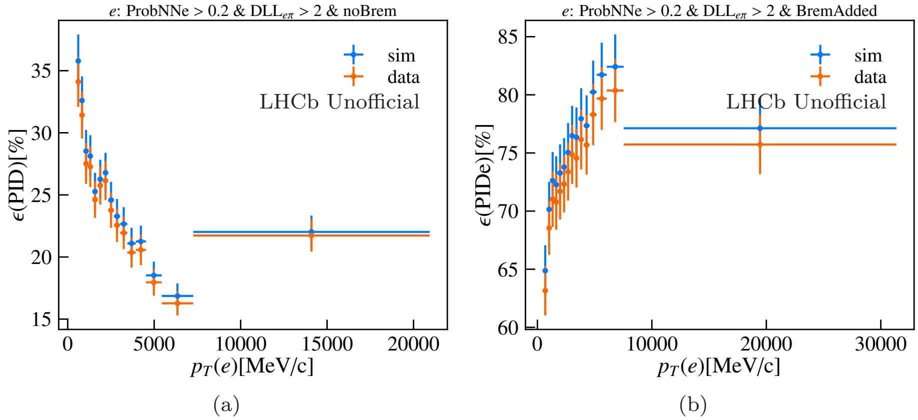
6.3 Tracking corrections in electrons

The performance of the track reconstruction algorithm for electrons depends on the precise simulation of the bremsstrahlung emission. To reduce any possible difference between data and simulation, the efficiency of reconstructing an electron track as *long* in simulation (see Sec. 3.2.1.5) is thus corrected in a data-driven way [146]. The track reconstruction efficiencies are obtained from a sample of $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ decays candidates using a *tag-and-probe* approach, where the tag candidates are selected exploiting the fully reconstructed information of the kaon and one of the electrons, while only the information of the VELO is used to reconstruct the second (probe) electron.[3] The tracking

---

[2] As already discussed in Sec. 3.2.1.6, electrons that are compatible with having emitted a photon detected in the ECAL are more likely to be identified as electrons.

[3] For this reason, this technically corresponds to a correction to the efficiency that a reconstructed electron VELO track is reconstructed as long, $\epsilon_{\rm TRK}(long|VELO)$.



(a)

(b)

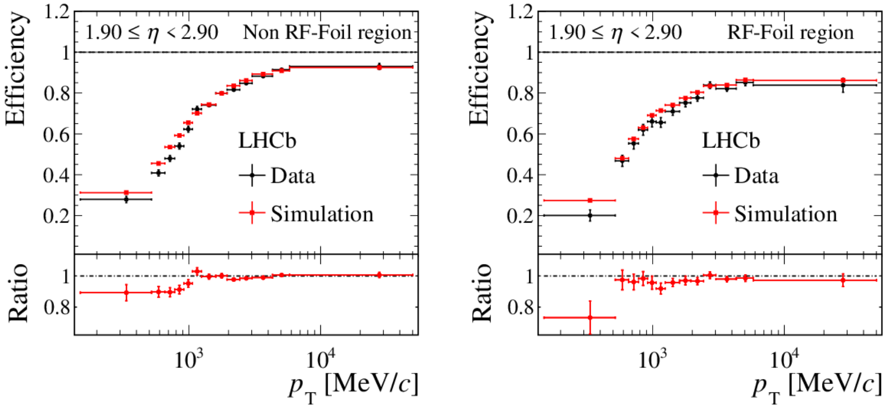
Figure 6.3: Comparison between the efficiency in data and simulation for the PID selection of electron candidates without (a) or with (b) a matching photon in the ECAL, as a function of the particle transverse momentum, for simulated events of $B^0 \to K^{*0}e^+e^-$ decays corresponding to the data-taking conditions of Run 2p1.

efficiencies are then obtained, in data, by performing a fit to the selected candidates for which a corresponding *long* track for the probe electron was found or not. The efficiencies are parametrized in bins of the transverse momentum, pseudorapidity and the azimuthal angle of the electron in the detector, ϕ. This last variable is needed to distinguish between the regions of the VELO where the RF-foil is installed (see Sec. 3.2.1.2), corresponding to an additional 10% radiation length for electrons and a different tracking performance. In simulation, the efficiencies $\epsilon^{sim}_{\mathrm{TRK}}$ are obtained in the same bins of data by means of a cut-and-count approach. Figure 6.4 shows the tracking efficiency in data and simulation, together with the corresponding ratio, as a function of transverse momentum for a specific bin of pseudorapidity. The impact of the presence or absence of the RF-foil for the efficiency can be seen comparing Figure 6.4(a) and (b).

The correction weight w^t_{TRK} for the electron track t corresponds to:

$$w^t_{\mathrm{TRK}} = \frac{\epsilon^{data}_{\mathrm{TRK}}(p^t_T, \eta^t, \phi^t; long|\mathrm{VELO})}{\epsilon^{sim}_{\mathrm{TRK}}(p^t_T, \eta^t, \phi^t; long|\mathrm{VELO})}, \tag{6.3}$$

while the total correction for a decay with two electrons in its final state corresponds to the product of the separate track weights $w_{\mathrm{TRK}} = w^{e^+}_{\mathrm{TRK}} \cdot w^{e^-}_{\mathrm{TRK}}$. No correction is considered for kaons, pions and muons. For the first two species, the same exact correction should be applied in decays that have muons and electrons in the final state and its impact is expected to cancel out in LFU observables. Tracking corrections are neglected for muons as they are expected to be quite close to unity in the region of phase space considered, and their impact is expected to be small.



(a)

(b)

Figure 6.4: (*a*) Measured electron efficiency as a function of the electron transverse momentum p_T, in a specific bin of pseudorapidity for electrons that do not travel parallel to the RF-foil. The ratios between the efficiencies in data and simulation are shown below. The analogous plot for electrons that travel parallel to the RF-foil is shown in (*b*) [146].

6.4 Trigger corrections

The trigger efficiency, *i.e.* the efficiency of triggering the signal candidates that have been reconstructed by the detector, is not well described in the simulation. To reduce the differences between data and simulation, a correction weight is assigned for the L0 and HLT trigger to each simulated event, referred to as w_{L0} and w_{HLT}, respectively. The correction is obtained separately for the different trigger categories L0I and L0L! and for electrons and muons. The weights are defined as data-over-simulation ratios of trigger efficiencies and, similarly to the PID corrections described for electrons in Sec 6.2, their aim is to ensure that the efficiency of the simulated sample matches the one observed in data, without losing the correlation with the remaining variables in the event.

Due to the fact that a recorded event must have satisfied at least one of the requirements of the LHCb trigger, normally it is not possible to study in data the absolute or *direct* efficiency of a trigger line. Thus, trigger efficiencies in data (ϵ_{data}) and simulation (ϵ_{MC}) are estimated using a *tag-and-probe* approach, based on the following expressions:

$$\epsilon^{\mathrm{data}}_{\mathrm{probe}} = \left(\frac{N_{\mathrm{tag\&probe}}}{N_{\mathrm{tag}}}\right)^{\mathrm{data}} \quad \text{and} \quad \epsilon^{\mathrm{MC}}_{\mathrm{probe}} = \left(\frac{N_{\mathrm{tag\&probe}}}{N_{\mathrm{tag}}}\right)^{\mathrm{MC}}. \tag{6.4}$$

In Eq. 6.4, N_{tag} corresponds to the number of signal candidates that have been observed after the *tag* selection, usually corresponding to a trigger selection that is not correlated with the line under study, while $N_{\mathrm{tag\&probe}}$ corresponds to the number of signal events that passed both the tag and the trigger selection of interest, here referred to as *probe*. The tag selection is chosen to be as uncorrelated as possible with respect the probe selection, in order to minimize any possible *tag bias*. The size of this bias can be assessed by measuring the efficiency of a trigger line using several tag selections.

In the following, data and simulation efficiencies for the different trigger selections used in the analysis are considered and the corresponding w_{L0} and w_{HLT} correction weights for the L0I and L0L! trigger categories discussed. For the L0M trigger selection, the *nominal* and *prior* trigger weights are computed, as already discussed in Sec. 6.1.

6.4.1 L0 trigger corrections

The L0 efficiencies are calibrated separately for each year of data-taking using the resonant modes $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ and $B^+ \to K^+ J/\psi(\to e^+e^-)$ after the full selection chain is applied (see Sec. 5.2), except for the L0 trigger requirements. Pure signal samples can be obtained from data without the use of the *sPlot* technique by applying a cut of ±45 MeV around the J/ψ constrained B^+ invariant mass. The simulated events are additionally truth-matched, as described in Sec. 5.1.2. Only the efficiencies for the inclusive categories L0I, L0E and L0M are computed, since the weights for the exclusive categories can be obtained by a proper combination of these efficiencies (see Sec. 6.4.1.4).

6.4.1.1 L0I calibration histograms

The probability that an event triggers independently of the specific decay of interest is studied as a function of the transverse momentum of the reconstructed B^0 candidate and `nTracks`. A dependency on these variables is expected, since the momentum of the reconstructed B^0 meson is correlated with the transverse momentum of the hadronization products of the opposite b-quark, and events with higher multiplicity have a higher probability to be triggered independently of the signal.

The efficiency is measured on events selected with the following three tag selections:

- **Lepton tag**, corresponding to events that have been triggered by one of the two leptons in the final state;
- **Hadron tag**, corresponding to events that have been triggered by a single hadron;
- **Combined tag**, events that have been triggered by the Lepton or Hadron tag.

Figure 6.5(a) shows the trigger efficiencies for L0I as a function of $p_T(B^+)$ and integrated in `nTracks`, as obtained from the muon samples for data (left)

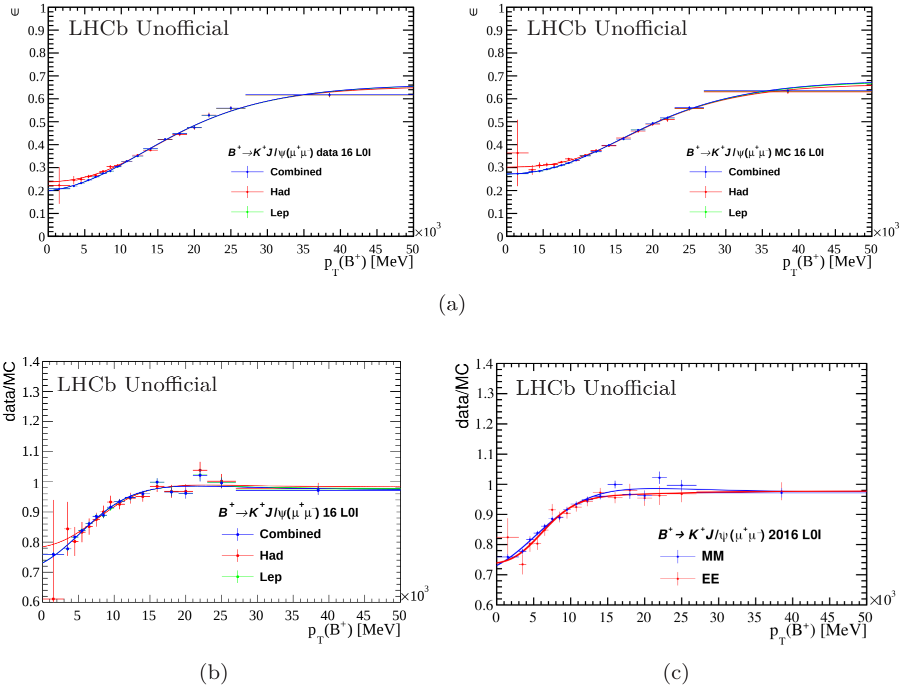


Figure 6.5: (a) Trigger efficiencies in $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ data (left) and simulated (right) candidates in 2016 for the L0I trigger category. The efficiency is shown as a function of $p_T(B^+)$ and integrated in `nTracks`. The parametrization obtained from a fit using the functional form introduced in Eq. 6.5 is overlaid. (b) Data-over-simulation ratio of the efficiencies shown in (a). (c) Comparison between ratio of efficiencies obtained for L0I in $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ and $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ decays for 2016. A good compatibility between the two is observed.

and simulation (right), and for the three different tag selections considered. Since a good agreement between the Lepton and Hadron tag is found, the default choice is to use the Combined tag to take advantage of the increased statistics. To reduce the dependency on the chosen binning, the trigger turn-on curves are fitted using a combination of an error function and a Gompertz function:

$$f_1 \cdot \left(1 + \mathrm{erf}\left(\frac{x-t}{\sqrt{2}\cdot\sigma_1}\right)\right) + f_2 \cdot s^{\exp\left(\frac{-(x-t)}{\sqrt{2}\cdot\sigma_2}\right)} + a\,, \tag{6.5}$$

where $f_{1,2}$ are the fractions of the two folded functions, $\sigma_{1,2}$ the resolutions, t the trigger threshold, s the skewness of the Gompertz function and a is a constant offset. The result of this fit can be seen in the solid curves in Figure 6.5(a) and in the corresponding data-over-simulation efficiency ratio in Figure 6.5(b). Since the L0I efficiency is expected to be independent of the signal final state, the corrections for electrons and muons are expected to be the same. Figure 6.5(c) shows that indeed the two corrections are compatible. It is therefore possible to take advantage of the clean and high statistics of muons sample and use this set of corrections also for the electron samples.

6.4.1.2 Single muon TOS calibration histogram

The probability that a muon fires the `L0Muon` trigger line depends on its transverse momentum p_T. For this reason, its response is parametrized as a function of $p_T(\mu)$ in three regions of pseudorapidity.

The efficiency of the L0M trigger line is studied in data and simulation using $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ candidates using three different tag selection:

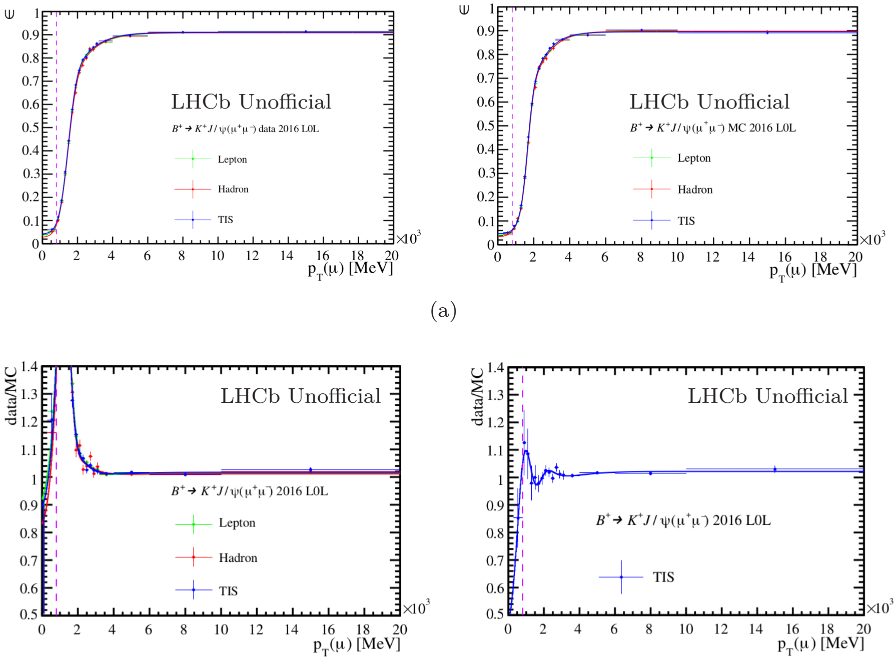
- **TIS tag**, corresponding to events that are triggered by a hadron or an electron not associated to the reconstructed signal;
- **Hadron tag**, corresponding to events triggered by a signal hadron;
- **Lepton tag**, corresponding to events triggered by the other muon from the decay of the J/ψ.

Also in this case, a Gompertz function is used to parametrize the turn-on curves of the obtained efficiencies. Figure 6.6(a) shows the efficiency for data (left) and simulation (right) for the three different tag selections considered in the central bin of pseudorapidity for the year 2016. The ratio of the two is shown in Figure 6.6(b). A good agreement between the ratios for the three different tag selections considered is found. Due to its higher statistics, the TIS tag is chosen as default in the following.

The corrections just presented are part of the *nominal* correction chain. The same strategy is used to obtain the efficiency ratios in the *prior* correction chain; the main difference between the two is a better alignment between the `TCK`s of data and simulation. Figure 6.6(c) shows the efficiency ratio obtained for the *prior* correction chain in the same bin of pseudorapidity shown in Figure 6.6(b). As it can be seen, the ratio obtained for the *prior* correction is much flatter than the nominal one, as expected.

6.4.1.3 Single electron TOS calibration histogram

The probability that an electron triggers an event depends on the region of the ECAL and on the amount of transverse energy deposited within 2×2 cells neighbouring the extrapolated trajectory. For this reason, the efficiency correction for the `L0Electron` trigger line is parametrized as a function of the region of the ECAL(inner, middle or outer) and of the transverse energy of the electron in the point where the electron track intersects the calorimeter, $E_T^{\rm L0}$. This last quantity represents an approximation for the transverse energy deposited in the calorimeter cells that caused the trigger to fire, since this quantity is not available for the offline analysis. The trigger efficiency is measured on data and



(a)

(b)

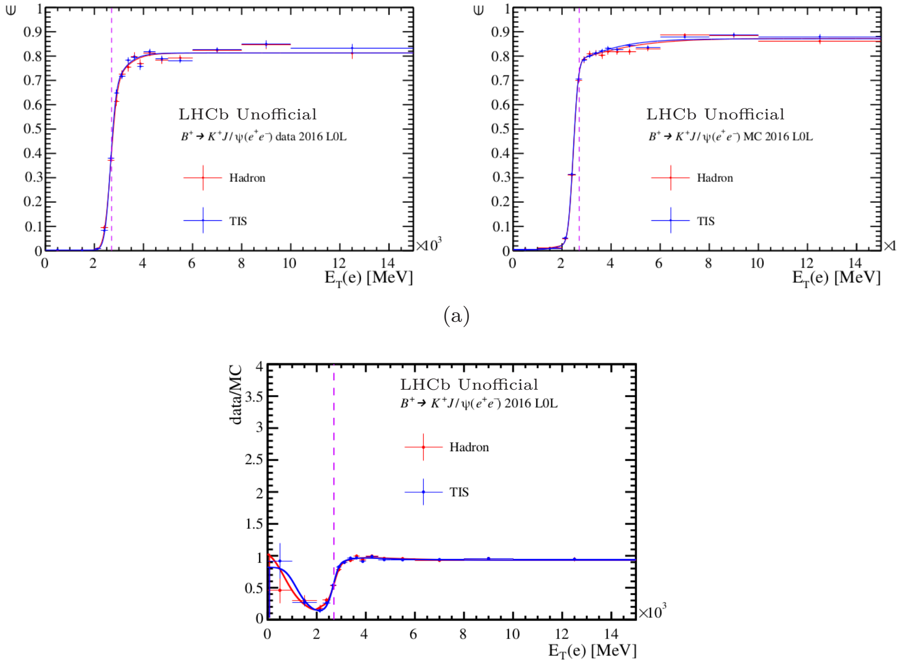
(c)

Figure 6.6: (a) Trigger efficiencies in $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ data (left) and simulated (right) candidates in 2016 for the L0M trigger category. The efficiency is shown as a function of $p_T(\mu)$ for the central bin in $\eta(\mu)$. The parametrization obtained from a fit using the functional form introduced in Eq. 6.5 is overlaid. (b) Data-over-simulation ratio of the efficiencies shown in (a). (c) Ratio of efficiencies obtained for the same central bin of $\eta(\mu)$ in the *prior* chain of corrections. Due to the improved alignment of `TCKs` between data and simulation, the ratio is much closer to unity than in (b), as expected.

simulation candidates of $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ decays with two different tags selections:

- **TIS tag**, corresponding to events that are triggered by a hadron or a muon not associated to the reconstructed signal;
- **Hadron tag**, corresponding to events triggered by a signal hadron.

A Gompertz function is used to parametrize the turn-on curves of the obtained efficiencies in the different regions of the electromagnetic calorimeter. Figure 6.7(a) shows the tag-and-probe efficiency computed for data (left) and simulation (right) in the middle region of the calorimeter for the year 2016. Figure 6.7(b) shows the ratio between the efficiencies obtained in data and simulation. A good agreement between the ratios obtained with different tag selections is observed. Due to the large statistics available, the TIS tag is chosen as default. Above the vertical dashed line, indicating the trigger threshold for 2016, the simulation agrees well with data. Below or near this line, the agreement becomes worse. There are two main reasons for this effect: a poorly simulated response of the ECAL and a dataset collected using different `TCKs`, corresponding to a mixture of different trigger thresholds in data. The additional threshold cut included in the definition of L0E! reduces the impact that such mismodellings have in the final result.



(a)

(b)

Figure 6.7: (a) Trigger efficiencies in $B^+ \to K^+ J/\psi(\to e^+e^-)$ data (left) and simulated (right) candidates in 2016 for the L0E trigger category. The efficiency is shown as a function of $E_T^{\rm L0}$ for the middle region of the ECAL. The parametrization obtained from a fit using the functional form introduced in Eq. 6.5 is overlaid. (b) Data-over-simulation ratio of the efficiencies shown in (a).

6.4.1.4 L0 correction weights

Using the data-over-simulation efficiency ratios just described, it is possible to assign event-by-event weights to the simulation samples depending on the classification of the L0 trigger response, *i.e.* $w_{\rm L0}^{\rm L0I}$, $w_{\rm L0}^{\rm L0M!}$ and $w_{\rm L0}^{\rm L0E!}$.

L0I weights

Simulated L0I events, *i.e.* events triggered by particles not reconstructed in the signal candidate, are corrected with the weight

$$w_{L0}^{\mathrm{L0I}} = \frac{\epsilon_{\mathrm{L0I}}^{data}(B^0 p_T)}{\epsilon_{\mathrm{L0I}}^{sim}(B^0 p_T)}\,. \tag{6.6}$$

This weight corresponds exactly to the ratio of efficiencies shown in Figure 6.5(b).

L0L! weights

Simulated L0L! events, *i.e.* events where at least one of the leptons in the signal candidate triggered but no muon, electron or hadron in the underlying event fired the corresponding trigger, are corrected with the weight

$$w_{L0}^{\mathrm{L0L!}} = \frac{(1-\epsilon_{\mathrm{L0I}}^{data}(B^0 p_T))}{(1-\epsilon_{\mathrm{L0I}}^{sim}(B^0 p_T))} \cdot \frac{(1-(1-\epsilon_{\ell TOS}^{data}(\ell^+))\cdot(1-\epsilon_{\ell TOS}^{data}(\ell^-)))}{(1-(1-\epsilon_{\ell TOS}^{sim}(\ell^+))\cdot(1-\epsilon_{\ell TOS}^{sim}(\ell^-)))}, \tag{6.7}$$

where the first factor corrects the efficiency of not having passed the L0I trigger selection, while the second one corrects the efficiency of having triggered on at least one of the final state leptons. The factorization of the different weights, used here to build $w_{\mathrm{L0}}^{\mathrm{L0E!}}$, is only exact if the trigger efficiencies of the two leptons are independent of each other and independent of L0I.

6.4.2 HLT trigger corrections

The HLT efficiencies are calibrated separately for each year of data taking and for the different trigger categories L0I, L0M! and L0E!, using the resonant modes $B^+ \to K^+ J/\psi(\to e^+e^-)$ and $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ after the full selection chain, with the exclusion of the HLT and the MVA selection. Due to a non-negligible contribution of combinatorial background under the signal, the calibration histograms for ϵ_{data} are obtained by performing separate fits to the *probe* and *tag* data samples as a function of `nTracks`. This choice is justified by the fact that Run 1 data shows a trend in this variable, not well reproduced in simulation, that needs to be corrected for. The use of additional variables, *e.g.* the transverse momentum of the B^+ meson, have been considered, but are only used in the *prior* correction. The efficiency histograms ϵ_{data} are then defined by the ratio of the extracted yields, with a binning scheme chosen to have a similar number of events in each bin. On the other hand, ϵ_{sim} is evaluated in the same bins of `nTracks` by looking at the ratio of the weighted sum of events between the *probe* and *tag* simulated samples. The weights used at this stage include the PID, kinematics and multiplicity, and L0 correction weights. The same selections are applied to data and simulation except for the tag selection, which corresponds to `HLTPHYS` and `HLTOR`, respectively, defined in Table 6.1. This choice is a compromise between using the `HLTOR` selection in data, that would greatly reduce the statistics available, and using `HLT{1,2}PHYS` in simulation, that would include a relevant fraction of HLT lines that are ignored in data.

Table 6.1: Tag categories definition used for the HLT correction for electron and muon mode in data (`HLTPHYS`) and in simulated events (`HLTOR`). The `HLT1OR` and `HLT2OR` categories are defined as the logical `OR` of their listed single lines.

Label	Run 1	15	16	17&18
HLT1PHYS	HLT1_PHYS_TIS	HLT1_PHYS_TIS	HLT1_PHYS_TIS	HLT1_PHYS_TIS
HLT2PHYS	HLT2_PHYS_TIS	HLT2_PHYS_TIS	HLT2_PHYS_TIS	HLT2_PHYS_TIS
HLTPHYS	HLT1PHYS & HLT2PHYS			
HLT1OR	HLT1TrackAllL0_TIS HLT1Track{Muon,DiMuon}_TIS	HLT1Track{MVA,Muon}_TIS HLT1TwoTrack_TIS	HLT1Track{MVA,Muon,MuonMVA}_TIS HLT1TwoTrack_TIS	HLT1Track{MVA,Muon,MuonMVA}_TIS HLT1TwoTrack_TIS
HLT2OR	HLT2Topo{2,3,4}BodyBBDT_TIS HLT2TopoMu{2,3,4}BodyBBDT_TIS HLT2TopoE{2,3,4}BodyBBDT_TIS HLT2DiMuonDetached_TIS	HLT2Topo{2,3,4}Body_TIS HLT2TopoMu2,3,4Body_TIS Hlt2DiMuonDetachedHeavy_TIS	HLT2Topo{2,3,4}Body_TIS HLT2TopoMu{2,3,4}Body_TIS HLT2TopoMuMu{2,3,4}Body_TIS Hlt2DiMuonDetachedHeavy_TIS HLT2TopoE{2,3,4}Body_TIS HLT2TopoEE{2,3,4}Body_TIS	HLT2Topo2,3,4,Body_TIS HLT2TopoMu2,3,4Body_TIS HLT2TopoMuMu2,3,4Body_TIS Hlt2DiMuonDetachedHeavy_TIS HLT2TopoE2,3,4Body_TIS HLT2TopoEE2,3,4Body_TIS
HLTOR	HLT1OR & HLT2OR			

Additionally, it is found that ϵ_{data} is compatible between the inclusive (L0L) and exclusive (L0L!) categories. Due to the larger statistics of the former, it is then preferred to correct the exclusive samples using the corrections obtained from the inclusive ones. The ratios between data and simulation are shown in Fig. 6.8 for the L0 inclusive categories L0I and L0L.

Similarly to the L0, a *prior* set of corrections is computed for the HLT trigger to improve the agreement between data and simulation in the L0M sample, used to obtain the generator level correction to the kinematics. The same procedure described above applies with two main differences: the corrections are performed as a function of `nTracks` and $p_T(B)$, thanks to the higher statistics available, and the simulation is weighted considering only the corrections obtained for PID and L0.

6.4.2.1 HLT correction weights

The L0I events in the simulation are corrected with the weight

$$w_{HLT}^{\text{L0I}} = \frac{\epsilon_{\text{L0I},data}^{HLT}(\texttt{nTracks})}{\epsilon_{\text{L0I},sim}^{HLT}(\texttt{nTracks})}, \tag{6.8}$$

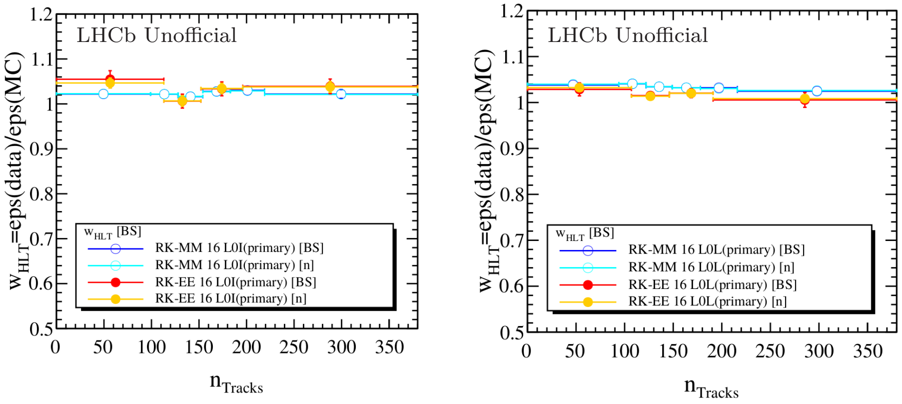
while the L0L! events are corrected with the weight

$$w_{HLT}^{\text{L0L!}} = \frac{\epsilon_{\text{L0L},data}^{HLT}(\texttt{nTracks})}{\epsilon_{\text{L0L},sim}^{HLT}(\texttt{nTracks})}. \tag{6.9}$$

6.5 Kinematic, multiplicity and reconstruction corrections

A good agreement between data and simulation in the variables that are used to select, describe or correct the signal is fundamental. Two set of corrections weights are discussed in the following:

- *generator level* correction weights, $w_{\text{Mult\&Kin}}$, whose aim is to correct the kinematics and event multiplicity of the B^0 meson generated in simulation;



(a) (b)

Figure 6.8: Nominal chain HLT efficiency correction map obtained using the $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ and $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ modes selected in the L0I (a) and L0L (b) inclusive category for the year 2016. The label [BS] highlights the results obtained enabling the bootstrapping of data fits and efficiency weight ratios in simulation, while [n] labels the nominal efficiency ratio between data and simulation.

- *reconstruction* correction weights, w_{Reco}, whose aim is to correct for the reconstructed χ^2_{IP} distribution of the B^0 meson and J/ψ resonance, together with any remaining difference in reconstructed kinematics of the B^0 meson.

Both corrections are obtained training Boosted Decision Trees (BDT) [147, 148] to reweight the simulated samples to match the selected data samples in multiple dimensions, ensuring that correlations between all input variables are properly taken into account. The algorithm used is the GBReweighter from the hep_ml library [149]. A *k-fold* approach is used also in this situation, with the number of folds used equal to four.

6.5.1 Generator level corrections

The first set of corrections focuses on the momenta, transverse momenta, pseudorapidity of the B^0 meson and `nTracks`. The last one is of particular importance, since the occupancy of the detector is known to be poorly described in simulation, mainly due to low momentum particles, backsplashes from particle showers, and secondary interactions. It is important to point out that different occupancy proxies are not all coherent between each other in data and simulation: in general correcting one does not guarantee a good agreement of the others. This is particularly relevant since both `nTracks` and `nSPDHits` are used in this analysis. The current approach is to correct `nTracks`, since this variable is used in the PID corrections, and to assign a systematics due to the remaining discrepancy observed between data and simulation in `nSPDHits`. It is however important to point out that the discrepancy observed in `nSPDHits` is coherent between rare and control mode and for muons and electrons, suggesting a minor impact on LFU observables. The data and simulation events used correspond to reconstructed $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ candidates, where the full selection except for the combinatorial MVA is used. To improve the agreement between data and simulation, only the data whose `TCK`s are matched in the simulated sample have been used. Additionally, a *prior* chain of trigger corrections is applied to the simulation. The only events used for data and simulation are the ones that correspond to the trigger category L0M. The BDT is trained to match the background-subtracted data and simulation after all these alignments, and any remaining difference is interpreted as a difference in the generated properties of the events in simulation, rather than a reconstruction effect. For this reason, when correcting the full simulation, the correction is applied to both electrons and muons based on the generated values of $p(B)$, $p_T(B)$, $\eta(B)$ and `nTracks`, rather than on their reconstructed quantities.

6.5.2 Reconstruction corrections

The second set of corrections focuses on some of the reconstructed kinematics variables of the B^0 meson, together with the χ^2_{IP} of the B^0 meson and J/ψ resonance. A good agreement between these features in data and simulation is key, since these variables enter the selections and the MVA used to reduce the combinatorial background. Due to the reconstruction differences between electrons and muons and between different L0 trigger categories, each correction is obtained separately. Data and simulation samples of $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ and $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ candidates are considered after the full selection and split in L0I and L0L! categories. Differently from the *generator level corrections*, clean samples of signal from data cannot be obtained using ${}_s\mathcal{P}lot$, due to the correlation of some of the reconstructed variables with the invariant mass of the $K\pi\ell^+\ell^-$ system. Instead, a tight cut on the MVA and an additional cut of $60\,\text{MeV}/c^2$ around the nominal mass of the B^0 meson in the $m_{DTF}^{J/\psi}(K\pi\ell^+\ell^-)$ is performed. The simulated samples have been selected accordingly, taking into account the whole chain of weights, from w_{PID} to w_{HLT}. Figure 6.9 shows in black the comparison between the distributions obtained from $B^+ \to K^+ J/\psi\,(\to \ell^+\ell^-)$ data candidates using ${}_s\mathcal{P}lot$ technique, while in blue and red are shown the distributions after and before the full chain of corrections is applied. As can be seen, a remarkable improvement in the agreement between data and simulation is achieved.

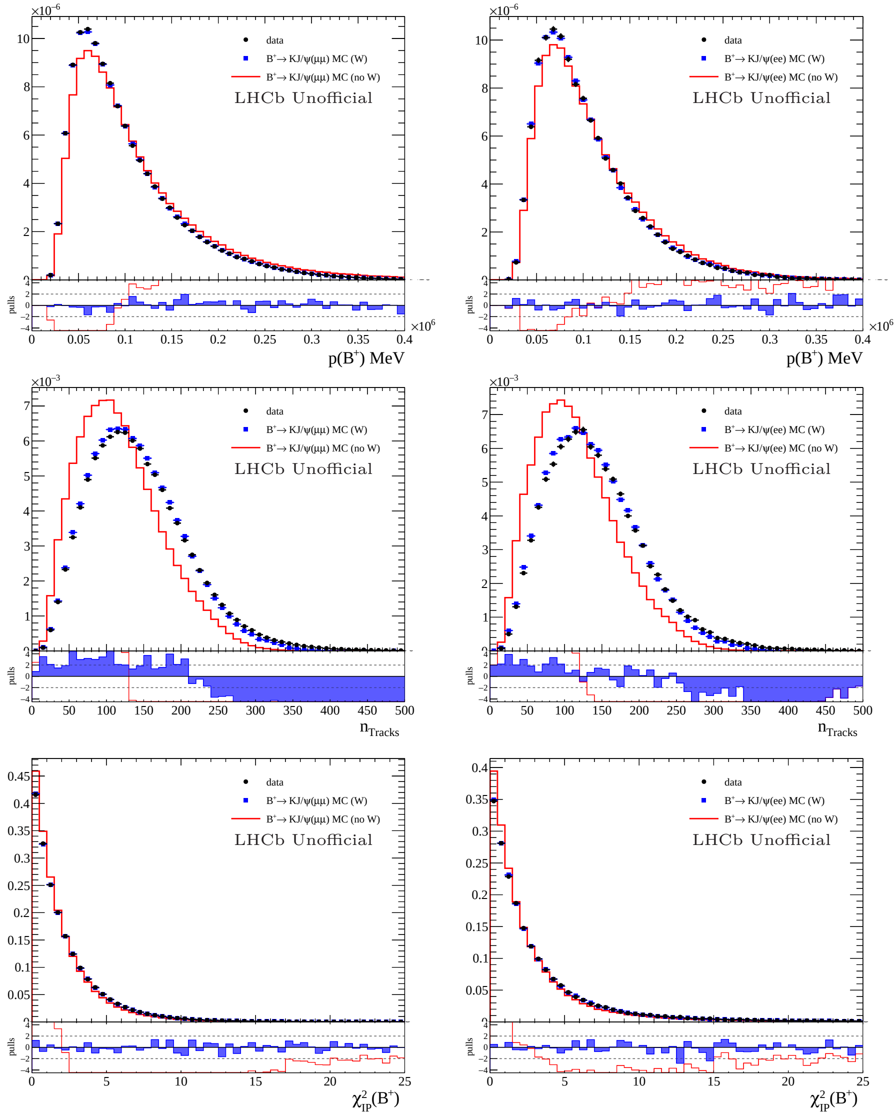


Figure 6.9: Data/simulation comparison plots for 2016 L0M! (left) and L0E! (right) $B^+ \to K^+ J/\psi (\to \ell^+\ell^-)$ samples. The variables shown are some the input variables used in the reweighter.

6.6 Invariant mass resolution correction

In simulation, the momentum resolution of the reconstructed electrons is better than in collision data. This difference in resolution comes from two main sources. First, differences in the material budget of the detector between data and simulation impact on the amount of bremsstrahlung emitted, which in turn impacts on the momentum resolution. Second, the poor description of the occupancy of the calorimeter response impacts on the bremsstrahlung recovery, which influences the mass shape. Therefore, any invariant mass built by using the dielectron system, e.g. $m(e^+e^-)$ or $m(K\pi e^+e^-)$ in $B^0 \to K^{*0} J/\psi(\to e^+e^-)$ decays, is narrower in simulation with respect to data, potentially biasing the efficiency estimation and the acceptance parametrization of the rare and control mode.

To study the resolution in simulation and data, a fit to the invariant mass of the dielectron system is performed on $B^0 \to K^{*0} J/\psi(\to e^+e^-)$ candidates. The samples are fully selected, except for the q^2 region and combinatorial MVA cut. The simulation is truth-matched and the full set of corrections up to w_{Reco} is included. Since in addition to the constrained dilepton invariant mass squared, q_c^2, several cross-checks are performed with the standard q^2 variable, the resolution is studied both in $m(e^+e^-) = \sqrt{q^2}$ and $m(e^+e^-)_{B^0} = \sqrt{q_c^2}$.

Fits to simulation are performed separately for different years, trigger and bremsstrahlung categories. The signal shapes are parametrized with a double-sided Crystal Ball [150](see Eq. 8.6) and fitted in the ranges corresponding to [2500, 3150] MeV/c^2 for $m(e^+e^-)_{B^0}$ and [2300,3500] MeV/c^2 for $m(e^+e^-)$. An additional requirement on the invariant $K\pi e^+e^-$ mass with the dilepton mass constrained to the J/ψ mass, $m_{DTF}^{J/\psi} \in [5150, 5680]\,\text{MeV}/c^2$, is applied in the fits to $m(e^+e^-)$ to reduce the partially reconstructed background.

The same shape is then used for the signal in the fit to data, allowing only three parameters to float:

- $\Delta\mu$, representing a shift between the peak position in data and simulation and defined as $\mu^{data} = \mu^{sim} + \Delta\mu$;
- s_σ, representing a scaling factor in the width of the Gaussian core between data and simulation and defined as $\sigma^{data} = s_\sigma \cdot \sigma^{sim}$;
- s_α, representing a scaling factor that controls the difference of the right polynomial tail in data and simulation and defined as $\alpha_R^{data} = s_\alpha \cdot \alpha_R^{sim}$ and $n_R^{data} = n_R^{sim}/s_\alpha$.

Since the right-hand tail plays no role in the 0γ bremsstrahlung category, s_α is not considered in those data fits. The residual combinatorial component is modelled using an exponential function. An example of the data fits obtained is shown on the top and bottom left part of Fig. 6.10 for the L0I trigger category in the year 2016 for $m(e^+e^-)$ and $m(e^+e^-)_{B^0}$, respectively.

In each year and bremsstrahlung category, the averages of s_σ, $\Delta\mu$ and μ_{sim} among the two trigger categories, denoted $\overline{s_\sigma}$, $\overline{\Delta\mu}$ and $\overline{\mu_{sim}}$, are used to approximately correct the simulation, by producing a smeared variable $m^{smrd}(e^+e^-)$ ($m^{smrd}(e^+e^-)_{B^0}$) defined as:

$$m^{smrd} = m_{preFSR}^{true} + \overline{s_\sigma} \cdot (m^{reco} - m_{preFSR}^{true}) + (1 - \overline{s_\sigma}) \cdot (\overline{\mu_{sim}} - M_{J/\psi}) + \overline{\Delta\mu}\,. \quad (6.10)$$

The m^{true}_{preFSR} is the true (generated) dilepton mass, evaluated from the generator level tuple by using the difference between the true kinematics of the parent B^0 meson and the true kinematics of the hadronic final states. To treat similarly all the reconstructed events independently of their truth-matching, m^{true}_{preFSR} is assigned to each event by matching the reconstructed and generated candidates. m^{reco} is the reconstructed mass in simulation and $M_{J/\psi}$ is the PDG value [129] of the J/ψ resonance mass. The first two terms in the formula represent a stretch in the m^{reco} distribution with respect to the point m^{true}_{preFSR} (this point will not move), the third term describes the shift needed to put back the stretched value of $\overline{\mu^{sim}}$ into its original position, and the last one describes the additional shift needed to have the stretched distribution centred in μ^{data}. If m^{reco} is described by a Gaussian with mean $\overline{\mu^{sim}}$ and standard deviation σ^{sim}, the transformed variable m^{smrd} will follow a Gaussian with mean $\overline{\mu^{sim}} + \Delta\mu$ and standard deviation $\sigma^{sim} \cdot \overline{s_\sigma}$.

The right plots in Fig. 6.10 show a comparison between the parametrization obtained in data (orange) and the shapes obtained from simulation before (black) and after (blue) the smearing introduced with Eq. 6.10. As it can be seen, the parametrization obtained from the smeared observables agrees better with the signal distribution directly obtained from data. This is true for all years and trigger categories, as the improved compatibility with zero (one) of the average value of $\Delta\mu$ (s_σ) after the correction shows in Tables 6.2 and 6.3. A small discrepancy remains however in the right tail of the distribution; this is expected since the smearing procedure of Eq. 6.10 does not impact the tail parameters. The systematic associated with this remaining difference between data and simulation is expected to be small.

Table 6.2: Values of the mass shift and sigma scale obtained from data fits to the invariant mass $m(e^+e^-)$, using signal shapes obtained from fits to simulated samples with (right) and without (left) resolution correction. The values reported, $<\Delta\mu>$ and $<s_\sigma>$, correspond to the mean and standard deviation of $\Delta\mu$ and s_σ obtained for the different years and trigger categories.

		$m(e^+e^-)$ w/o smearing		$m(e^+e^-)$ with smearing	
		Run1	Run2	Run1	Run2
0γ	$<\Delta\mu>$	-0.04 ± 2.12	-4.77 ± 1.14	0.20 ± 2.23	0.01 ± 0.66
	$<s_\sigma>$	1.06 ± 0.05	1.13 ± 0.05	1.00 ± 0.04	1.00 ± 0.02
1γ	$<\Delta\mu>$	-0.51 ± 3.77	-9.68 ± 3.95	-0.59 ± 3.45	-0.50 ± 2.69
	$<s_\sigma>$	1.13 ± 0.03	1.14 ± 0.04	1.01 ± 0.02	1.00 ± 0.03
2γ	$<\Delta\mu>$	3.41 ± 1.72	-10.85 ± 4.64	-0.11 ± 0.80	-0.23 ± 3.76
	$<s_\sigma>$	1.10 ± 0.05	1.10 ± 0.06	1.00 ± 0.03	1.00 ± 0.03

Table 6.3: Values of the mass shift and sigma scale obtained from data fits to the invariant mass $m(e^+e^-)_{B^0}$, using signal shapes obtained from fits to simulated samples with (right) and without (left) resolution correction. The values reported, $<\Delta\mu>$ and $<s_\sigma>$, correspond to the mean and standard deviation of $\Delta\mu$ and s_σ obtained for the different years and trigger categories.

		$m(e^+e^-)_{B^0}$ w/o smearing		$m(e^+e^-)_{B^0}$ with smearing	
		Run1	Run2	Run1	Run2
0γ	$<\Delta\mu>$	-0.70 ± 0.68	-0.99 ± 0.32	0.05 ± 0.67	-0.01 ± 0.29
	$<s_\sigma>$	1.05 ± 0.06	1.12 ± 0.05	1.00 ± 0.05	1.00 ± 0.03
1γ	$<\Delta\mu>$	-1.42 ± 0.56	-0.07 ± 0.30	-0.02 ± 0.27	-0.00 ± 0.10
	$<s_\sigma>$	1.08 ± 0.13	1.12 ± 0.03	1.02 ± 0.12	1.00 ± 0.03
2γ	$<\Delta\mu>$	-0.91 ± 0.29	-0.17 ± 0.31	-0.04 ± 0.28	0.01 ± 0.22
	$<s_\sigma>$	1.14 ± 0.08	1.10 ± 0.04	1.00 ± 0.05	1.00 ± 0.03

6.7 Impact of the correction chain on the distributions

The different correction steps discussed in this chapter can have a sizable impact on the distributions of the resonant and rare modes of interest. In the following, the variation in the variables $\cos\theta_K$, $\cos\theta_\ell$, ϕ, q^2 (q_c^2), $m(K\pi\mu\mu)$ ($m(K\pi ee)$) and $m(K\pi)$ due to the correction chain applied is studied in fully corrected and selected simulated samples of $B^0\to K^{*0}\mu^+\mu^-$ ($B^0\to K^{*0}e^+e^-$) decays. The most important features observed are summarized as follows:

- The use of the PID efficiencies from dedicated calibration samples, rather than performing a cut on the PID variables obtained from simulation, does not have a strong impact on the variables of interest. The most relevant effect can be seen in the $\cos\theta_K$ distribution of $B^0\to K^{*0}\mu^+\mu^-$ decays, as shown in Figure 6.11(a) for Run 1.
- The tracking corrections, applied only to the electron samples, have no impact except for a small shaping in $\cos\theta_\ell$ for $B^0\to K^{*0}e^+e^-$ decays in Run 2, as shown in Figure 6.11(b). No effect is noticeable in Run 1.
- The kinematic and multiplicity corrections affect similarly muons and electrons for the different runs and trigger categories. They tend to reduce the number of events for $\cos\theta_K\sim 1$ and $\cos\theta_\ell\sim -1$, as can be seen in Figure 6.11(c)-(d). These trends can be explained by the correlation that the fiducial cuts $p_T(K) > 250\,\mathrm{MeV}/c$ and $p_T(\pi) > 250\,\mathrm{MeV}/c$ introduce between the transverse momentum of the B^0 meson, $p_T(B)$, and $\cos\theta_K$. As Figure 6.13 shows, the $\cos\theta_K$ distribution distortion increases as the bin of $p_T(B)$ considered decreases, while no distortion is present if the fiducial cut is not applied. As a consequence, by using a reweighting scheme that decreases the importance of events with higher $p_T(B)$, as it is done for $w_{\mathrm{Mult\&Kin}}$, the impact that the fiducial cut has on the integrated $\cos\theta_K$

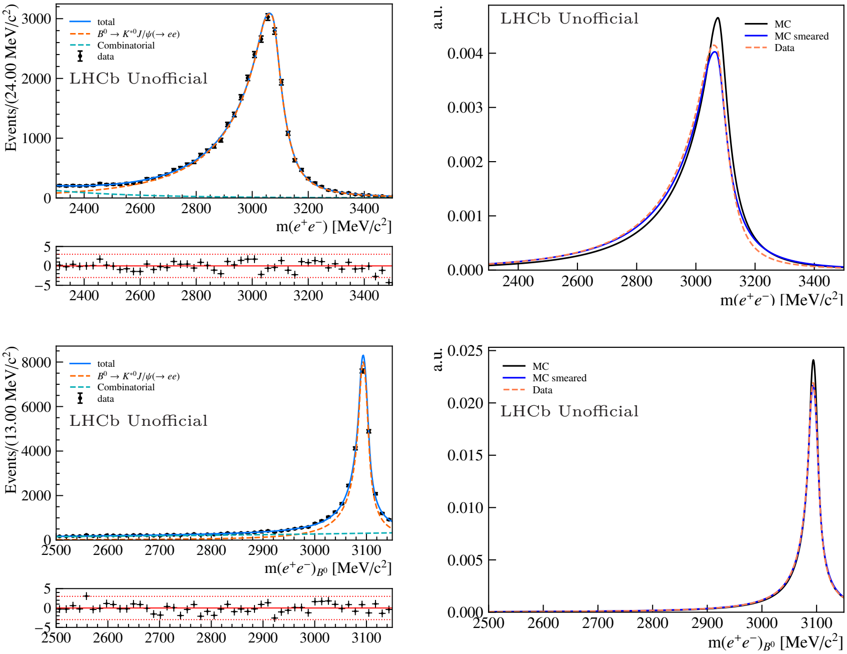
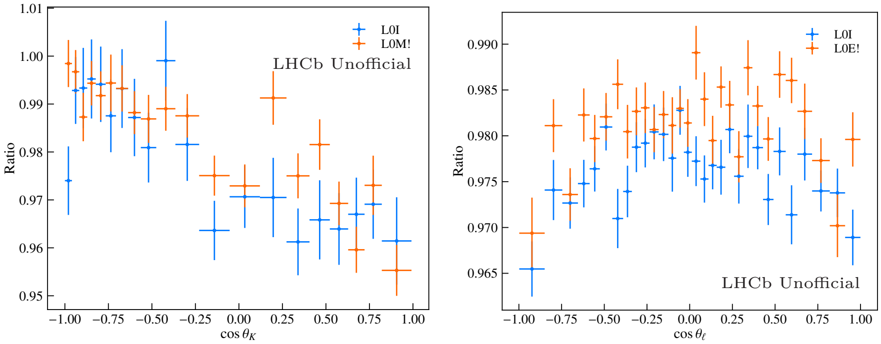


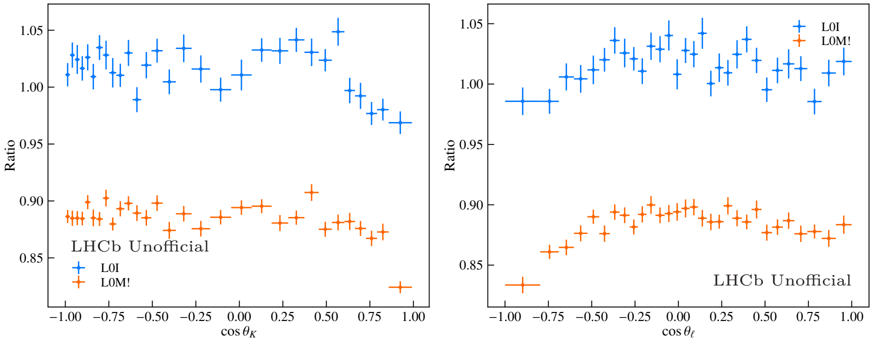
Figure 6.10: On the left are shown the results of fits performed to $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ candidates for the year 2016 in the L0I category for the variables $m(e^+e^-)$, on the top, and $m(e^+e^-)_{B^0}$, on the bottom. The sum of the fit results obtained separately for the different bremsstrahlung categories 0γ, 1γ and 2γ is shown here together with the joint dataset. On the right are shown the comparisons between signal shape obtained directly from the fits on the left (orange) and the shapes obtained from simulation before (black) and after (blue) the smearing correction is applied.

distribution increases. A similar interpretation holds for the behaviour observed when correcting for the L0 trigger.

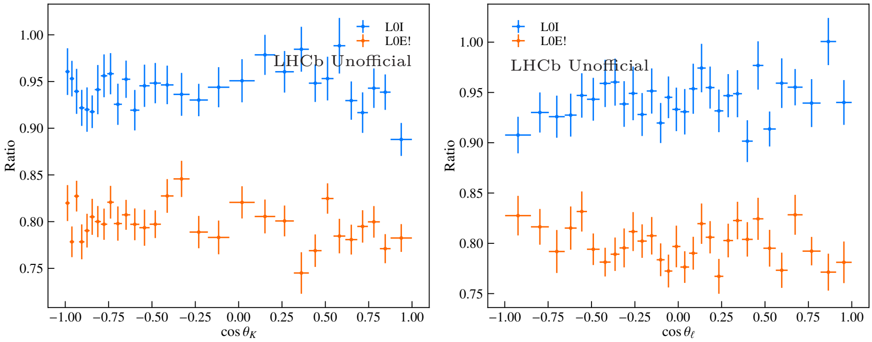
- The `L0` corrections seem to have the opposite effect for the L0M! and L0I categories in $B^0 \to K^{*0}\mu^+\mu^-$ decays (see Figure 6.12(a)): the first one is qualitatively similar to the impact of $w_{\text{Mult\&Kin}}$, while the second one increases the number of events in $\cos\theta_K \sim 1$ and $\cos\theta_\ell \sim -1$. A similar behaviour is seen in Figure 6.12(b) for L0I in the electron case, while no clear trend can be seen for L0E!.
- No clear trend is observed when considering the HLT corrections.
- The only trend visible when using the correction weight w_{Reco} is in $m(K\pi\ell\ell)$, as shown in Figure 6.12(c).



(a) Example of the impact of w_{PID} weights (b) Example of the impact of w_{TRK} weights

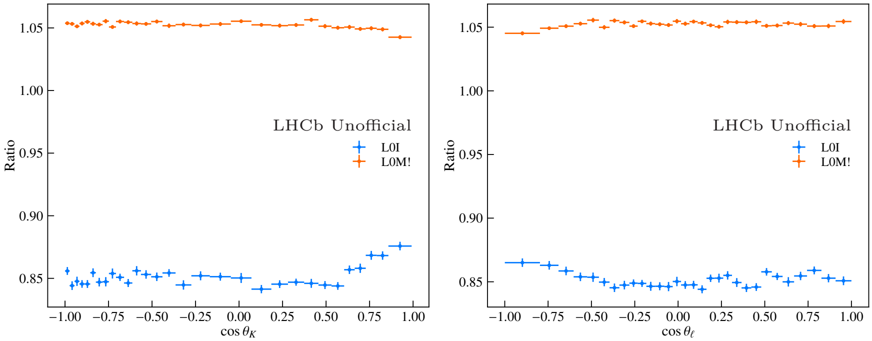


(c) Example of the impact of $w_{\text{Mult\&Kin}}$ corrections in $B^0 \to K^{*0}\mu^+\mu^-$ decays

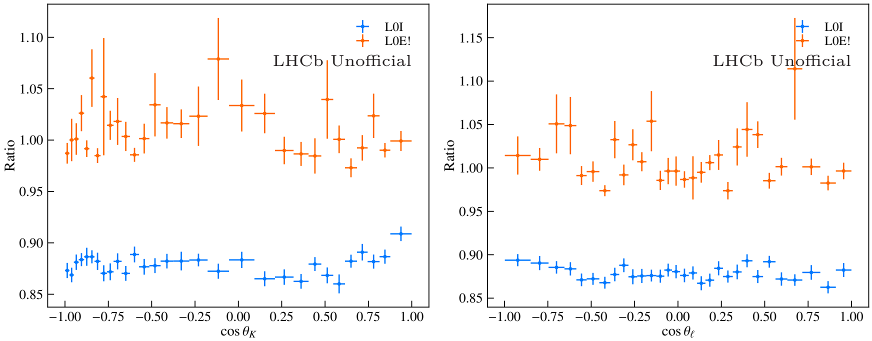


(d) Example of the impact of $w_{\text{Mult\&Kin}}$ corrections in $B^0 \to K^{*0}e^+e^-$ decays

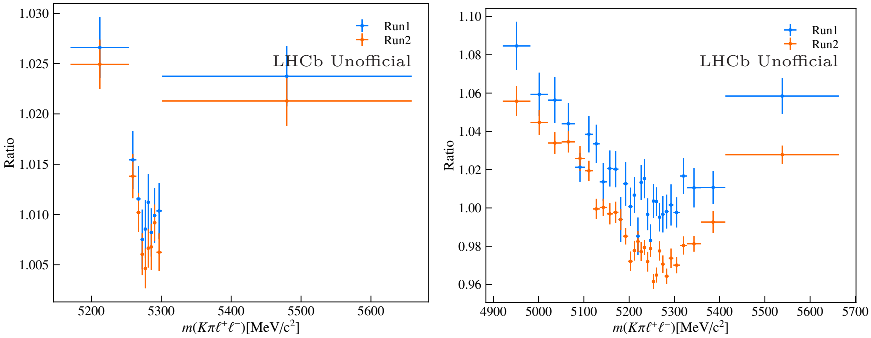
Figure 6.11: Collection of the most relevant trends in angles, q^2, $m(K\pi\ell\ell)$ and $m(K\pi)$ caused by the use of w_{PID}, w_{TRK} and $w_{\text{Mult\&Kin}}$ correction weights on simulated $B^0 \to K^{*0}\ell^+\ell^-$ decays.



(a) Example of the impact of w_{L0} corrections in $B^0\to K^{*0}\mu^+\mu^-$ decays



(b) Example of the impact of w_{L0} corrections in $B^0\to K^{*0}e^+e^-$ decays



(c) Example of the impact of w_{Reco} corrections in $B^0\to K^{*0}\mu^+\mu^-$ and $B^0\to K^{*0}e^+e^-$ decays

Figure 6.12: Collection of the most relevant trends in angles, q^2, $m(K\pi\ell\ell)$ and $m(K\pi)$ caused by the use of w_{L0} and w_{Reco} correction weights on simulated $B^0\to K^{*0}\ell^+\ell^-$ decays.

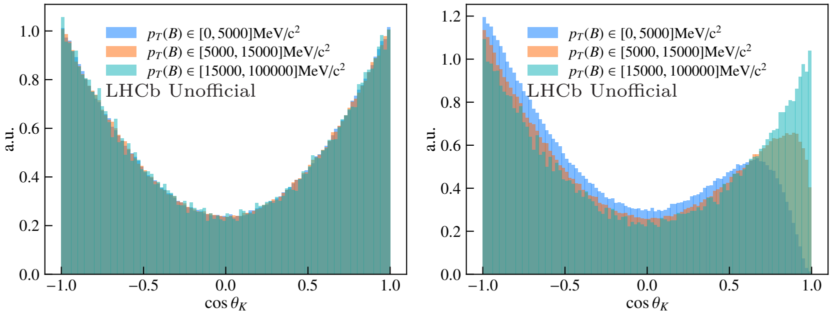


Figure 6.13: Variation of $\cos\theta_K$ in three bins of $p_T(B)$ with and without the fiducual cuts $p_T(K) > 250\,\text{MeV}/c$ and $p_T(\pi) > 250\,\text{MeV}/c$ applied. The sample used for this check corresponds to a private production of generator level MC for $B^0 \to K^{*0}e^+e^-$ in the central-q^2_c region.

When considering the overall impact of all the corrections steps discussed, the main effect seems to be a decreasing behaviour in $\cos\theta_K$ in $B^0 \to K^{*0}\mu^+\mu^-$ decays, common to L0I and L0M!, and a decrease of $\cos\theta_\ell$ in the region close to -1 for L0M! only. In the electron samples, probably due to the reduced statistics, no clear trend in visible after the full chain of corrections is applied.

Chapter 7

Efficiency

The selections described in Sec. 5.2 aim to isolate the decays of $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ from potential backgrounds. At the same time, they inevitably distort their distributions of interest, reducing the amount of B^0 decays available to study with respect to the ones that were originally produced. The description of these two effects is the aim of the following chapter. Firstly, the strategy for the determination of the absolute efficiencies of signal and background decays from simulation is discussed in Sec. 7.1. A special attention is given to the ratio of absolute efficiencies between $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ channels, called R_ε, used to impose a constraint on the observed yield of the signal in the final fit to data (see Section 10.3). Secondly, the impact of the selections in distributions of interest is studied and parametrized with a four-dimensional effective acceptance function. This function is multiplied together with the signal *pdf* in the amplitude fits to data (see Sec. 9.2.3 and Ch. 10).

7.1 Integrated and relative efficiency

The total efficiency of a decay (ϵ_{tot}) corresponds to the probability that an event generated by a *p*-*p* collision in LHCb is reconstructed and survives the full chain of selections described in Sec. 5.2. It can be measured from fully corrected and smeared simulation samples as:

$$\epsilon_{\text{tot}} = \varepsilon_{\text{geo}} \cdot \varepsilon_{\text{full}|\text{geo}}\,, \tag{7.1}$$

where

- ε_{geo} represents the efficiency of having all final state particles in the acceptance of the LHCb detector, corresponding to a polar angle θ between 10 and 400 mrad with respect to the axis of the detector. For $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ decays, this corresponds to $\varepsilon_{\text{geo}} \approx 15 - 17\%$, with a relative uncertainty between 0.2 and 0.4%.
- $\varepsilon_{\text{full}|\text{geo}}$, also referred to as *selection efficiency*, corresponds to the efficiency of applying the full selection chain, including the q^2 selection, fit mass range and MVA selection, for the events that are already in the LHCb acceptance.

Eq. 7.1 needs to be slightly modified when computing the total efficiencies of some of the simulated backgrounds, due to the use of an intermediate filtering step applied during the production. The expression becomes

$$\epsilon_{\text{tot}} = \varepsilon_{\text{geo}} \cdot \varepsilon_{\text{flt}|\text{geo}} \cdot \varepsilon_{\text{full}|\text{flt}}\,, \tag{7.2}$$

where $\varepsilon_{\text{flt}|\text{geo}}$ corresponds to the efficiency of the filtering step applied, given that the events are already in the LHCb acceptance. The use of filtering steps in the generation of samples has the advantage of reducing the amount of disk space used for storing the simulation, without loosing any of the events that pass the full final selection.

7.1.1 Selection efficiency

The selection efficiency $\varepsilon_{\text{full}|\text{geo}}$ ($\varepsilon_{\text{full}|\text{flt}}$ if a filtering step is present) is computed using the fully corrected and smeared simulated samples thanks to the following expression:

$$\varepsilon_{\text{full}|\text{geo}} = \frac{A}{B} \cdot \frac{C}{D} \ , \tag{7.3}$$

$$\begin{aligned}
A &= \sum_{w}^{\texttt{DecayTuple}} (w_{PID} \cdot w_{TRK} \cdot w_{L0} \cdot w_{HLT} \cdot w_{Mult\&Kin} \cdot w_{Reco})|Selection(A) \ , \\
B &= \sum_{w}^{\texttt{MCDecayTuple}} (w_{Mult\&Kin})|Selection(B) \ , \\
C &= \sum_{w}^{\texttt{DecayTuple}} (w_{PID} \cdot w_{TRK} \cdot w_{L0} \cdot w_{HLT} \cdot w_{Mult\&Kin})|Selection(NORM) \ , \\
D &= \sum_{w}^{\texttt{DecayTuple}} (w_{PID} \cdot w_{TRK} \cdot w_{L0} \cdot w_{HLT} \cdot w_{Mult\&Kin} \cdot w_{Reco})|Selection(NORM) \ ,
\end{aligned} \tag{7.4}$$

where:

- `DecayTuple` and `MCDecayTuple` represent the samples over which the sums are computed, respectively the fully reconstructed simulated samples and the generator level samples inside the LHCb acceptance before any reconstruction.
- w_i are the weights defined in Ch. 6, describing each of the corrections performed to improve the agreement between simulation and data. Notice that the weight $w_{Mult\&Kin}$ appears in B to normalize the expression in A, since $w_{Mult\&Kin}$ is not normalized by default.
- $Selection(A)$ is the full selection described in Sec. 5.2, except for the PID cuts of kaons, pions and muons, whose efficiency in simulation is taken into account with the use of w_{PID}.
- In order to estimate the full efficiency, $Selection(B)$ should correspond to no selection at all.[1] However, as the `nSPDHits` variable is badly modelled

---

[1] The geometric and filtering selections are however already applied to the `MCDecayTuple` sample.

in simulation, the effect of the corresponding cut is removed from $\varepsilon_{\mathrm{full|geo}}$ by defining $Selection(B)$ as `nSPDHits` >600(450) for RUN 1 (RUN 2). Ignoring the efficiency of this cut, introduces a bias in the estimation of the total absolute efficiency. However, since no much difference is expected in the `nSPDHits` distribution of the rare and resonant mode, it can be assumed safely that the main quantity of interest, the ratio between the two efficiencies, is unbiased.

- The cut used to normalize w_{Reco}, $Selection(NORM)$, is a combination of truth-matching, preliminary selection, L0, HLT and occupancy cuts. The additional terms C and D (both of which operate on fully reconstructed `DecayTuple` samples) are necessary because the BDT which generates the w_{Reco} simulation weights is not normalized, and the inputs to this BDT are not available in the generator level `MCDecayTuple` samples. The C/D ratio therefore acts as an additional normalization for this weight.

Total efficiencies are calculated per-run and per-polarity and, if multiple candidates survive the full chain of selections, one is chosen randomly. Tables 7.1 and 7.2 report the total efficiencies for the resonant modes after the set of selections corresponding to the data fits in Sec. 8.3.2 and 8.3.3. The composite efficiencies that combine years and polarities are computed as a luminosity weighted sum:

$$\epsilon_{\mathrm{tot}} = \frac{\sum_{y,m} \mathcal{L}_j^i \cdot \epsilon_{tot}^{y,m}}{\sum_{y,m} \mathcal{L}_m^y} , \tag{7.5}$$

where $\mathcal{L}_m^y$ corresponds to the integrated luminosity for the year y and magnet polarity m. Due to the correlations between the different terms of Eq. 7.3, the uncertainty is obtained using a *bootstrapping* procedure that consists in assigning to each reconstructed candidate 100 different Poisson distributed weight values, with a mean value of 1. In this way, 100 different efficiencies can be obtained from the 100 data samples generated by the use of these weights. The mean and standard deviation of the efficiencies obtained is then used as nominal value and uncertainty for the estimated total efficiency. Note that a similar procedure can be used to obtain the uncertainties and correlations of the data-simulation corrections weights, and a straightforward combination of these two sources of systematics uncertainties on the absolute efficiency is possible.

Table 7.1: Total efficiencies ϵ_{tot} for the selection used in the mass fit of Sec.8.3.2.

Channel	Year	L0I	L0L!
$B^0 \to K^{*0} J/\psi(\to e^+e^-)$	RUN 1	$(6.67 \pm 0.02)\cdot 10^{-4}$	$(7.23 \pm 0.02)\cdot 10^{-4}$
	RUN 2	$(10.49 \pm 0.03)\cdot 10^{-4}$	$(11.95 \pm 0.03)\cdot 10^{-4}$
$B^0 \to K^{*0} J/\psi(\to \mu^+\mu^-)$	RUN 1	$(17.18 \pm 0.04)\cdot 10^{-4}$	$(44.66 \pm 0.06)\cdot 10^{-4}$
	RUN 2	$(23.70 \pm 0.05)\cdot 10^{-4}$	$(57.17 \pm 0.08)\cdot 10^{-4}$

Table 7.2: Total efficiencies ϵ_{tot} for the selection used in the mass fit of Sec.8.3.3.

Channel	Year	L0I	L0L!
$B^0 \to K^{*0} J/\psi(\to e^+e^-)$	Run 1	$(6.60 \pm 0.02)\cdot 10^{-4}$	$(7.18 \pm 0.02)\cdot 10^{-4}$
	Run 2	$(10.41 \pm 0.03)\cdot 10^{-4}$	$(11.87 \pm 0.03)\cdot 10^{-4}$
$B^0 \to K^{*0} J/\psi(\to \mu^+\mu^-)$	Run 1	$(17.00 \pm 0.04)\cdot 10^{-4}$	$(44.18 \pm 0.06)\cdot 10^{-4}$
	Run 2	$(23.45 \pm 0.05)\cdot 10^{-4}$	$(56.58 \pm 0.08)\cdot 10^{-4}$
$B^0 \to K^{*0} \psi(2S)(\to e^+e^-)$	Run 1	$(5.34 \pm 0.03)\cdot 10^{-4}$	$(7.09 \pm 0.04)\cdot 10^{-4}$
	Run 2	$(8.34 \pm 0.03)\cdot 10^{-4}$	$(11.35 \pm 0.04)\cdot 10^{-4}$
$B^0 \to K^{*0} \psi(2S)(\to \mu^+\mu^-)$	Run 1	$(15.72 \pm 0.05)\cdot 10^{-4}$	$(42.35 \pm 0.11)\cdot 10^{-4}$
	Run 2	$(21.22 \pm 0.05)\cdot 10^{-4}$	$(53.31 \pm 0.09)\cdot 10^{-4}$

7.1.2 Relative efficiency to control channel

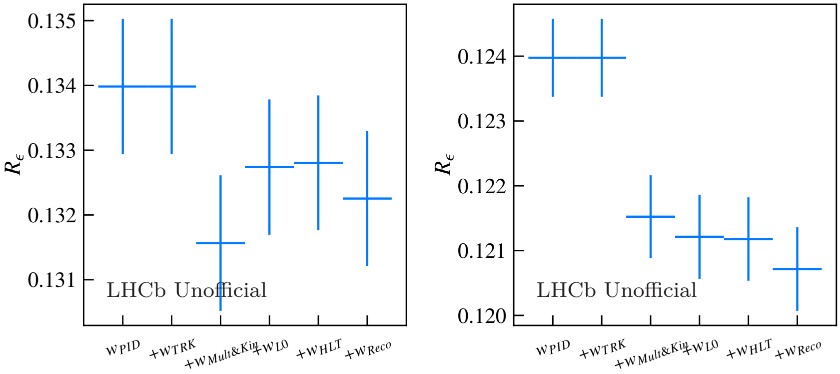
The relative efficiency between the rare and control mode is a fundamental ingredient of the measurement, as it allows estimating the expected number of events for the signal rare mode as a function of the parameters that describe the differential decay rate. A good control of this quantity over the different q^2 regions of interest, both in muons and electrons, is of utmost importance. The relative efficiency is defined as

$$R_\varepsilon(i,j,k) = \frac{\epsilon_{\text{tot}}^{sig(i,j,k)}}{\epsilon_{\text{tot}}^{J/\psi(j,k)}} \,, \tag{7.6}$$

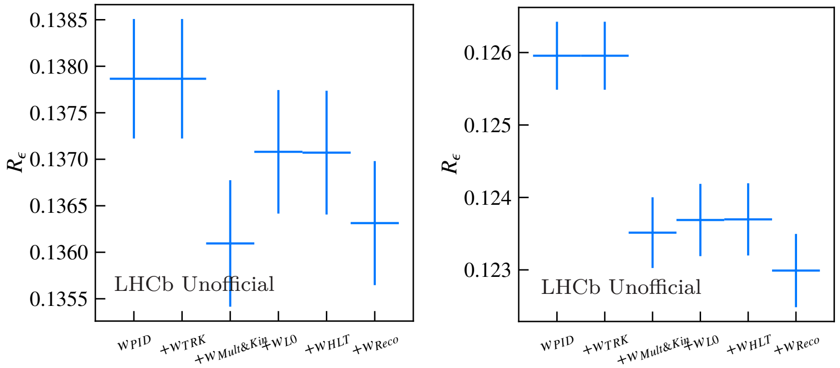
where the indices i,j and k refer to the q^2 region, run and trigger category of the computed ratio, respectively. Figures 7.1 and 7.2 show the evolution of these ratios after each of the corrections steps for $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays. The value of the ratios has been multiplied with a random number to avoid unblinding the value of absolute efficiencies of the rare mode. The maximum variation observed in each channel, run and trigger category is below 4%, with a coherent decreasing behaviour that is common between all of them. This suggests that any systematic in the estimation of the branching ratio coming from data-simulation differences should be below this level. At the same time, the impact of such a systematic on the final quantities is expected to be even lower, since part of the variation observed is correlated between muons and electrons, and is thus expected to largely cancel when looking at LFU quantities.

7.2 Effective acceptance

Bremsstrahlung, selections and reconstruction effects distort the distributions of the angles and q^2 of the signal decay, defined in Sec. 2.2. This distortion can be organized, for simplification, in three subsequent steps: final state radiation, acceptance and resolution, as shown in Eq. 7.7. Final state radiation corresponds to the bremsstrahlung emission of one or multiple photons by any of the charged



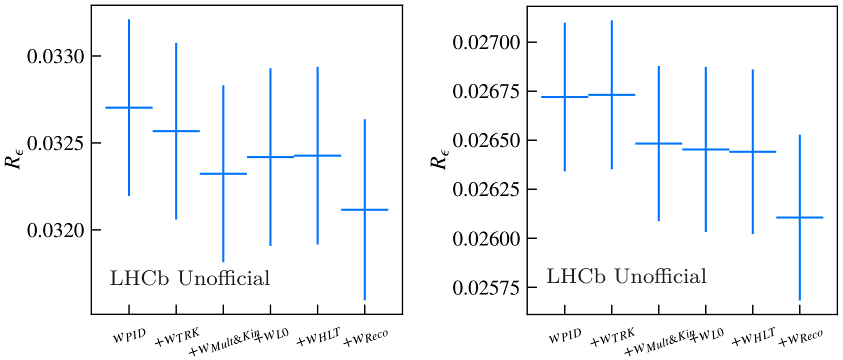
(a) Run 1 L0I (b) Run 1 L0M!



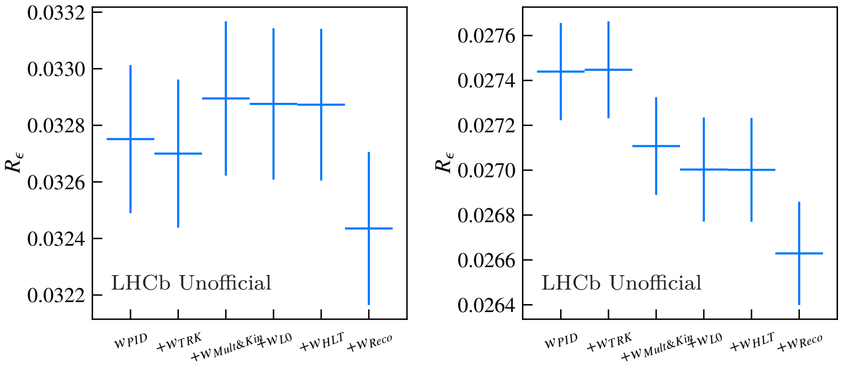
(c) Run 2 L0I (d) Run 2 L0M!

Figure 7.1: Evolution of the blinded efficiency ratio between $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ in the trigger categories L0I (left) and L0M! (right) and in Run 1 (top) and Run 2 (bottom).

particles of the signal final state. These emissions affect the measurement of the decay branching ratio and angular distribution, and can be particularly dangerous in LFU tests since they can induce non-universal corrections of the order of $\mathcal{O}(\alpha)\ln(m_\ell/M_B)$, which can be large between different lepton generations [31, 55]. These effects, not included in the differential decay rate of Eq. 2.2, are corrected through the use of the software package Photos [133]. The agreement between dedicated theoretical studies and the results provided by Photos ensures the possibility of effectively removing these QED effects and retrieve the observables of interest from data using simulation [55, 151]. The impact of FSR, whose exact functional form is not really of interest for this discussion, is described in the rightmost part of Eq. 7.7, with a generic convolution of the theory decay rate $f_{th}(\Omega_t, q_t^2)$ with the kernel $\mathcal{F}_{\text{QED}}$. The second contribution from the right in



(a) Run 1 L0I (b) Run 1 L0E!



(c) Run 2 L0I (d) Run 2 L0E!

Figure 7.2: Evolution of the blinded efficiency ratio between $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ in the trigger categories L0I (left) and L0E! (right) and in Run 1 (top) and Run 2 (bottom).

Eq. 7.7 represents the impact of our full selection on the distribution after the emission of FSR. This, usually referred to as *acceptance*, is indicated as $\epsilon(\Omega_f, q_f^2)$. The last contribution describes the *resolution*, that is the effect of the smearing of different features of the decay chain due to the physical characteristics of the LHCb detector. Notice that the boundaries between these three steps are not always well-defined, and is possible to find multiple situations where the order of the terms might be inverted, or two more steps mixed.

$$f(\Omega_r, q_r^2) = \underbrace{\int \mathrm{d}\Omega_r \mathrm{d}q_r^2 K(\Omega_r, q_r^2; \Omega_f, q_f^2)}_{\text{resolution}} \cdot \overbrace{\epsilon(\Omega_f, q_f^2)}^{\text{acceptance}} \cdot \underbrace{\int \mathrm{d}\Omega_t \mathrm{d}q_t^2 \mathcal{F}_{\text{QED}}(\Omega_f, q_f^2; \Omega_t, q_t^2)}_{\text{FSR}} f_{th}(\Omega_t, q_t^2) \tag{7.7}$$

The main drawback of using Eq. 7.7 directly in the measurement is its computational complexity, since it involves two eight-rank probability transition tensors and a four-dimensional acceptance function.

In the following work, a simplified approach is preferred: the transformation $f_{th}(\Omega_t, q_t^2) \to f(\Omega_r, q_r^2)$ is approximated by the *effective acceptance* function

$$\epsilon_{eff}(\Omega, q^2) = f(\Omega, q^2)/f_{th}(\Omega, q^2)\,, \tag{7.8}$$

which encodes the combined effect of FSR, selection and resolution. It is important to point out that for this expression to become exact, not only the acceptance, FSR and resolution kernels obtained from the simulation, but also the model used in the generation must be compatible with what is observed in data. The simplification obtained by the use of Eq. 7.8 brings thus an additional model dependence in the analysis, not present in Eq. 7.7. This model dependence is treated here as a source of systematic uncertainty, discussed more in detail in Sec. 10.6, and is expected to be more important as the width of the different convolution kernels grows. For muon decays, where both FSR and resolution effects are less relevant, the effective acceptance equals the selection acceptance, *i.e.* $\epsilon_{eff}(\Omega_r, q_r^2) = \epsilon_{eff}(\Omega_t, q_t^2)$, and almost no model dependence is expected in the parametrization of ϵ_{eff}. This strategy corresponds to what has already been used in the previous binned analysis of $B^0 \to K^{*0}\mu^+\mu^-$ [29]. Figure 7.3 shows a comparison of the resolution in the angles and dilepton mass squared for muons and electrons. As it can be seen, the resolution in the angles $\cos\theta_K$ and ϕ is almost identical between the two lepton generations, while a sizable difference can be observed for $\cos\theta_\ell$. The largest difference comes however from the dilepton mass squared, where the resolution in the electron mode is roughly four times that observed for muons. This last contribution is expected to be the main potential source of systematic uncertainty in the parametrization of the electron acceptance. However, thanks to the corrections described in Sec. 6.6, the resolution kernel in simulation matches closely what is observed in data, suggesting that the use of the effective acceptance is expected to be a good approximation of Eq. 7.7 also for the electron channel. Any residual model dependence on the underlying physics model will be considered as a source of systematics and is currently expected to be subleading with respect to the statistical uncertainty of the measurement.

7.2.1 Method of moments

The effective acceptance is parametrized using a method of moments approach in the three decay angles $\cos\theta_K$, $\cos\theta_\ell$, ϕ and q^2, without assuming factorization. This results in the expression

$$\epsilon_{eff}(\cos\theta_K, \cos\theta_\ell, \phi, q^2) = \sum_{k,l,m,n} c_{klmn} L(\cos\theta_K, k) L(\cos\theta_\ell, l) L(\phi', m) L({q^2}', n), \tag{7.9}$$

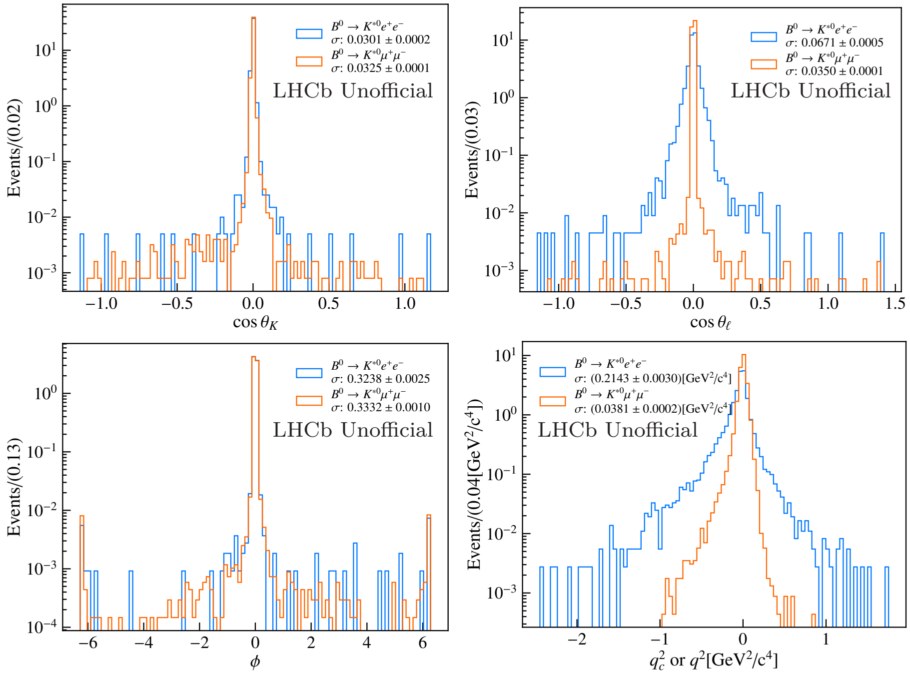


Figure 7.3: Resolution distributions for simulated $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays, defined as the difference between reconstructed and generator level variables after QED corrections for (top left) $\cos\theta_K$, (top right) $\cos\theta_\ell$, (bottom left) ϕ and (bottom right) q^2 and q^2_c. Note that the resolution in the angular distributions for muons and electrons is quite similar, with the biggest difference observed for $\cos\theta_\ell$.

where $L(x,a)$ is the Legendre polynomial of order a in the variable x, with $x \in [-1,1]$. Since the variables ϕ and q^2 are not naturally defined in the range desired, the parametrization is performed in ϕ' and $q^{2\prime}$, defined as

$$\begin{cases} q^{2\prime} & = \frac{2q^2 - {q^2}_{max} - {q^2}_{min}}{{q^2}_{max} - {q^2}_{min}} \\ \phi' & = \phi/\pi. \end{cases} \tag{7.10}$$

The coefficients c_{klmn} are determined using Monte Carlo integration and the orthogonality properties of the Legendre polynomials. Their definition corresponds to

$$
\begin{aligned}
c_{klmn} =& \frac{1}{N'}\sum_{i=1}^{N} w_i \cdot \tilde{c}^{i}_{klmn}(\cos\theta_K, \cos\theta_\ell, \phi, q^2) = \\
& \frac{1}{N'}\sum_{i=1}^{N} w_i \left[\left(\frac{2k+1}{2}\right)\left(\frac{2l+1}{2}\right)\left(\frac{2m+1}{2}\right)\left(\frac{2n+1}{2}\right)\right. \\
& \left. \times L(\cos\theta_K, k) L(\cos\theta_\ell, l) L(\phi, m) L(q^2, n)\right] ,
\end{aligned}
\tag{7.11}
$$

where the sum is performed over each i^{th} simulated event, N is the total number of events in the sample used and w_i is the per-event weight. The normalization is given by $N' = \sum_{i=1}^{N} w_i$.

7.2.2 Generator level parametrization

As described in Eq. 7.8, before the parametrization of ϵ_{eff} is necessary to obtain the inverse of the functional form of the generator level distribution used in the simulation of the decays of interest, $f^{gen}_{eff}(\cos\theta_K, \cos\theta_\ell, \phi, q^2)$. This is obtained by applying the method of moments to a set of generator level samples of $B^0 \to K^{*0}\mu^+\mu^-$, $B^0 \to K^{*0}e^+e^-$, $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ and $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays, where no QED corrections are included. The samples are produced using the LHCb software with the same configurations used for the generation of the reconstructed samples. The orders of the polynomials used to parametrize the generator level function are chosen to provide a satisfactory description of the distributions of interest and correspond to:

- 5, 4, 8, 18 in $\cos\theta_K$, $\cos\theta_\ell$, ϕ, q^2 for $B^0 \to K^{*0}\mu^+\mu^-$, where $q^2 \in [0.7, 8.5]\mathrm{GeV}^2/\mathrm{c}^4$;
- 5, 4, 8, 12 in $\cos\theta_K$, $\cos\theta_\ell$, ϕ, q^2 for $B^0 \to K^{*0}e^+e^-$, where $q^2 \in [0.5, 10]\mathrm{GeV}^2/\mathrm{c}^4$;
- 5, 4, 8 in $\cos\theta_K$, $\cos\theta_\ell$, ϕ both for $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ and $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$.

The ranges in the parametrization have been chosen to contain the q^2 regions of interest, and to have a buffer region at the boundaries to avoid being influenced by the border instability of the parametrization. Figure 7.4 shows the parametrization used for the generator level functions for $B^0 \to K^{*0}e^+e^-$ (left) and $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ (right) decays. The corresponding muon distributions are analogous and are not shown for brevity. A general agreement between the simulated datasets and the projected parametrization is observed. However, inspecting the distributions, it can be noticed that in $B^0 \to K^{*0}e^+e^-$ simulation there is a step in the q^2 region slightly below 6 GeV^2/c^4. This effect is related to the discontinuity of the effective Wilson coefficients $\mathcal{C}_7$ and $\mathcal{C}_9$ used in the parametrization of the $B^0 \to K^{*0}\ell^+\ell^-$ decay form factors [136] in `EvtGen`; the NNLO corrections, as obtained from Ref. [152] are only available up to roughly 5.7 Gev^2/c^4 and are afterwards turned off in the simulation, causing the discontinuity observed in q^2. The effect of this discontinuity is expected to have a marginal effect in the parametrization of the acceptance, thanks to the high order polynomials that closely mimic this step.

7.2.3 Effective acceptance parametrization

Once the generator level parametrization is available, the effective acceptance ϵ_{eff} can then be parametrized as described in Eq .7.9. The samples used correspond to truth-matched, fully weighted and smeared simulation candidates of $B^0 \to K^{*0}\mu^+\mu^-$, $B^0 \to K^{*0}e^+e^-$, $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ and $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays. The selections applied correspond to the full selection described in Sec. 5.2, where in particular:

- For $B^0 \to K^{*0}e^+e^-$ ($B^0 \to K^{*0}J/\psi(\to e^+e^-)$) decays, it is required a MVA greater than 0.97, a mass window of $m(K\pi ee) \in [4900, 5700]\,\text{MeV}/c^2$ ($m(K\pi ee) \in [4600, 5700]\,\text{MeV}/c^2$) and the region of q_c^2 corresponding to central-q_c^2 (J/ψ-q_c^2). In $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays is additionally requested that $m_{DTF}^{J/\psi} > 5150\,\text{MeV}/c^2$, to reduce to a negligible level the partially reconstructed component in the fit.
- For $B^0 \to K^{*0}\mu^+\mu^-$ ($B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$) decays, it is required a MVA greater than 0.8, a mass window of $m(K\pi\mu\mu) \in [5170, 5700]\,\text{MeV}/c^2$ and the region of q^2 corresponding to central-q^2 (J/ψ-q^2).

The per-event weight w_{acc} used in simulation corresponds to:

$$w_{acc} = \frac{w_{tot} \cdot w_{mix}}{f_{eff}^{gen}(\cos\theta_{K\,r}, \cos\theta_{\ell r}, \phi_r, {q^2}_r)}\,, \tag{7.12}$$

where the weight w_{mix} allows the correct combination of simulated samples for different years and magnet polarities, taking into account their relative integrated luminosities and total efficiencies. To avoid instabilities in the parametrization, it is additionally required that $f_{eff}^{gen}(\cos\theta_K, \cos\theta_\ell, \phi, q^2) \leq 0$, to remove negative and infinite weights, and that the top 0.05%, 0.2% and 0.003% highest weights for $B^0 \to K^{*0}\mu^+\mu^-$, $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays are rejected. These thresholds have been motivated by an empirical inspection of the 1D-projection of the acceptance parametrizations, together with the negligible impact that they seem to have on the generator level observables as shown in Sec. 9.2.2. It is important to point out that the higher the statistics in the simulated sample, the smaller is the percentage rejected, suggesting that the need to remove more events is mainly due to high weights that are not properly averaged out in some corners of the phase space.

The acceptance is modelled using the lowest orders of the polynomials that show good description of the angular distributions, corresponding to Legendre

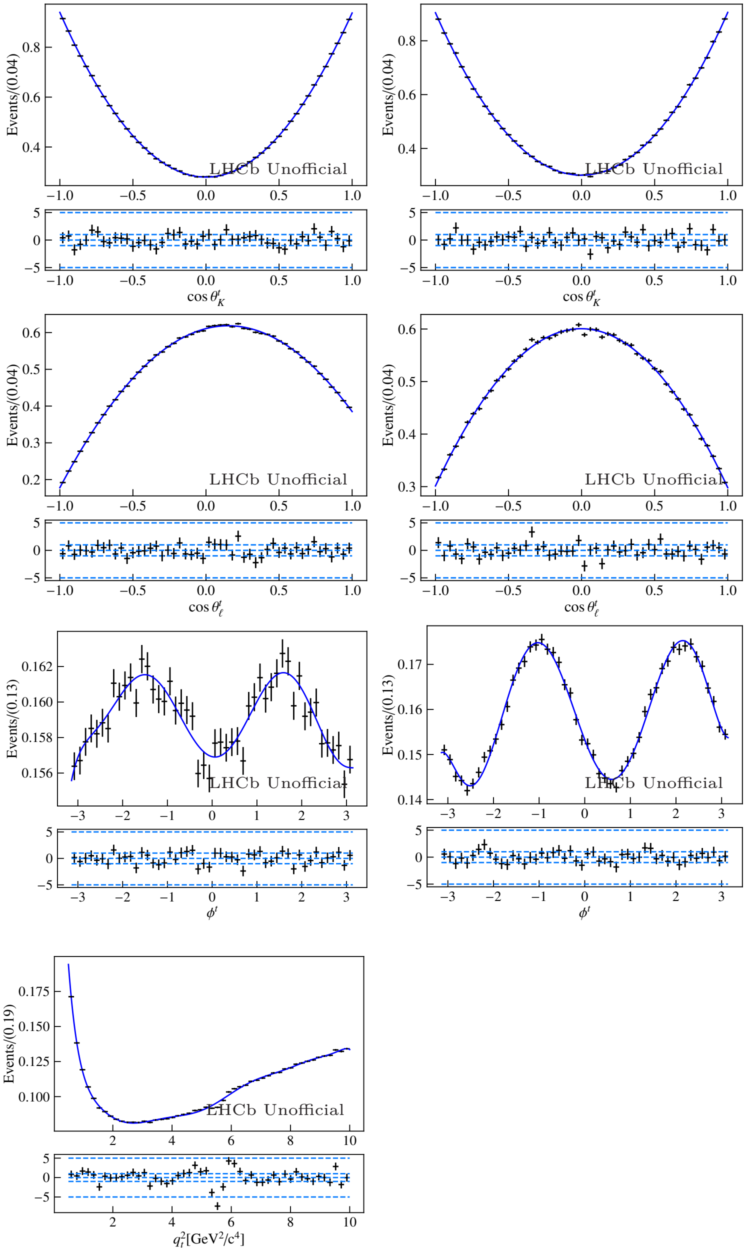


Figure 7.4: Generator level distributions of $\cos\theta_K$, $\cos\theta_\ell$, ϕ and q^2 for $B^0 \to K^{*0}e^+e^-$ (left) and $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ (right) simulated decays (in black) with the parametrization performed with the method of moments in blue. A similar parametrization is performed for $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$ decays. No distribution is shown for q^2 in $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ since its parametrization is not used in the determination of the effective acceptances.

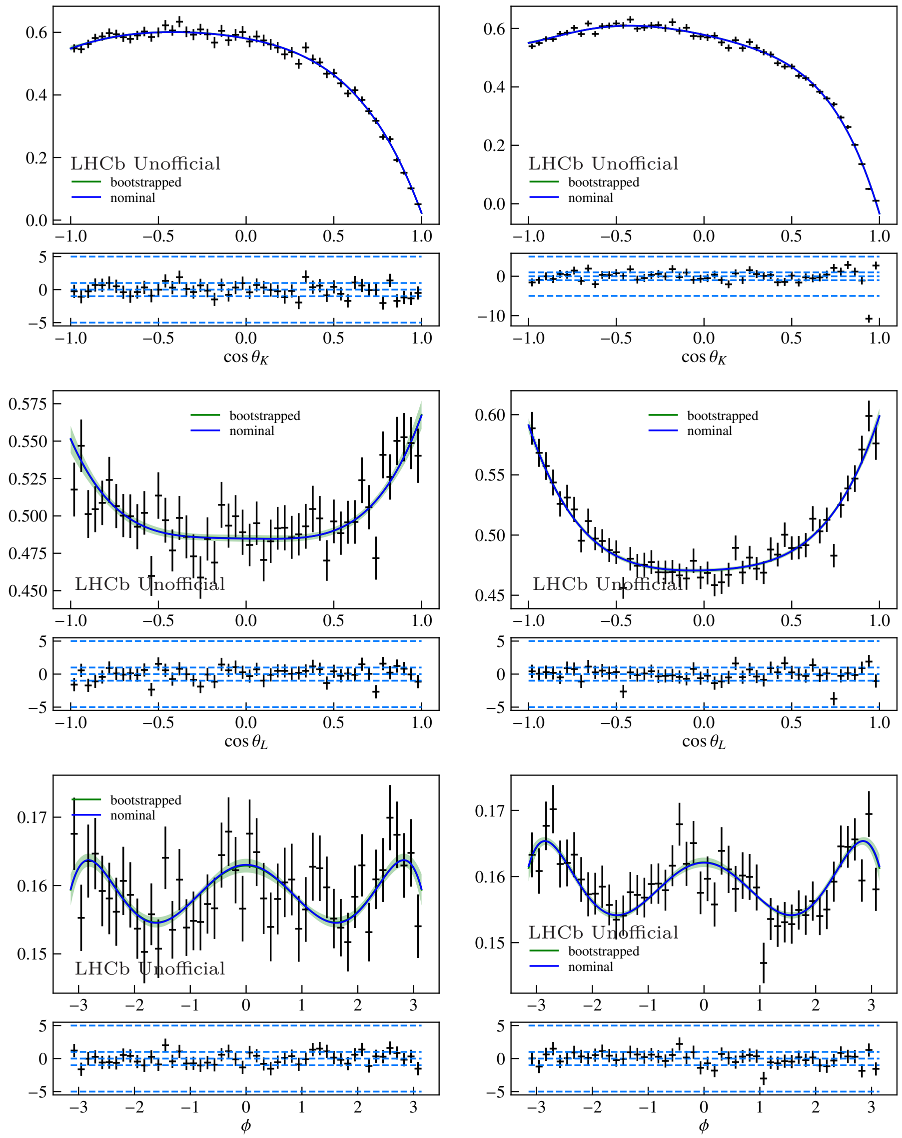


Figure 7.5: Acceptance parametrization for $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ (left) and $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ (right) in Run 2 for the trigger category L0I. In blue is shown the nominal acceptance, while in green is shown the 1σ contour given by the bootstrapping of the samples used to parametrize the acceptance.

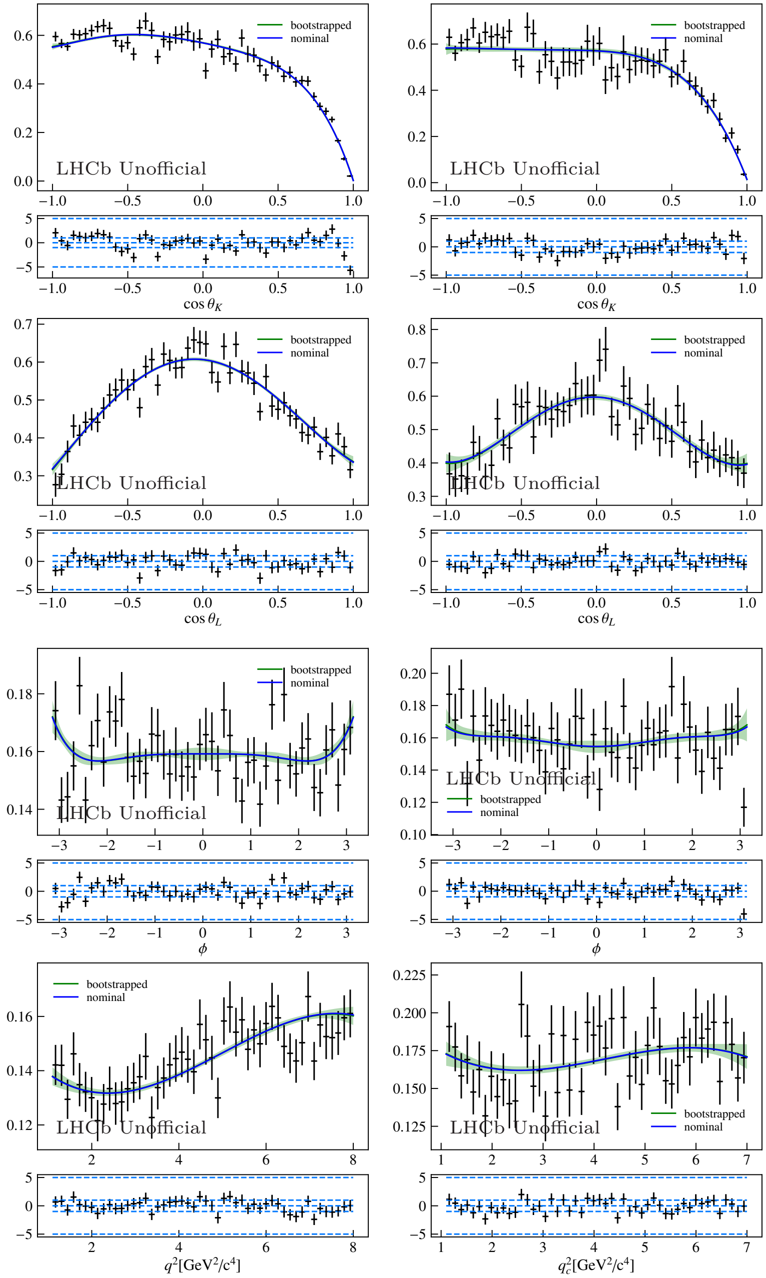


Figure 7.6: Acceptance parametrization for $B^0 \to K^{*0}\mu^+\mu^-$ (left) and $B^0 \to K^{*0}e^+e^-$ (right) in Run 2 for the trigger category L0I. In blue is shown the nominal acceptance, while in green is shown the 1σ contour given by the bootstrapping of the samples used to parametrize the acceptance.

polynomials of order five or less for $\cos\theta_K$, four or less for $\cos\theta_\ell$, six or less for ϕ. Additionally, to enforce a symmetric behaviour in ϕ, the odd terms in its parametrization have been set to zero. Legendre polynomials of order three are used for q^2 and q_c^2 in $B^0\to K^{*0}\mu^+\mu^-$ and $B^0\to K^{*0}e^+e^-$ decays, respectively, while no dependency on these variables is present for the parametrization of the acceptance in the resonant channels. In the rare (control) mode, these correspond to a total of 840 (210) coefficients, some of which are set to zero due to the symmetry of ϕ. The acceptance is parametrized separately for each run of data-taking and for each trigger category of interest. Figures 7.5 and 7.6 show the parametrization of the acceptance for $B^0\to K^{*0}J/\psi(\to\ell^+\ell^-)$ and $B^0\to K^{*0}\ell^+\ell^-$ decays in Run 2 for the L0I trigger category. As can be seen, the projections of the parametrization closely match the reweighted sample in most of the phase space with one clear exception: the high region of $\cos\theta_K$ in the muon channels. The impact that this mismodelling could have is discussed in Sec. 9.2.3, where the possibility to use higher order parametrizations of the acceptance to estimate the size of the bias introduced is investigated. Similar behaviours are observed for the other trigger categories and runs and are not shown here for brevity.

Chapter 8

Invariant mass parametrization and yield extraction from fits

The following chapter discusses two fundamental ingredients that enter the amplitude fits to the rare mode: the parametrization of the rare mode invariant mass distributions for $B^0 \to K^{*0}\ell^+\ell^-$ decays, and the determination of the yields from $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ reconstructed decay candidates. The former are described in Sec. 8.2 together with the invariant mass parametrizations for $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays. The latter are obtained in Sec. 8.3.2 by performing simultaneous fits to $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ decay candidates in the invariant mass $m(K\pi\ell\ell)$. The chapter ends with Sec. 8.3.3 describing an additional set of fits to $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays performed in the $m_{DTF}^{J/\psi}$ and $m_{DTF}^{\psi(2S)}$ variables. As discussed in more detail in Sec. 9.1, these fits can be used to test the quality of the correction chain in regions of q^2 different from where they have been derived.

8.1 Fit generalities

All mass fits are obtained through the maximization of an unbinned extended maximum likelihood $\mathcal{L}$, defined as

$$\mathcal{L}(x; N_{tot}, \theta) = \frac{e^{-N_{tot}} \cdot N_{tot}^{N_{obs}}}{N_{obs}!} \times \prod_{i}^{N_{obs}} \mathcal{P}(x_i; \theta)\,, \tag{8.1}$$

where the first term corresponds to the Poissonian probability of observing a total number of events N_{obs} if the expected number corresponds to N_{tot}, while the second corresponds to the probability that the observed events $x = \{x_i\}_{i=1}^{N_{obs}}$ are obtained from the *probability density function* (*pdf*) $\mathcal{P}(x_i; \theta)$ with the set of parameters θ. The values of the parameters that maximize the likelihood $\mathcal{L}$ represent the best estimate of the true parameters for N_{tot} and θ, and are indicated as $\hat{N}_{tot}$ and $\hat{\theta}$.

In fits to data, the chosen mass range can be populated not only by signal, but also by several background contributions. The function $\mathcal{P}(x_i; \theta)$ can then be re-written as

$$\mathcal{P}(x_i; \theta) = \frac{N_S}{N_{tot}} S(x_i; \theta) + \sum_k \frac{N_B^k}{N_{tot}} B_k(x_i; \theta)\,, \tag{8.2}$$

where N_S and $S(x_i; \theta)$ correspond to the expected yield and *pdf* of the signal component, respectively, while N_B^k and $B_k(x_i; \theta)$ correspond respectively to the expected yield and *pdf* of the k^{th} background component considered in the fit.

Some backgrounds components might not have a distinctive feature in their invariant mass and, when included in the fit, could cause some instability. This issue is greatly reduced by constraining the expected size of these backgrounds with respect to the resonant modes in the fit. For the decay of a generic b-hadron H_b, the expected yield can be expressed as:

$$\mathcal{N}_{H_b}^{\psi_n} = \frac{f_{H_b}}{f_d} \cdot \frac{\mathcal{B}(H_b \to X)}{\mathcal{B}(B^0 \to K^{*0}\psi_n(\to \ell^+\ell^-))} \cdot \frac{\epsilon(H_b \to X)}{\epsilon_{\psi_n}} \cdot \mathcal{N}_{B^0 \to K^{*0}\psi_n(\to \ell^+\ell^-)} \,, \tag{8.3}$$

where:

- $N_{H_b}^{\psi_n}$ corresponds to the expected yield of the decay mode $H_b \to X$, calibrated from the observed yield in the resonant mode $\mathcal{N}_{B^0 \to K^{*0}\psi_n(\to \ell^+\ell^-)}$;
- f_{H_b} and f_d correspond to the fragmentation fractions to produce a H_b and B^0 hadron, respectively;
- $\mathcal{B}(H_b \to X)$ and $\mathcal{B}(B^0 \to K^{*0}\psi_n(\to \ell^+\ell^-))$ indicate the branching ratios of the corresponding modes;
- $\epsilon(H_b \to X)$ and ϵ_{ψ_n} correspond to the total efficiencies of the background and resonant mode, respectively, as obtained from simulated samples. Only PID and tracking corrections are considered for this calculation.[1]

The total expected yield $\mathcal{N}_{H_b}^{\psi_n}$ can then be constrained by applying a set of Gaussian constraints to the reconstruction efficiencies, fragmentation fractions and branching ratios of Eq. 8.3. The total likelihood is then modified with the multiplication of an additional penalty term of the form:

$$\mathcal{L}_{\text{constr}} = \prod_j \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{1}{2}\left(\frac{\eta_j - \eta_j^{\text{constr}}}{\sigma_j^{\text{constr}}}\right)^2} \,, \tag{8.4}$$

where η_j corresponds to the additional fit parameters introduced with the j^{th} Gaussian constraint, and η_j^{constr} and σ_j^{constr} to then mean and standard deviation of the parameter, as determined either from simulation or from literature.

In the case in which different fits use the same parameters, *e.g.* multiple backgrounds share the same fragmentation fraction, it is sometimes convenient to perform these fits *simultaneously*, thus allowing the optimization to make use of the higher statistics that comes from the combination of different samples. A simultaneous fit simply corresponds to the optimization of the joint likelihood

$$\mathcal{L}_{\text{sim}} = \mathcal{L}_{\text{constr}} \times \prod_k \mathcal{L}_k \,, \tag{8.5}$$

---

[1] The size of the uncertainty on the efficiencies obtained from the backgrounds considered is expected to be larger than the variation due to the different corrections steps. To simplify the analysis strategy only the first correction steps are considered.

equal to the multiplication of the single likelihoods $\mathcal{L}_k$. The term describing the Gaussian constraints, if present, should be included without repetition.

All likelihoods and *pdf* used in this work are implemented in `TensorFlow` [153] within a framework developed specifically for the amplitude analysis described in Ref. [144] and in this work, while the optimization of the likelihood is carried out in `MINUIT` [154]. The performance of this framework has been cross-checked and verified against the fitting package RooFit [155], widely used in the HEP community.

8.2 Mass signal parametrization from simulation

From first principles, the invariant mass lineshape for the decays $B^0 \to K^{*0}\psi_n(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\ell^+\ell^-$ should correspond to a narrow peak at $5279.6{\pm}0.15\,\mathrm{MeV}/c^2$, as reported by Ref. [129]. However, due to a combination of effects such as bremsstrahlung emission and the finite momentum resolution of the tracking in LHCb, the sharp peak is modified into a distribution determined by the details of the reconstruction. To account for these multiple effects, the main strategy is to use the fully corrected and smeared simulation to obtain an effective parametrization of the invariant mass line shape, and then to allow some residual freedom when fitting the data. This approach has, however, the disadvantage of breaking the correlation between the reconstructed invariant mass of the B^0 and the remaining variables of the fit. The impact of such effect is thought to be small and neglected in this work.

As already discussed in Sec. 5.2.4.2, only two types of invariant masses are used in the fits: $m(K\pi\ell\ell)$ and $m^{\psi_n}_{DTF}$, with $\psi_n = J/\psi\,, \psi(2S)$. The former corresponds to the latter if the decay is refitted constraining the dilepton pair mass to the nominal mass of the resonances. Their shape is modelled from truth-matched fully corrected simulated samples as a superposition of Gaussians and double-sided Crystal Ball (DSCB) functions, as explained more in detail in the following. A DSCB corresponds to a variation of the standard Crystal Ball function [150], characterized by a common Gaussian core shared by two independent power law tails on the opposite sides of the peak and defined as:

$$\mathcal{P}_{\mathrm{DSCB}}(m|\mu,\sigma,\alpha,n) = \begin{cases} \dfrac{a_L}{\left(b_L - \frac{(m-\mu)}{\sigma}\right)^n} & \text{if} \quad \frac{(m-\mu)}{\sigma} < \alpha_L \\ e^{\frac{-(m-\mu)^2}{2\sigma^2}} & \text{if} \quad -\alpha_L < \frac{(m-\mu)}{\sigma} < \alpha_R \\ \dfrac{a_R}{\left(b_R + \frac{(m-\mu)}{\sigma}\right)^n} & \text{if} \quad \frac{(m-\mu)}{\sigma} > \alpha_R \end{cases} \quad (8.6)$$

with

$$\begin{aligned} a_{L,R} =& \left(\frac{n_{L,R}}{|\alpha_{L,R}|}\right)^{n_{L,R}} e^{-\frac{1}{2}\alpha^2_{L,R}} \,, \\ b_{L,R} =& \frac{n}{|\alpha_{L,R}|} - \alpha_{L,R} \,, \end{aligned} \tag{8.7}$$

where μ and σ are the mean value and resolution of the Gaussian part of the function, the sign of the α parameter governs the right-handed or left-handed location of the tail and the parameter n drives the power law associated to the decrease of the tail.

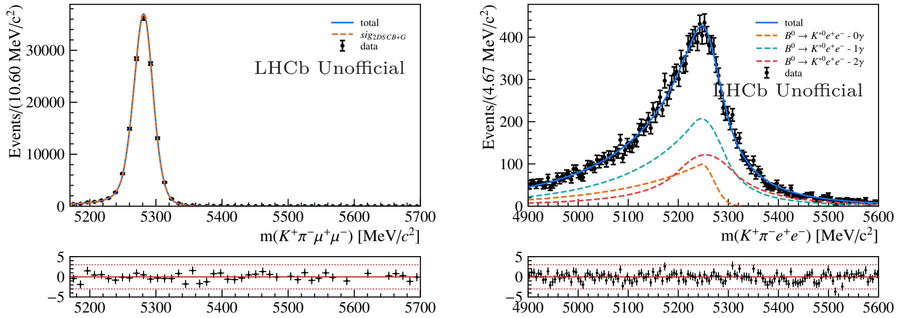
The invariant mass shape used in fits to the muon channels is parametrized with two DSCB and a Gaussian, corresponding to:

$$\begin{aligned} S^{\mu}(m) =& f_2 \cdot \big(f_1 \cdot \mathcal{P}_{\text{DSCB}}(m|\mu, \sigma_1, \alpha_L, \alpha_R, n_L, n_R) \\ &+ (1-f_1) \cdot \mathcal{P}_{\text{DSCB}}(m|\mu, \sigma_2, , \alpha_L, \alpha_R, n_L, n_R)\big) \\ &+ (1-f_2)\mathcal{G}(m|\mu_G = \mu + \Delta\mu_G, \sigma_G) \;. \end{aligned} \tag{8.8}$$

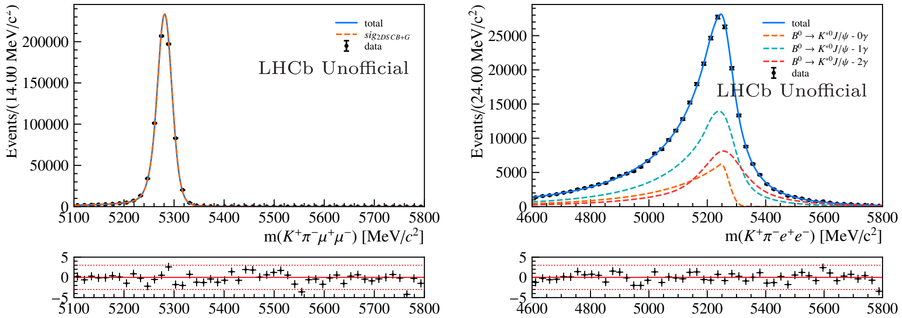
In the electron case, the invariant mass shape strongly depends on the number of electrons in the final state that have been corrected by the bremsstrahlung recovery algorithm. To improve the description of the mass shape, the events are therefore separately parametrized in three *bremsstrahlung categories*: 0γ, 1γ and 2γ, corresponding to events for which none, one or more than one photon cluster has been added to the dielectron candidate to correct its momentum due to bremsstrahlung. The full mass model can be then written as

$$S^e(m) = f_{0\gamma}S_{0\gamma}(m) + f_{1\gamma}S_{1\gamma}(m) + (1 - f_{0\gamma} - f_{1\gamma})S_{2\gamma}(m) \;, \tag{8.9}$$

where $f_{0\gamma}$ and $f_{1\gamma}$ correspond to the fractions of events belonging to the 0γ and 1γ categories in simulation. In fits to $B^0 \to K^{*0}e^+e^-$ decays, the shape used to parametrize the invariant mass of the single bremsstrahlung category is a DSCB. An additional Gaussian is added for the shapes of the 1γ and 2γ categories for the fits to $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays in the mass $m(K\pi ee)$. To reduce the differences caused by the mismodelling of the resolution in the dielectron momenta, the signal shapes are parametrized using the smeared mass $m^{smrd}(K\pi e^+e^-)$ rather than $m(K\pi ee)$ (see Sec. 6.6). In fits to $m(K\pi ee)^{\psi(2S)}_{DTF}$, the combination of a DSCB and a Gaussian is used for all bremsstrahlung categories, while in fits to $m(K\pi ee)^{J/\psi}_{DTF}$ the parametrization is extended including a second additional Gaussian in each category. In all cases, the choice of including an additional Gaussian is based on the improvement that such term provides on the parametrization of the simulation in the tails of the distributions. Figure 8.1 shows the results of the fit results to simulated candidates for $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays in the $m(K\pi\ell\ell)$ system, while Figure 8.2 shows the corresponding results for the fits to $m(K\pi\ell\ell)^{J/\psi}_{DTF}$ in $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $m(K\pi\ell\ell)^{\psi(2S)}_{DTF}$ in $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays.



(a) Fits to simulated $B^0 \to K^{*0}\ell^+\ell^-$ decays in $m(K\pi\ell\ell)$

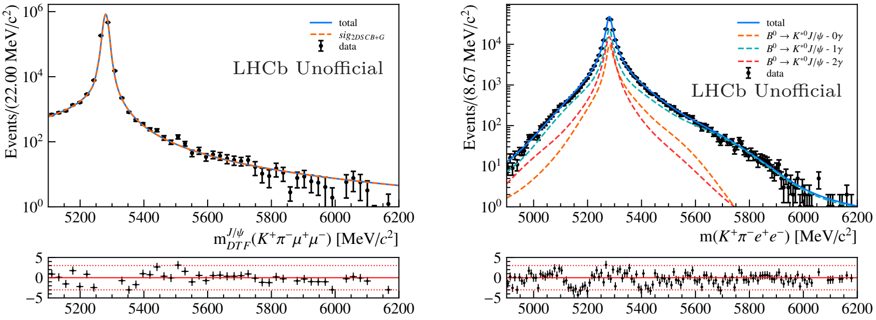


(b) Fits to simulated $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays in $m(K\pi\ell\ell)$

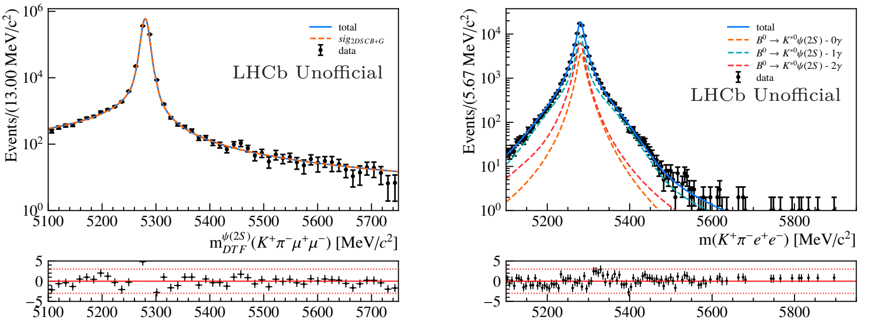
Figure 8.1: (a) Fits to the invariant mass distributions $m(K\pi\ell\ell)$ of the simulated decays of $B^0 \to K^{*0}\mu^+\mu^-$ (left) and $B^0 \to K^{*0}e^+e^-$ (right) in the central-q^2 and central-q^2_c regions, respectively. (b) Fits to the invariant mass distributions of the simulated decays of $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ (left) and $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ (right) in the J/ψ-q^2 and J/ψ-q^2_c regions, respectively. All fits shown correspond to samples selected in L0I trigger category for Run 2 conditions.

The left columns correspond to the muons, while the right ones to the electron modes.

The parametrizations thus obtained are used as starting point to describe the mass lineshape of the different signal channels in data. All parameters obtained from simulation are fixed to their best fit values. To account for residual simulation-data differences, two additional parameters are introduced and allowed to float in the fit: $\Delta\mu$ and s_σ, corresponding to a shift and a stretching of the distribution. Additional care must be taken when stretching *pdf*s that contain multiple Gaussians together with a DSCB: since these are not forced to be centred at the peak of the DSCB, their centre must be transformed as $\mu_G \to (\mu_G \cdot s_\sigma - (s_\sigma - 1) \cdot \mu_{DSCB}) + \Delta\mu$. In the fits to $m(K\pi e^+e^-)$ for $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays, a further parameter is allowed to float: $s^e_{\alpha_R}$,



(a) Fits to simulated $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ decays in $m(K\pi\ell\ell)^{J/\psi}_{DTF}$



(b) Fits to simulated $B^0 \to K^{*0} \psi(2S)(\to \ell^+\ell^-)$ decays in $m(K\pi\ell\ell)^{\psi(2S)}_{DTF}$

Figure 8.2: (a) Fits to the invariant mass distributions $m(K\pi\ell\ell)^{J/\psi}_{DTF}$ of the simulated decays of $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ (left) and $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ (right) in J/ψ-q^2 ($R_{K^{*0}}$). (b) Fits to the invariant mass distributions $m(K\pi\ell\ell)^{\psi(2S)}_{DTF}$ of the simulated decays of $B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-)$ (left) and $B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$ (right) in $\psi(2S)$-q^2 ($R_{K^{*0}}$). All fits shown correspond to samples selected in L0I trigger category for RUN 2 conditions.

representing an additional degree of freedom for the right-hand tail of the DSCB. This parameter is introduced in $S^e(m)$ through the transformation $\alpha_R^{data} = s^e_{\alpha_R} \cdot \alpha_R^{sim}$ and $n_R^{data} = n_R^{sim}/s^e_{\alpha_R}$ and is needed to improve the description of the right-hand tail in data.

8.3 Mass fits to $B^0 \to K^{*0}\psi_n$ decays

In the following, the two set of mass fits performed on the resonant modes are presented. Due to the similarities between the fits in $m(K\pi\ell\ell)$ in the $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ channel and the fits in $m(K\pi\ell\ell)^{J/\psi}_{DTF}$ and $m(K\pi\ell\ell)^{\psi(2S)}_{DTF}$ for the $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ channels, the discussion of the background composition used in the mass fits is kept together and discussed in Sec. 8.3.1. The details of the different fits are instead discussed in Secs. 8.3.2 and 8.3.3.

8.3.1 Background parametrization

The backgrounds components considered in fits to $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ data candidates are:

- *Combinatorial*:
  The combinatorial background is modelled with a decreasing exponential of slope λ^ℓ, separately for different lepton generations $\ell = \mu, e$, runs and trigger categories. The expected yield $\mathcal{N}^\ell_{comb}$ and the slope λ^ℓ are left free to float in the fits.

- $\Lambda^0_b \to pK\psi_n(\to \ell^+\ell^-)$, with $\psi_n = J/\psi\,, \psi(2S)$:
  This contribution is modelled with a binned KDE lineshape, separately for different trigger categories, but joining the different years of simulation available regardless of the different integrated luminosities. The samples have been truth-matched allowing single ($p_{\to\pi}$) and double mis-ID ($p_{\to K}, K_{\to\pi}$) contributions. Fig. C.1(a) shows an example of such parametrization for the L0I trigger category for muons and electrons, respectively. To improve the quality of the simulation used, two set of corrections have been considered: a data-driven correction for the underlying dynamics of the two-dimensional plane of $m(pK)$ and $m(pJ/\psi)$ based on Ref. [156], and a two-dimensional kinematic correction of the $p_T(\Lambda^0_b)$ and $\eta(\Lambda^0_b)$ spectrum of the Λ^0_b at production based on Ref. [62]. The shape obtained is kept fixed in fits to data. Due to the small expected size of this background in data (*e.g.* see Tables 5.9), its yield cannot be allowed to float freely. Instead, it is parametrized with respect to the observed yield of the corresponding signal resonant mode using Eq. 8.3, where the parameters $\frac{f_{\Lambda^0_b}}{f_d}$, $\frac{\mathcal{B}(\Lambda^0_b \to pK\psi_n(\to \ell^+\ell^-))}{\mathcal{B}(B^0 \to K^{*0}\psi_n(\to \ell^+\ell^-))}$ and $\frac{\epsilon(\Lambda^0_b \to pK\psi_n(\to \ell^+\ell^-))}{\epsilon_{\psi_n}}$ are Gaussian constrained to their known values of Ref. [129] and to the values obtained from the available simulated samples.

- $B^0_s \to K^{*0}\psi_n(\to \ell^+\ell^-)$, with $\psi_n = J/\psi\,, \psi(2S)$:
  This background is modelled using the same mass shape used for the signal resonant mode, but with means shifted upwards of the quantity $\Delta m_{B^0_s}$ that corresponds to the mass difference between B^0_s and B^0 [129]. The size of this background contribution is parametrized with respect to observed yield of the resonant mode through the fraction $f_{B^0_s}$ and $f_{B^0_s} \cdot \frac{\mathcal{B}B^0_s \to K^{*0}\psi(2S)}{\mathcal{B}(B^0_s \to K^{*0} J/\psi\,)}$ for fits to $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays, respectively. The fraction $f_{B^0_s}$ is allowed to float freely in data while the fraction $\frac{\mathcal{B}(B^0_s \to K^{*0}\psi(2S))}{\mathcal{B}(B^0_s \to K^{*0} J/\psi\,)}$ is Gaussian constrained to the values obtained from Ref. [129]. Notice that, while floating the $f_{B^0_s}$ in an electron-only

fit would be unfeasible, this can be done in the fits of Secs. 8.3.2 and 8.3.3, since this parameter is always shared with the corresponding muon channel.

- $B_s^0 \to \phi J/\psi(\to \ell^+\ell^-)$:
  This contribution is modelled with a binned KDE lineshape, separately for different runs and trigger categories. Fig. C.1(b) shows, as an example, the parametrization obtained for the Run 1 L0I simulated samples of muons and electrons, respectively. Notice that in the $\psi(2S)-q^2$-region this component is not included, as it is already absorbed in the $B_s^0 \to XJ/\psi(\to \ell^+\ell^-)$ simulated sample. The yield of this background is parametrized with respect to the observed yield in the corresponding resonant mode as described by Eq. 8.3, where the parameters $\frac{f_s}{f_d}$ and $\frac{\mathcal{B}(B_s^0\to\phi J/\psi(\to\ell^+\ell^-))}{\mathcal{B}(B^0\to K^{*0}J/\psi(\to\ell^+\ell^-))}$ are constrained to their known values [129] and $\frac{\epsilon(B_s^0\to\phi J/\psi(\to\ell^+\ell^-)))}{\epsilon_{\psi_n}}$ to the values obtained from the available simulated samples.

- $\mathrm{K}-\pi$ swap:
  This background is parametrized using a binned KDE lineshape, separately for different runs and trigger categories. Fig. C.1(c) shows, as an example, the parametrization obtained for the Run 1 L0I simulated samples of muons and electrons, respectively. The samples used for the parametrization are the same ones used for the resonant signal simulation, with the difference that the events are truth-matched requiring the double swap $K \leftrightarrow \pi$. The background yield is parametrized with respect to the observed yield in the corresponding resonant mode as described by:

$$\mathcal{N}^{\psi_n}_{\mathrm{K}-\pi\,\mathrm{swap}} = \frac{\epsilon_{swap}}{\epsilon_{\psi_n}} \cdot \mathcal{N}_{B^0\to K^{*0}\psi_n(\to\ell^+\ell^-)}\,, \tag{8.10}$$

  where the parameter $\frac{\epsilon_{swap}}{\epsilon_{\psi_n}}$ is constrained to the value obtained from the available simulated samples.

- Partially reconstructed $B \to XJ/\psi(\to \ell^+\ell^-)$ decays:
  This contribution is modelled with a binned KDE lineshape, separately for different trigger categories but joining the different years of simulation available, regardless of the different integrated luminosities. An example of such parametrization is shown in Fig. C.1(d) for the L0I simulated samples of muons and electrons, respectively. The samples used in muon fits correspond to an inclusive simulation of B^0 and B^+ decays, while for electrons an additional inclusive B_s^0 component is considered. The different samples and channels that compose them are reweighted considering their relative efficiencies, hadronization fractions, isospin factors and total branching ratios. In fits to $B^0\to K^{*0}J/\psi\,(\to e^+e^-)$ candidates using $m(K\pi ee)^{J/\psi}_{DTF}$, due to the larger dependency of the lineshape on the presence or absence of higher $c\bar{c}$ resonances, two different components are considered to parametrize the partially reconstructed: decays where the J/ψ comes directly from the B, referred to as $B \to XJ/\psi(\to e^+e^-)$, and

decays where the J/ψ comes from a higher $c\bar{c}$ resonance, referred to as $B \to Y(\to J/\psi X)Z$. For the remaining $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ fits, a single component parametrized from corrected simulation is considered. This source of background is not included in fits to $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ candidates, due to their negligible size. The shapes obtained are fixed to the simulation, while the observed yield is allowed to float freely in the fit to data.

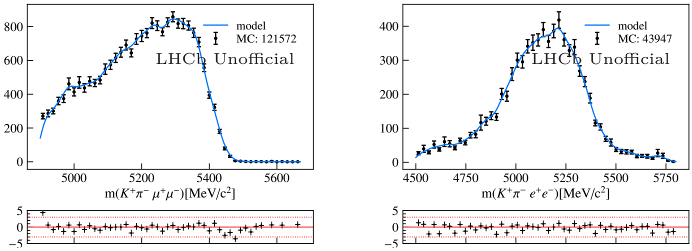
Extra components are considered in the fits to $B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$ decays using $m_{DTF}^{\psi(2S)}(K\pi e^+e^-)$:

- $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ leakage:
  This background is parametrized with a binned KDE lineshape, separately for different runs and trigger categories. An example of such parametrization is shown Fig. 8.4(a) for the L0I trigger category in Run 1. The yield of this contribution is allowed to float freely in the fit, while its shape is fixed to simulation.

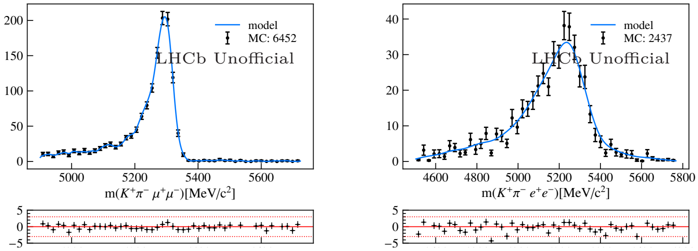
- *Partially reconstructed* $B^0 \to K^{*0}\psi(2S)(\to XJ/\psi(\to e^+e^-))$:
  This contribution is modelled from simulation with a binned KDE lineshape, separately for different runs and trigger categories. An example of such parametrization is shown Fig. 8.4(b) used in the Run 1 data fit to the L0I trigger category. The samples used for the parametrization are simulated events of $B^0 \to K^{*0}\psi(2S)(\to \pi^+\pi^- J/\psi(\to e^+e^-))$ decays, where the decay $\psi(2S) \to \pi^+\pi^- J/\psi$ constitutes roughly half of the inclusive decays of a charmonium $\psi(2S)$ into a J/ψ state. This is expected to be a good proxy for the mass shape of the remaining decays that involve neutral particles instead of the two pions. The yield of this background is parametrized with respect to the observed yield in the corresponding resonant mode and described by:

$$\begin{aligned}\mathcal{N}_{bkg}^{\psi(2S)} &= \frac{\mathcal{B}(\psi(2S) \to \pi^+\pi^- J/\psi) \cdot \mathcal{B}(J/\psi \to e^+e^-)}{\mathcal{B}(\psi(2S) \to e^+e^-)} \\ &\times \frac{\epsilon(B^0 \to K^{*0}\psi(2S)(\to \pi^+\pi^- J/\psi(\to e^+e^-)))}{\epsilon_{\psi_n}} \\ &\times \mathcal{N}_{B^0 \to K^{*0}\psi(2S)(\to e^+e^-)}\,,\end{aligned} \tag{8.11}$$

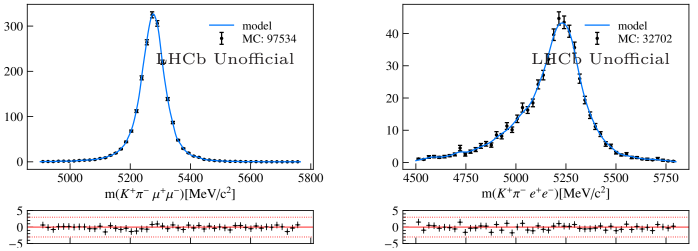
  where the parameter $\frac{\epsilon(B^0 \to K^{*0}\psi(2S)(\to \pi^+\pi^- J/\psi(\to e^+e^-)))}{\epsilon_{\psi_n}}$ is Gaussian constrained to value obtained from the available simulated samples. To allow the fit to compensate for the fact that only half of the $\psi(2S) \to XJ/\psi$ decays are considered in the branching ratio and in the efficiency, the constraint applied on $\frac{\mathcal{B}(\psi(2S) \to \pi^+\pi^- J/\psi) \cdot \mathcal{B}(J/\psi \to e^+e^-)}{\mathcal{B}(\psi(2S) \to e^+e^-)}$ is centred at the values reported in Ref. [129], but its width corresponds to the central value increased by an arbitrary factor 1.5. This last choice allows the background to float almost freely in data within a factor two or three from its central value, without allowing the fit to confuse this component with the $B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$ signal decays.



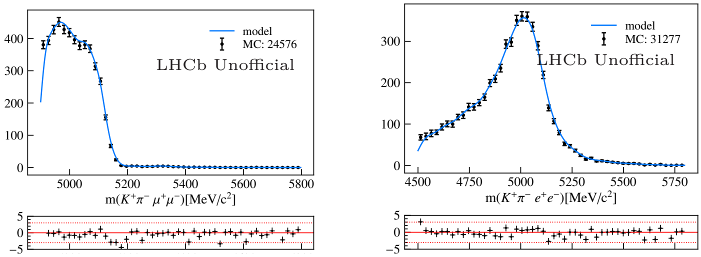
(a) $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ and $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$



(b) $B_s^0 \to \phi J/\psi(\to \mu^+\mu^-)$ and $B_s^0 \to \phi J/\psi(\to e^+e^-)$

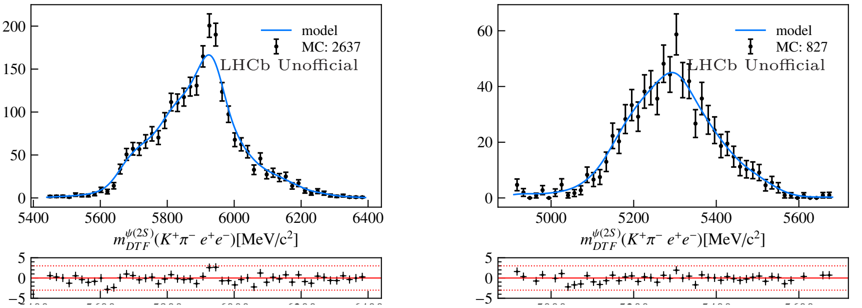


(c) K $-\ \pi$ swap



(d) $B \to XJ/\psi(\to \mu^+\mu^-)$ and $B \to XJ/\psi(\to e^+e^-)$

Figure 8.3: Comparison between the simulated distribution and the binned KDE parametrization for some of the backgrounds discussed in Sec. 8.3.1. The lack of agreement at the borders of the range shown is expected and does not affect the region where the parametrization is used in fits.



(a) $B^0 \to K^{*0} J/\psi(\to e^+e^-)$ leakage (b) $B^0 \to K^{*0}\psi(2S)(\to \pi^+\pi^- J/\psi(\to e^+e^-))$

Figure 8.4: Binned KDE parametrization, in blue, and the simulated distribution, in black, for $B^0 \to K^{*0} J/\psi(\to e^+e^-)$ leakage and $B^0 \to K^{*0}\psi(2S)(\to \pi^+\pi^- J/\psi(\to e^+e^-))$ events reconstructed as signal for Run 1 in the L0I trigger category. Both of them are parametrized for $m_{DTF}^{\psi(2S)}(K\pi e^+e^-)$.

8.3.2 Simultaneous data fits to $B^0 \to K^{*0} J/\psi(\to \ell^+\ell^-)$ decays in $m(K\pi\ell^+\ell^-)$

Simultaneous fits to $B^0 \to K^{*0} J/\psi(\to \mu^+\mu^-)$ and $B^0 \to K^{*0} J/\psi(\to e^+e^-)$ decay candidates are performed in the invariant mass $m(K\pi\ell^+\ell^-)$, separately for each run and L0I and L0L! trigger category, in order to extract the corresponding signal yields for electrons and muons. The choice to perform these fits simultaneously between electrons and muons is justified by the possibility to share the Gaussian constraints introduced in Sec. 8.3.1 and the fraction of B_s^0 between the two modes, further stabilizing the fit. The two channels are selected according to the full selection described in Sec. 5.2 and requiring in particular:

- for $B^0 \to K^{*0} J/\psi(\to \mu^+\mu^-)$ decays, an invariant mass range from 5100 to 5800 MeV/c^2, a q^2 range between 8 and 11 GeV^2/c^4 (corresponding to the J/ψ-q^2 region) and a cut on the combinatorial MVA greater than 0.2;
- for $B^0 \to K^{*0} J/\psi(\to e^+e^-)$ decays, an invariant mass range from 4600 to 5800 MeV/c^2, a q^2 range between 7 and 11 GeV^2/c^4 (corresponding to the J/ψ-q_c^2 region) and a cut on the combinatorial MVA greater than 0.2.

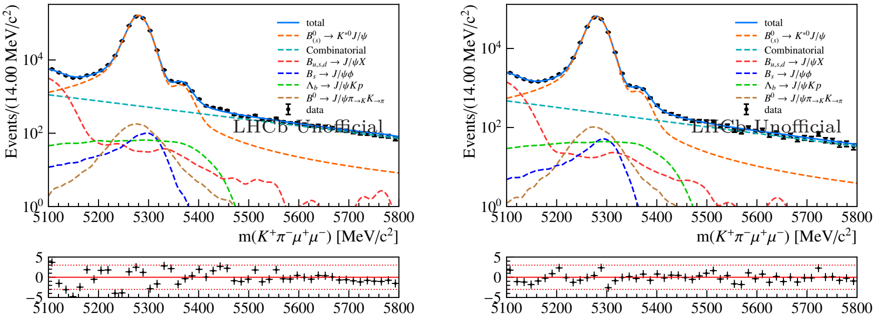
The background components included are the same between the two channels, and correspond to combinatorial, $\Lambda_b^0 \to pK J/\psi(\to \ell^+\ell^-)$, $B_s^0 \to K^{*0} J/\psi(\to \ell^+\ell^-)$, $B_s^0 \to \phi J/\psi(\to \ell^+\ell^-)$, K − π swap and partially reconstructed decays $B \to X J/\psi(\to \ell^+\ell^-)$. Table 8.1 recaps the parameters floating in the fit, specifying which ones float and which ones are Gaussian constrained. The result of the data fits is shown in Fig. 8.5, while the yields extracted are summarized in Table 8.2. An overall agreement can be seen between the fit results and the data.

Table 8.1: Summary of the parameters floated in the $m(K\pi\ell^+\ell^-)$ fits to $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays.

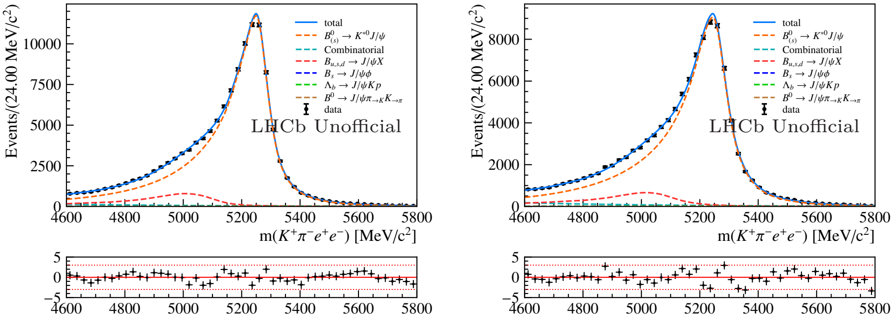
	Parameters
Floating	• s^{μ}_{σ}, $\Delta\mu^{\mu}$, s^{e}_{σ}, $\Delta\mu^{e}$, $s^{e}_{\alpha_R}$, • $f_{B^0_s}$, • $\mathcal{N}_{B^0\to K^{*0}J/\psi(\to\mu^+\mu^-)}$, $\mathcal{N}_{comb}$, $\mathcal{N}_{B\to XJ/\psi(\to\mu^+\mu^-)}$, • $\mathcal{N}_{B^0\to K^{*0}J/\psi(\to e^+e^-)}$, $\mathcal{N}_{comb}$, $\mathcal{N}_{B\to XJ/\psi(\to e^+e^-)}$, • λ^{μ}, λ^{e}
Gaussian constr.	• $\frac{f_{\Lambda_b}}{f_d}$, $\frac{f_s}{f_d}$, • $\frac{\mathcal{B}(\Lambda_b^0\to pKJ/\psi(\to\ell^+\ell^-))}{\mathcal{B}(B^0\to K^{*0}J/\psi(\to\ell^+\ell^-))}$, • $\frac{\mathcal{B}(B_s^0\to\phi J/\psi(\to\mu^+\mu^-))}{\mathcal{B}(B^0\to K^{*0}J/\psi(\to e^+e^-))}$, • $\epsilon(\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-))$, $\epsilon(\Lambda_b^0 \to pKJ/\psi(\to e^+e^-))$, • $\epsilon(B_s^0 \to \phi J/\psi(\to \mu^+\mu^-))$, $\epsilon(B_s^0 \to \phi J/\psi(\to e^+e^-))$, • $\epsilon(\mathrm{K}-\pi\,\mathrm{swap})$, $\epsilon(\mathrm{K}-\pi\,\mathrm{swap})$, • $\epsilon(B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-))$, $\epsilon(B^0 \to K^{*0}J/\psi(\to e^+e^-))$

Table 8.2: Measured values of yields for $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ in the J/ψ-q^2 region and $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ in the J/ψ-q^2_c region.

Channel	Run	L0L!	L0I
$B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$	Run 1	177742 ± 467	67913 ± 288
	Run 2	529707 ± 789	215941 ± 507
$B^0 \to K^{*0}J/\psi(\to e^+e^-)$	Run 1	29455 ± 367	25465 ± 343
	Run 2	111333 ± 619	95677 ± 683



(a) Fits to $B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$ data candidats in $m(K\pi\mu\mu)$



(b) Fits to $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ data candidates in $m(K\pi ee)$

Figure 8.5: Fits to $B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$ (a) and $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ (b) decay candidates for the L0L! (left) and L0I (right) trigger category. The samples correspond to the data collected in Run 2.

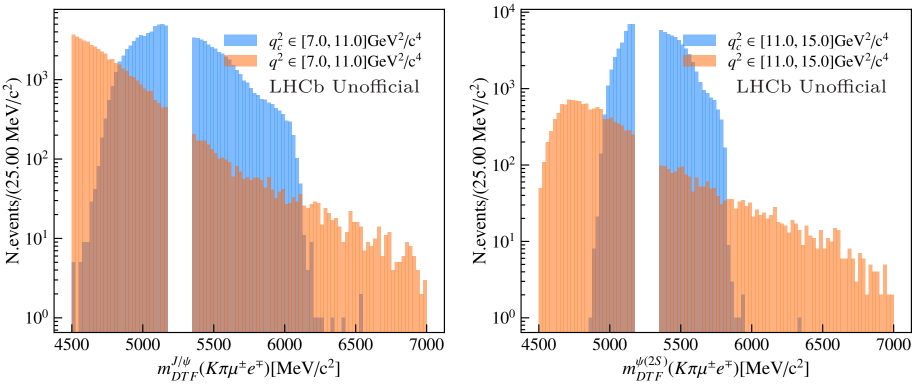
8.3.3 Simultaneous data fit to $B^0 \to K^{*0}\psi_n(\to \ell^+\ell^-)$ decays in $m_{DTF}^{\psi_n}(K\pi\ell^+\ell^-)$, with $\psi_n = J/\psi\,, \psi(2S)$

The simultaneous fits to $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays are performed in the invariant masses $m_{DTF}^{J/\psi}$ and $m_{DTF}^{\psi(2S)}$, respectively, separately for each run and combination of trigger categories. The datasets used have passed all selections and, for the $B^0 \to K^{*0}J/\psi\,(\to \ell^+\ell^-)$ channel, are additionally required to have a $m_{DTF}^{\psi(2S)}$ that is more than 200 MeV/c^2 away from the nominal mass of the B^0, needed to suppress some of the partially reconstructed contributions that are more difficult to model (see Sec. 5.2.5). The four channels are then selected requiring:

- for $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ decays, an invariant mass range from 5100 to 6200 MeV/c^2, a q^2 range between 8.98 and 10.22 GeV^2/c^4 and a cut on the combinatorial MVA greater than 0.2;

- for $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays, an invariant mass range from 4900 to 6200 MeV/c^2, a q^2 range between 6 and 11 GeV$^2/c^4$ and a cut on the combinatorial MVA greater than 0.2;
- for $B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-)$ decays, an invariant mass range from 5100 to 5750 MeV/c^2, a q^2 range between 12.86 and 14.33 GeV$^2/c^4$ and a cut on the combinatorial MVA greater than 0.2;
- for $B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$ decays, an invariant mass range from 5100 to 5950 MeV/c^2, a q^2 range between 11 and 15 GeV$^2/c^4$ and a cut on the combinatorial MVA greater than 0.3.

The choice of using regions of dilepton invariant mass squared based on the q^2 variable rather than q^2_c for the electron channels, is justified by the distortion that such a selection would cause on the invariant masses $m(K\pi ee)^{J/\psi}_{DTF}$ and $m(K\pi ee)^{\psi(2S)}_{DTF}$ for the combinatorial background component. The effect of this distortion can be seen when using the LFV sample after the full chain of selections as a proxy for the combinatorial background in data, as shown in Figure 8.6. When selecting on q^2_c, the invariant mass distribution of the combinatorial proxy is greatly distorted, with almost no region of the spectrum that follows an exponential behaviour. The effect is particularly strong for $q^2_c \in [11, 15]$ GeV$^2/c^4$. When selecting in q^2, instead, the combinatorial proxy behaves exponentially in the region used for the fits described here. This different behaviour justifies the choice to use q^2 rather than q^2_c to select the electron modes.



(a) (b)

Figure 8.6: Effect of the choice of applying a dilepton invariant mass squared selection on q^2 or q^2_c for the invariant masses $m(K\pi ee)^{J/\psi}_{DTF}$ and $m(K\pi ee)^{\psi(2S)}_{DTF}$. The samples used correspond to the reconstructed decays of $B^0 \to K^{*0}\mu^\pm e^\mp$ after the full selection. The region corresponding a real $B^0 \to K^{*0}\mu^\pm e^\mp$ decay peak is not shown.

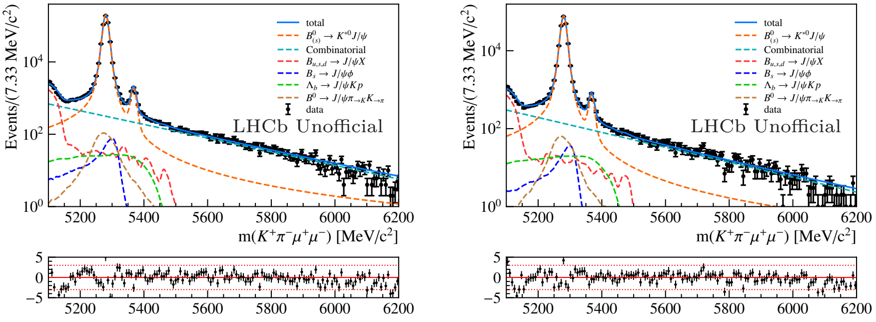
Background components such as combinatorial, $\Lambda_b^0 \to pK\psi_n(\to \ell^+\ell^-)$ decays, $B_s^0 \to K^{*0}\psi_n(\to \ell^+\ell^-)$ decays, $\mathrm{K}-\pi$ swap are present in each of the four channels. The $B_s^0 \to \phi J/\psi(\to \ell^+\ell^-)$ components is included only in $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ fits. Partially reconstructed decays are not included when fitting $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays due to their negligible contribution if added, but they are included as a single component in $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ decays and as two separate components in $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays: $B \to XJ/\psi$ and $B \to Y(\to J/\psi X)Z$. The contributions coming from the $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ leakage and partially reconstructed $B^0 \to K^{*0}\psi(2S)(\to XJ/\psi(\to e^+e^-))$ decays are considered only in $B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$ fits. Table 8.3 recaps the parameters floating in the fit, specifying which ones float and which ones are Gaussian constrained. The result of the data fits are shown Figs. 8.7 and 8.8, while the yields extracted are summarized in Table 8.4. An overall good agreement is observed between the fit results and the data, with a main exception: some discrepancy with data can be noticed in the mass regions on the left of the signal peak. These discrepancies are expected to have a minor effect on the values of the extracted yields, and can be attributed to the complex composition of the partially reconstructed backgrounds, for which an exact parametrization from simulation is very difficult to achieve.

Table 8.3: Summary of the parameters floated in the $m_{DTF}^{J/\psi}(K\pi\ell^+\ell^-)$ fits to $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays and $m_{DTF}^{\psi(2S)}(K\pi\ell^+\ell^-)$ fits to $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays.

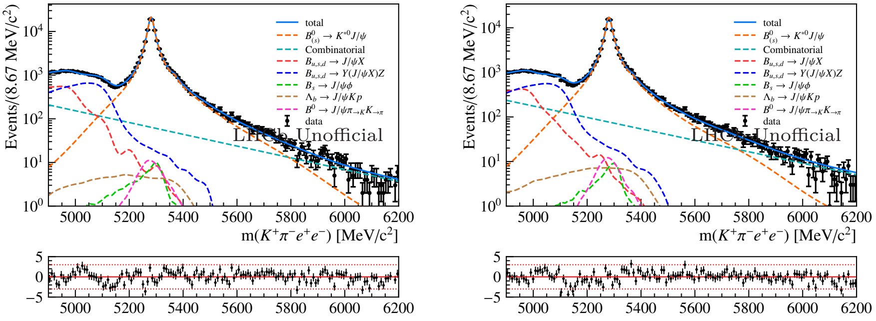
	Parameters
Floating	• s_σ^μ, $\Delta\mu^\mu$,s_σ^e, $\Delta\mu^e$, s_σ^e, $\Delta\mu^e$, • $f_{B_s^0}$, • $\mathcal{N}_{B^0\to K^{*0}J/\psi(\to\mu^+\mu^-)}$,$\mathcal{N}_{comb}$, $\mathcal{N}_{B\to X\,J/\psi(\to\mu^+\mu^-)}$, • $\mathcal{N}_{B^0\to K^{*0}J/\psi(\to e^+e^-)}$, $\mathcal{N}_{comb}$, • $\mathcal{N}_{B\to X\,J/\psi(\to e^+e^-)}$, $\mathcal{N}_{B\to Y(J/\psi X)Z}$, • λ^μ, λ^e , • $\mathcal{N}_{B^0\to K^{*0}\psi(2S)(\to\mu^+\mu^-)}$,$\mathcal{N}_{comb}$, • $\mathcal{N}_{B^0\to K^{*0}\psi(2S)(\to e^+e^-)}$, $\mathcal{N}_{comb}$, • $\mathcal{N}_{J/\psi\ \text{leakage}}$, • λ^μ, λ^e ,
Gaussian constr.	• $\frac{f_{\Lambda_b}}{f_d}$, $\frac{f_s}{f_d}$, • $\frac{\mathcal{B}(\Lambda_b^0\to pK\,J/\psi(\to\ell^+\ell^-))}{\mathcal{B}(B^0\to K^{*0}J/\psi(\to\ell^+\ell^-))}$, $\frac{\mathcal{B}(B_s^0\to\phi J/\psi(\to\ell^+\ell^-))}{\mathcal{B}(B^0\to K^{*0}J/\psi(\to\ell^+\ell^-))}$, • $\epsilon(\Lambda_b^0 \to pK\,J/\psi(\to \mu^+\mu^-))$, $\epsilon(\Lambda_b^0 \to pK\,J/\psi(\to e^+e^-))$, • $\epsilon(B_s^0 \to \phi J/\psi(\to \mu^+\mu^-))$, $\epsilon(B_s^0 \to \phi J/\psi(\to e^+e^-))$, • $\epsilon(\mathrm{K}-\pi\,\mathrm{swap})$, $\epsilon(\mathrm{K}-\pi\,\mathrm{swap})$, • $\epsilon(B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-))$, $\epsilon(B^0 \to K^{*0}J/\psi(\to e^+e^-))$, • $\frac{\mathcal{B}(\Lambda_b^0\to pK\psi(2S)(\to\ell^+\ell^-))}{\mathcal{B}(B^0\to K^{*0}\psi(2S)(\to\ell^+\ell^-))}$, $\frac{\mathcal{B}(B_s^0\to\phi\psi(2S)(\to\ell^+\ell^-))}{\mathcal{B}(B^0\to K^{*0}\psi(2S)(\to\ell^+\ell^-))}$, • $\frac{\mathcal{B}(B^0\to K^{*0}\psi(2S)(\to\pi^+\pi^-\,J/\psi(\to\ell^+\ell^-)))}{\mathcal{B}(B^0\to K^{*0}\psi(2S)(\to\ell^+\ell^-))}$, • $\epsilon(\Lambda_b^0 \to pK\psi(2S)(\to \mu^+\mu^-))$, $\epsilon(\Lambda_b^0 \to pK\psi(2S)(\to e^+e^-))$, • $\epsilon(\mathrm{K}-\pi\,\mathrm{swap})$, $\epsilon(\mathrm{K}-\pi\,\mathrm{swap})$, • $\epsilon(B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-))$, $\epsilon(B^0 \to K^{*0}\psi(2S)(\to e^+e^-))$, • $\epsilon(B^0 \to K^{*0}\psi(2S)(\to \pi^+\pi^-\,J/\psi(\to \ell^+\ell^-)))$,

Table 8.4: Measured values of yields for $B^0 \to K^{*0}\psi_n(\to \ell^+\ell^-)$ fits in $m_{DTF}^{\psi_n}$ with $\psi_n = J/\psi\,, \psi(2S)$.

Channel	Run	L0L!	L0I
$B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$	RUN 1	174972 ± 442	66795 ± 267
	RUN 2	521588 ± 732	212355 ± 468
$B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$	RUN 1	27394 ± 215	24715 ± 221
	RUN 2	108108 ± 394	93006 ± 396
$B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-)$	RUN 1	10675 ± 111	4075 ± 65
	RUN 2	31628 ± 183	12382 ± 117
$B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$	RUN 1	1714 ± 43	1243 ± 45
	RUN 2	6444 ± 91	4846 ± 85

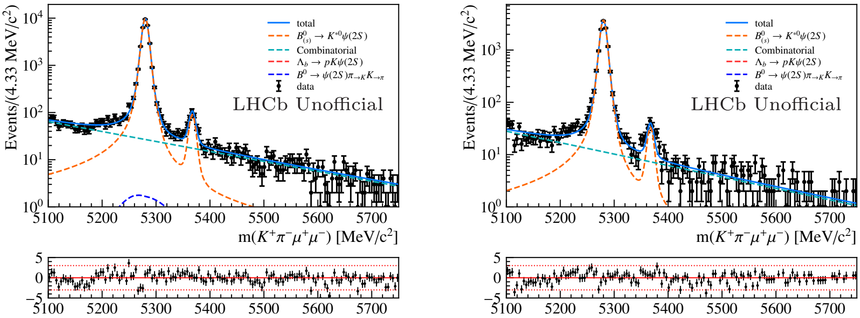


(a) Fits to $B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$ candidates in $m(K\pi\mu\mu)_{DTF}^{J/\psi}$

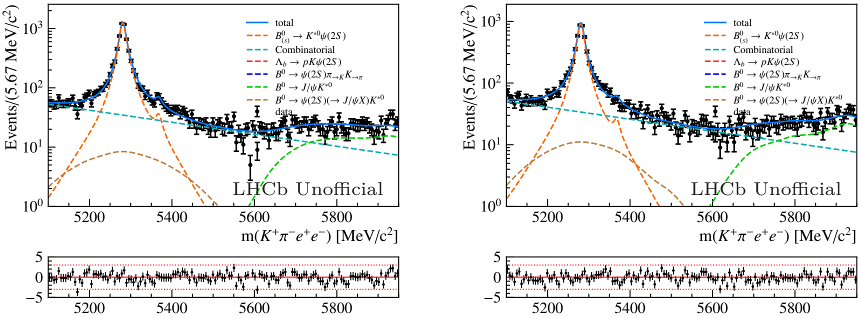


(b) Fits to $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ data candidates in $m(K\pi ee)_{DTF}^{J/\psi}$

Figure 8.7: Fits to $B^0 \to K^{*0}J/\psi\,(\to \mu^+\mu^-)$ (a) and $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ (b) decay candidates for the L0L! (left) and L0I (right) trigger category. The samples correspond to the data collected in RUN 2.



(a) Fits to $B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-)$ data candidates in $m(K\pi\mu\mu)^{\psi(2S)}_{DTF}$



(b) Fits to $B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$ data candidates in $m(K\pi ee)^{\psi(2S)}_{DTF}$

Figure 8.8: Fits to $B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-)$ (a) and $B^0 \to K^{*0}\psi(2S)(\to e^+e^-)$ (b) decay candidates for the L0L! (left) and L0I (right) trigger category. The samples correspond to the data collected in Run 2.

Chapter 9

LFU cross-checks

The analysis strategy outlined in Ch. 4 and described in detail in Ch. 10 relies on the assumption that the understanding of the signal total efficiency and acceptances obtained from simulation can be trusted to recover the original properties of the underlying decay, without introducing any spurious lepton flavour non-universal behaviour. To verify this, two sets of checks on the reconstructed data candidates of $B^0 \to K^{*0} J/\psi (\to \mu^+\mu^-)$ and $B^0 \to K^{*0} J/\psi (\to e^+e^-)$ decays are presented in the following. The first set of checks, discussed in Sec. 9.1, looks for sources of potential LFU-breaking introduced by the use of the yields of the $B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$ decays to compute the constraint on the expected number of rare signal events. The second set of checks, discussed in Sec. 9.2, is focused on the possible sources of lepton flavour non-universality introduced by the procedure used for the amplitude fit. The same framework of signal *pdf* and acceptance parametrization is used, with minor modifications, to perform an amplitude fit of the $B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$ decays in data. Since no difference is expected in the description of the electron and muon resonant channel, the compatibility between the observables of the two modes ensures the reliability of the procedure proposed.

9.1 LFU cross-checks on ratios of branching fractions

The following section presents three cross-checks that aim to test the ability of the correction chain to reliably correct for data-simulation differences in the regions of q^2 dominated by the charmonium resonances. The first test, presented in Secs. 9.1.1, corresponds to the ratio of the estimated branching ratios of $B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$ decays between muons and electrons. It is important to stress that, since the electrons and muons behave so differently, such a check is far from trivial and requires a remarkable control of many aspects of the analysis. The control is thought to be satisfactory when the ratio $r_{J/\psi}$ is compatible with one, its expected value in the SM. The second check, described in Sec. 9.1.2, is the measurement of the ratio of branching fractions between $\mathcal{B}(B^0 \to K^{*0}\psi(2S))$ and $\mathcal{B}(B^0 \to K^{*0}J/\psi)$ decays, obtained separately for the decays of the two resonances into electrons or muons. This type of ratio is quite interesting because, when $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays are substituted with $B^0 \to K^{*0}\ell^+\ell^-$, it corresponds to one of the main ingredients in the constraint of the expected yield used in the final data fit, as discussed in Sec. 9.1.2. The last test, described in Sec 9.1.3, corresponds to the ratio of the quantities obtained in the second check, when they are computed using muons and electrons. While this corresponds just to a further cross-check, when substituting again the decay of the resonant $\psi(2S)$ with the rare mode, such a quantity will correspond to the LFU test of the SM called R_{K^*}. Such a quantity plays an important role in this analysis, since any difference in the branching ratios of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays will shift $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ from their SM value of zero.

9.1.1 Integrated $r_{J/\psi}$

The single ratio $r_{J/\psi}$, defined as

$$r_{J/\psi} = \frac{N_{B^0\to K^{*0}J/\psi(\to\mu^+\mu^-)}}{N_{B^0\to K^{*0}J/\psi(\to e^+e^-)}} \cdot \frac{\epsilon^{tot}_{B^0\to K^{*0}J/\psi(\to e^+e^-)}}{\epsilon^{tot}_{B^0\to K^{*0}J/\psi(\to\mu^+\mu^-)}} \quad , \tag{9.1}$$

is well-known in the SM and, due to LFU in this class of decays, it is expected be one to high degree of precision. Since fits to $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays are used as a reference to extract the branching ratio of the rare mode in the additional constraint to the likelihood (see Sec. 10.3), it is important to verify that no lepton flavour non-universal behaviour is introduced at this step. The yields $N_{B^0\to K^{*0}J/\psi\,(\to\ell^+\ell^-)}$ are extracted using $m(K\pi\ell\ell)$ fits, as described in Sec. 8.3.2 and reported in Table 8.2, while the efficiencies $\epsilon^{tot}_{B^0\to K^{*0}J/\psi\,(\to\ell^+\ell^-)}$ are obtained using simulated samples with the nominal correction chain applied, as described in Sec. 7.1 and reported in Table 7.1. The cross-check is performed in each of the two runs and trigger categories L0I and L0L! and shown in Table 9.1. The two uncertainties correspond to the statistical uncertainty on the extracted yields and the bootstrapped systematic uncertainty obtained from the estimation of the efficiencies. The latter is expected to be a small fraction of the overall systematic uncertainty, which was evaluated be around 2% in Ref. [128]. As it can be seen, all the measured values agree with unity within 4%. Moreover, the combination of different runs and trigger categories shows a compatibility with unity within a single statistical standard deviation. This test confirms that no lepton flavour non-universal behaviour is introduced by using the channels $B^0 \to K^{*0}J/\psi\,(\to \ell^+\ell^-)$ to calibrate the expected number of signal events in the final fit.

Table 9.1: Measured values of $r_{J/\psi}$ obtained from fits to the invariant mass $m(K\pi\ell^+\ell^-)$ (see Sec. 8.3.2).

Type	L0M!/L0E!	L0I/L0I	All triggers
Run 1	0.977 ± 0.012 ± 0.003	1.036 ± 0.015 ± 0.004	1.002 ± 0.009 ± 0.002
Run 2	0.995 ± 0.006 ± 0.003	0.999 ± 0.008 ± 0.003	0.997 ± 0.005 ± 0.003
All Runs	0.986 ± 0.005 ± 0.002	1.016 ± 0.007 ± 0.003	0.999 ± 0.004 ± 0.002

For completeness, a similar cross-check can be performed in the fits to $m(K\pi\ell\ell)^{J/\psi}_{DTF}$ and $m(K\pi\ell\ell)^{\psi(2S)}_{DTF}$ invariant masses described in Sec. 8.3.3, using the measured yields of Table 8.4 and the efficiencies reported in Table 7.2. The result of this cross-check is shown in Table 9.2. All the values obtained agree with unity within 5%, however, in this case the combination of the different runs and trigger categories is compatible with one only when considering the expected systematic uncertainty computed in Ref. [128]. While the difference between the two LFU test discussed could be used to assign an additional systematic based on the fit strategy chosen, this deviation does not enter directly in our final measurement and is only used in combination with $r_{\psi(2S)}$ to obtain the cross-check $R_{\Psi(2S)}$, discussed in Sec. 9.1.3. For this reason, no further studies on the origins of the deviation observed were considered.

Table 9.2: Measured values of $r_{J/\psi}$ obtained from fits to the invariant mass $m_{DTF}^{J/\psi}$ (see Sec. 8.3.3).

Type	L0M!/L0E!	L0I/L0I	All triggers
Run 1	$1.039 \pm 0.009 \pm 0.003$	$1.049 \pm 0.010 \pm 0.004$	$1.043 \pm 0.007 \pm 0.003$
Run 2	$1.012 \pm 0.004 \pm 0.003$	$1.014 \pm 0.005 \pm 0.003$	$1.013 \pm 0.003 \pm 0.003$
All Runs	$1.025 \pm 0.004 \pm 0.002$	$1.029 \pm 0.004 \pm 0.003$	$1.028 \pm 0.003 \pm 0.002$

9.1.2 Integrated $\mathcal{B}(B^0 \to K^{*0}\psi(2S))/\mathcal{B}(B^0 \to K^{*0}J/\psi)$

A measurement of the ratio of branching fractions of $B^0 \to K^{*0}\psi(2S)$ and $B^0 \to K^{*0}J/\psi$ decays can be obtained as:

$$\frac{\mathcal{B}(B^0 \to K^{*0}\psi(2S))}{\mathcal{B}(B^0 \to K^{*0}J/\psi)} = \frac{N_{\psi(2S)}}{N_{J/\psi}} \cdot \frac{\mathcal{B}(J/\psi \to \ell^+\ell^-)}{\mathcal{B}(\psi(2S) \to \ell^+\ell^-)} \cdot \frac{1}{R_{\varepsilon}^{\psi(2S)}}\,, \tag{9.2}$$

where $N_{J/\psi}$ and $N_{\psi(2S)}$ are the signal yields obtained from the mass fits of Sec. 8.3.3, $R_{\varepsilon}^{\psi(2S)}$ is the efficiency of $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays relative to the $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ mode, and the ratio [129]

$$\frac{\mathcal{B}(J/\psi \to \ell^+\ell^-)}{\mathcal{B}(\psi(2S) \to \ell^+\ell^-)} = 7.69 \pm 0.19\,, \tag{9.3}$$

assumes LFU in charmonium decays, *i.e.* $\mathcal{B}(J/\psi \to \mu^+\mu^-)/\mathcal{B}(\psi(2S) \to \mu^+\mu^-) = \mathcal{B}(J/\psi \to e^+e^-)/\mathcal{B}(\psi(2S) \to e^+e^-)$.

Figure 9.1 shows the results of Eq. 9.2, separately for the Run 1 and Run 2 datasets in the trigger categories L0I and L0L! for muons and electron, where the uncertainty must be understood as statistical-only, compared to the world average of [129]

$$\frac{\mathcal{B}(B^0 \to K^{*0}\psi(2S))}{\mathcal{B}(B^0 \to K^{*0}J/\psi)}^{[\mathrm{PDG}]} = 0.487 \pm 0.018 \pm 0.011\,. \tag{9.4}$$

The first uncertainty combines the statistical and systematic uncertainties from Ref. [157] and the second isolates the systematic uncertainty due to uncertainty on $\mathcal{B}(J/\psi \to \ell^+\ell^-)/\mathcal{B}(\psi(2S) \to \ell^+\ell^-)$ of Eq. 9.3. Note that the latter uncertainty enters identically both in the world average of Eq. 9.4 and in our determination from Eq. 9.2, and is therefore removed. Finally, the check is based on the mass fit results of Sec. 8.3 and assumes the same fraction of S-wave (F_S) in both $B^0 \to K^{*0}J/\psi$ and $B^0 \to K^{*0}\psi(2S)$ decays. The excellent agreement between the two datasets and the world average fully accomplishes the purpose of this cross-check.

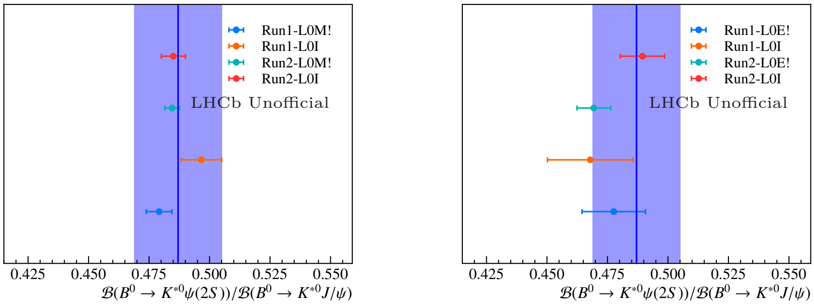


Figure 9.1: Determination of the ratio of branching fraction $\mathcal{B}(B^0 \to K^{*0}\psi(2S))/\mathcal{B}(B^0 \to K^{*0}J/\psi)$ in muons and electrons respectively, separately for the Run 1 and Run 2 datasets in the trigger categories L0I and L0L!. The blue shaded area corresponds to PDG value of Ref. [129].

9.1.3 Integrated $R_{\psi(2S)}$

The double ratio $R_{\psi(2S)}$ is defined as

$$
\begin{aligned}
R_{\psi(2S)} = & \frac{N_{B^0\to K^{*0}\psi(2S)(\to\mu^+\mu^-)}}{N_{B^0\to K^{*0}\psi(2S)(\to e^+e^-)}} \cdot \frac{N_{B^0\to K^{*0}J/\psi(\to e^+e^-)}}{N_{B^0\to K^{*0}J/\psi(\to\mu^+\mu^-)}} \\
& \times \frac{\epsilon^{tot}_{B^0\to K^{*0}\psi(2S)(\to e^+e^-)}}{\epsilon^{tot}_{B^0\to K^{*0}\psi(2S)(\to\mu^+\mu^-)}} \cdot \frac{\epsilon^{tot}_{B^0\to K^{*0}J/\psi(\to\mu^+\mu^-)}}{\epsilon^{tot}_{B^0\to K^{*0}J/\psi(\to e^+e^-)}} \, .
\end{aligned} \tag{9.5}
$$

This corresponds to the ratio of ratios $R_{\psi(2S)} = r_{\psi(2S)}/r_{J/\psi}$ or to $R_{\psi(2S)} = \frac{\mathcal{B}(B^0\to K^{*0}\psi(2S)(\to\mu^+\mu^-))}{\mathcal{B}(B^0\to K^{*0}J/\psi(\to\mu^+\mu^-))} \times \frac{\mathcal{B}(B^0\to K^{*0}J/\psi(\to e^+e^-))}{\mathcal{B}(B^0\to K^{*0}\psi(2S)(\to e^+e^-))}$, where the single ratio $r_{\psi(2S)}$ is evaluated on $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays, similarly to what is done for $r_{J/\psi}$ in $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$. Thanks to the double ratio, this check is particularly insensitive to systematic uncertainties that affect the single ratios and that were ignored in the Tables 9.1 and 9.2. This is because, when the same procedure is used to parametrize the signal in the fits and to correct the efficiencies in $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$ decays, most of the systematic uncertainties are common between the two modes, shifting the two channels similarly but leaving the ratio largely unchanged. The ratios obtained for $r_{\psi(2S)}$ and $R_{\psi(2S)}$ are reported in Table 9.3 and Table 9.4, respectively. As can be seen, while the values of $r_{\psi(2S)}$ are systematically larger than one and the combination of different runs and trigger categories is incompatible with unity, the values obtained for the double ratio $R_{\psi(2S)}$ are in general closer to unity with the final weighted average that is within two standard deviations from one. The compatibility of the last result with its expected value, suggests that the use of the correction chain developed in Ch. 6 in the $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays can be extended successfully to other regions of q^2 without the introduction of spurious non-LFU behaviours.

Table 9.3: Measured values of $r_{\Psi(2S)}$ obtained from fits to the invariant mass $m(K\pi\ell\ell)_{DTF}^{\psi(2S)}$ (see Sec. 8.3.3).

Type	L0M!/L0E!	L0I/L0I	All triggers
Run 1	1.042 ± 0.028 ± 0.007	1.114 ± 0.044 ± 0.008	1.074 ± 0.024 ± 0.006
Run 2	1.045 ± 0.016 ± 0.004	1.004 ± 0.020 ± 0.004	1.027 ± 0.012 ± 0.003
All Runs	1.045 ± 0.014 ± 0.003	1.031 ± 0.018 ± 0.003	1.036 ± 0.011 ± 0.002

Table 9.4: Measured values of $R_{\Psi(2S)}$ obtained from fits to the invariant masses $m(K\pi\ell\ell)_{DTF}^{J/\psi}$ and $m(K\pi\ell\ell)_{DTF}^{\psi(2S)}$ (see Sec. 8.3.3).

Type	L0M!/L0E!	L0I/L0I	All triggers
Run 1	1.003 ± 0.029 ± 0.008	1.061 ± 0.043 ± 0.009	1.029 ± 0.024 ± 0.006
Run 2	1.032 ± 0.016 ± 0.004	0.991 ± 0.020 ± 0.005	1.015 ± 0.013 ± 0.003
All Runs	1.027 ± 0.014 ± 0.004	1.010 ± 0.018 ± 0.004	1.019 ± 0.011 ± 0.003

9.2 LFU cross-check on the fit procedure: amplitude fits to $B^0 \to K^{*0} J/\psi$ decays

The control mode $B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$ is an ideal candidate to validate the amplitude fit strategy discussed in Ch. 10 for the rare mode. Having the same final states of $B^0 \to K^{*0}\ell^+\ell^-$ decays, it is characterized by very similar efficiency and reconstruction effects. At the same time, occurring via a tree-level transition, it is expected to be significantly more abundant than the rare mode, providing enough statistics to validate the proposed fitting strategy. The muon mode has also been studied extensively in the past via several angular analyses, all of which found results consistent with the SM [158–161]. Besides published results, this mode was used for validation purposes as part of the angular analyses of $B^0 \to K^{*0}\mu^+\mu^-$ decays [29]. Table 9.5 reports the result obtained by LHCb amplitude analysis with 1 fb^{-1} of data taken in 2011 [158] and the set of internal results obtained from Ref. [29]. The compatibility between the observables obtained for the muon and electron channels, together with the compatibility with the results previously obtained from the literature, is used to determine if the check can be considered passed or not.

9.2.1 Fit generalities

All amplitude fits to the $B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$ decay modes are obtained through the maximization of an unbinned extended maximum likelihood $\mathcal{L}$, defined as in Eq. 10.35, where the term x corresponds to the five-dimensional vector $x \equiv \{\cos\theta_K, \cos\theta_\ell, \phi, m(K\pi\ell\ell), m(K\pi)\}$, needed to fully describe the

	LHCb 1fb^{-1}	$B^0 \to K^{*0} J/\psi(\to \mu^+\mu^-)$
	$m_{K\pi} \in [825.9, 965.9]\mathrm{MeV/c}^2$	Simultaneous fit Ref. [29]
F_L	0.572 ± 0.008	0.5629 ± 0.0010
S_3	-0.013 ± 0.010	-0.0004 ± 0.0013
S_4	-0.250 ± 0.006	-0.2507 ± 0.0015
S_5	–	0.0003 ± 0.0014
A_{FB}	–	0.0006 ± 0.0009
S_7	–	0.0016 ± 0.0015
S_8	-0.048 ± 0.007	-0.0547 ± 0.0015
S_9	-0.084 ± 0.006	-0.0863 ± 0.0013
F_S	0.064 ± 0.010	0.0554 ± 0.0019
S_{S_1}	–	-0.2273 ± 0.0026
S_{S_2}	–	0.0261 ± 0.0017
S_{S_3}	–	0.0014 ± 0.0015
S_{S_4}	–	0.0002 ± 0.0016
S_{S_5}	–	-0.0635 ± 0.0017
$\lvert A_\parallel\rvert^2$	0.227 ± 0.004 (0.011)	–
$\lvert A_\perp\rvert^2$	0.201 ± 0.004 (0.008)	–
$\delta_\parallel$	-2.94 ± 0.02 (0.03)	–
$\delta_\perp$	2.94 ± 0.02 (0.02)	–
$\lvert A_S\rvert^2_{\Delta_1}$	0.115 ± 0.007 (0.020)	0.1234 ± 0.0072
$\lvert A_S\rvert^2_{\Delta_2}$	0.049 ± 0.004 (0.007)	0.0454 ± 0.0036
$\lvert A_S\rvert^2_{\Delta_3}$	0.052 ± 0.006 (0.009)	0.0531 ± 0.0034
$\lvert A_S\rvert^2_{\Delta_4}$	0.105 ± 0.014 (0.008)	0.1142 ± 0.0065
$\delta_S^{\Delta_1}$	3.09 ± 0.10 (0.08)	–
$\delta_S^{\Delta_2}$	2.66 ± 0.06 (0.06)	–
$\delta_S^{\Delta_3}$	1.94 ± 0.03 (0.09)	–
$\delta_S^{\Delta_4}$	1.53 ± 0.03 (0.11)	–

Table 9.5: Summary table of external results discussed in this section. The first column contains the results of the amplitude fit to $B^0 \to K^{*0} J/\psi(\to \mu^+\mu^-)$ [158] (on the bottom) and the translation to the S-basis angular observables (on the top). The set of values shown correspond to the $m_{K\pi}$ range of 825.9 to 965.9 MeV/c^2, with the exception of the $|A_S|^2$ and δ_S that are given in four bins of $m_{K\pi}$: [826, 861] MeV/c^2, [861, 896] MeV/c^2, [896, 931] MeV/c^2 and [931, 966] MeV/c^2. We will refer to these bins as Δ_i, where the index i goes from 1 to 4. The P-wave observables S_5, A_{FB} and S_7 are zero: this is a consequence of the purely vectorial nature of the amplitudes that describe the decay. The second column contains a cross-check results obtained as part of the $B^0 \to K^{*}\mu^+\mu^-$ angular analysis for Run 1 and 2016 datasets in the same $m_{K\pi}$ bin [29]. The values reported have been obtained from the internal documentation of the LHCb experiment.

decay. The total *pdf* can be described as a linear composition of signal, $\mathcal{P}_{sig}(x;\theta)$, and background components, $\mathcal{P}_k(x;\theta)$, as already discussed in Eq. 8.2.

The signal is parametrized with the *pdf*:

$$\mathcal{P}_{sig}(x;\theta) = \mathcal{N}_{norm} \cdot \Big(\frac{\mathrm{d}^4\Gamma[\bar{B}^0 \to K^+\pi^- J/\psi]}{\mathrm{d}m_{K\pi}\mathrm{d}\vec{\Omega}} \cdot \mathcal{S}^{\ell}(m_{K\pi\ell^+\ell^-}) \cdot \epsilon_{eff}(\vec{\Omega}) \Big) , \quad (9.6)$$

where:

- $\frac{\mathrm{d}^4\Gamma[\bar{B}^0 \to K^+\pi^- J/\psi]}{\mathrm{d}m_{K\pi}\,\mathrm{d}\vec{\Omega}}$ corresponds the differential decay amplitude of the decay defined in Eq. 9.9 of Sec. 9.2.2;
- $\mathcal{S}^{\ell}(m_{K\pi\ell^+\ell^-})$ is the invariant mass distribution for $\ell = e, \mu$, parametrized as described in Sec. 8.2;
- $\epsilon_{eff}(\vec{\Omega})$ is the three-dimensional effective acceptance described in Sec. 7.2, representing the distortion of the theory *pdf* due to the details of selection and reconstruction;
- $\mathcal{N}_{norm}$ is a factor representing the analytical normalization of the *pdf*.

The background component $\mathcal{P}_k$ is parametrized assuming factorization between the five variables considered, using the functional forms that better describe the distributions observed. More details on the specific parametrization used in the fits to data are discussed in Sec. 9.2.3 and Appendix. C.

9.2.2 Parametrization of the signal amplitude

The angular structure of $B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$ decays is determined by kinematics and conservation of angular momentum. It is thus possible to obtain its differential decay rate starting from the Eq. 2.14, derived for $B^0 \to K^{*0}\ell^+\ell^-$ decays, and taking the limit for which a vector current dominates the decay. By asking that the transversity amplitudes respect the following relations[1]:

$$\mathcal{A}^L_{0,\perp} = \mathcal{A}^R_{0,\perp} = \frac{i}{\sqrt{2}}\tilde{A}_{0,\perp} \quad \mathcal{A}^L_{\parallel} = \mathcal{A}^R_{\parallel} = \frac{-i}{\sqrt{2}}\tilde{A}_{\parallel} \,, \quad (9.7)$$

---

[1] The i factors are introduced to counteract the -1 terms in Eq. 2.15 and stay explicitly aligned with the definition of transversity amplitudes used in Refs. [162, 163].

the differential decay rate in Eq. 2.14 collapses to:

$$
\begin{aligned}
\frac{\mathrm{d}^3\Gamma[\bar{B}^0 \to \bar{K}^{*0}\ell^+\ell^-]}{\mathrm{d}\vec{\Omega}}\Big|_{\substack{A^L_\lambda = A^R_\lambda \\ \beta_\ell \approx 1}} = & \frac{\mathrm{d}^3\Gamma[\bar{B}^0 \to \bar{K}^{*0}J/\psi(\to \ell^+\ell^-)]}{\mathrm{d}\vec{\Omega}} = \\
= \frac{9}{32\pi}\Big[& \frac{1}{2}(|\tilde{A}_\perp|^2 + |\tilde{A}_\parallel|^2)\sin^2\theta_K(1+\cos^2\theta_\ell) \\
& + 2|\tilde{A}_0|^2\cos^2\theta_K\sin^2\theta_\ell \\
& + \frac{1}{2}(|\tilde{A}_\perp|^2 - |\tilde{A}_\parallel|^2)\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi \\
& + \frac{1}{\sqrt{2}}\mathcal{R}e(\tilde{A}_0\tilde{A}^*_\parallel)\sin 2\theta_K\sin 2\theta_\ell\cos\phi \\
& + \frac{1}{\sqrt{2}}\mathcal{I}m(\tilde{A}_0\tilde{A}^*_\perp)\sin 2\theta_K\sin 2\theta_\ell\sin\phi \\
& + \mathcal{I}m(\tilde{A}_\perp\tilde{A}^*_\parallel)\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big] , \qquad (9.8)
\end{aligned}
$$

where the dependency on the q^2 variable of Eq. 2.14 has been dropped since the function is evaluated at $q^2 = M^2_{J/\psi}$ and $\beta_\ell \approx 1$.

In fits to data, the contribution from decays where the $K\pi$ system is produced with zero angular momentum (S-wave) corresponds to a background that can bias the observables of interest if not taken into account properly. This contribution can be included in the differential decay rate by introducing an additional transversity amplitude, $\tilde{A}_{S0}$, and the two functions, $A_P(m_{K\pi})$ and $A_S(m_{K\pi})$, needed to parametrize the distribution of the S- and P-wave in $m(K\pi)$, respectively. The discussion of the exact parametrization of each of these additional functions is deferred to Sec. 10.1, since the same formalism can be used in the decays of $B^0 \to K^{*0}\ell^+\ell^-$. Those results can be extended to $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays by simply performing the substitution $\mathcal{A}^L_{S0} = \mathcal{A}^R_{S0} = \frac{i}{\sqrt{2}}\tilde{A}_{S0}$, analogously to what has been done for the P-wave transversity amplitudes in Eq. 9.7. The full differential decay rate, including the P, S and P-S wave interference terms corresponds to:

$$
\begin{aligned}
\frac{\mathrm{d}^4\Gamma[\bar{B}^0 \to K^+\pi^- J/\psi]}{\mathrm{d}m_{K\pi}\mathrm{d}\vec{\Omega}} = & \Big[|A_P(m_{K\pi})|^2 \cdot \frac{\mathrm{d}^3\Gamma}{\mathrm{d}\vec{\Omega}}\Big|_{\mathrm{P-wave}} \\
& + \frac{2}{3}|\tilde{A}_{S0}A_S(m_{K\pi})|^2\sin^2\theta_\ell \\
& + \frac{4}{\sqrt{3}}\mathcal{R}e\Big[\tilde{A}_{S0}\tilde{A}^*_0 A_S(m_{K\pi})A^*_P(m_{K\pi})\Big]\sin^2\theta_\ell\cos\theta_K \\
& + \sqrt{\frac{2}{3}}\mathcal{R}e\Big[\tilde{A}_{S0}\tilde{A}^*_\parallel A_S(m_{K\pi})A^*_P(m_{K\pi})\Big]\sin 2\theta_\ell\sin\theta_K\cos\phi \\
& + \sqrt{\frac{2}{3}}\mathcal{I}m\Big[\tilde{A}_{S0}\tilde{A}^*_\perp A_S(m_{K\pi})A^*_P(m_{K\pi})\Big]\sin 2\theta_\ell\sin\theta_K\sin\phi\Big] .
\end{aligned}
\qquad (9.9)
$$

The two quantities $\mathcal{A}_\lambda$ and $\tilde{A}_\lambda$ represent the transversity amplitudes that describe the decay of $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$, respectively.

However, the former corresponds to a complex function that depends on the hadronic matrix elements and Wilson coefficients, while the latter correspond to a complex parameter that can be floated or fixed in the fit. When performing fits, the differential decay rate in Eq. 9.9 needs to be normalized into a *pdf*, allowing only three of the four complex transversity amplitudes to float freely in the fit, *i.e.* $\tilde{A}_{\perp,\parallel,S0}$. To simplify the comparison with Ref. [158], these parameters can be combined to construct the four real P-wave observables:

$$|A_{\perp,\parallel}|^2 = \frac{|\tilde{A}_{\perp,\parallel}|^2}{|\tilde{A}_0|^2 + |\tilde{A}_\perp|^2 + |\tilde{A}_\parallel|^2} \quad \text{and} \quad \delta_{\perp,\parallel} = Arg(\tilde{A}_0^*\tilde{A}_{\perp,\parallel})\,, \tag{9.10}$$

and the two real S-wave observables:

$$\begin{aligned} |A_{S0}|^2_{\Delta_{m_{K\pi}}} = F_S\big|_{\Delta_{m_{K\pi}}} &= \frac{|\tilde{A}_{S0}|^2 \int_{\Delta_{m_{K\pi}}} \mathrm{d}m_{K\pi} |A_S(m_{K\pi})|^2}{(|\tilde{A}_0|^2 + |\tilde{A}_\perp|^2 + |\tilde{A}_\parallel|^2)\cdot \int_{\Delta_{m_{K\pi}}} \mathrm{d}m_{K\pi} |A_P(m_{K\pi})|^2} \\ \delta_S^{\Delta_{m_{K\pi}}} &= Arg\Big(\tilde{A}_{S0}\tilde{A}_0^* \cdot \int_{\Delta_{m_{K\pi}}} \mathrm{d}m_{K\pi} A_S(m_{K\pi}) A_P^*(m_{K\pi})\Big)\,. \end{aligned} \tag{9.11}$$

The normalization condition imposes $|A_0|^2 = 1 - |A_\perp|^2 - |A_\parallel|^2$, while, since only phase differences are measurable, no imaginary part is allowed for $\tilde{A}_0$. Starting from the amplitudes $\tilde{A}_\lambda$, the conventional angular observables defined in Sec. 2.3.2 can be obtained as described in Eq. 2.24.

The parametrization described in Eq. 9.9 is tested in reconstructed simulated samples of $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ and $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays. Since these samples are the same ones used to parametrize the acceptances used in the fit, special care must be taken to extract a statistically meaningful check. This issue can be solved using a *k-folding* approach, where five of the six total folds are used to parametrize the acceptance, while the last one is used as dataset in the fit. For simplicity, no weights are considered when parametrizing the acceptances and in the fit to the datasets. Figure 9.2 shows the difference between the result of the amplitude fits performed to $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ and $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ decays and the generator values of the parameters, as defined in `EvtGen`. A reasonable agreement is found between all channels and years for the trigger category L0L!, confirming the validity of the approach chosen. A similar agreement is obtained when considering the trigger category L0I. Notice that, since the floating parameters in the signal-only *pdf* are the real and complex part of the amplitudes $\tilde{A}_{\perp,\parallel}$ (no S-wave is present in the simulation), only the last four columns of Fig. 9.2 come directly from the fit, while the remaining are derived as described by Eq. 2.24.

9.2.3 Amplitude fits to data

Extended maximum likelihood fits to data samples of $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decay candidates are performed in $\cos\theta_K$, $\cos\theta_\ell$, ϕ, $m(K\pi\ell\ell)$ and $m(K\pi)$, simultaneously between Run 1 and Run 2, but separately for different trigger categories. The $m(K\pi)$ range chosen is reduced to $[825.9,\ 965.9]\,\mathrm{MeV}/c^2$ in

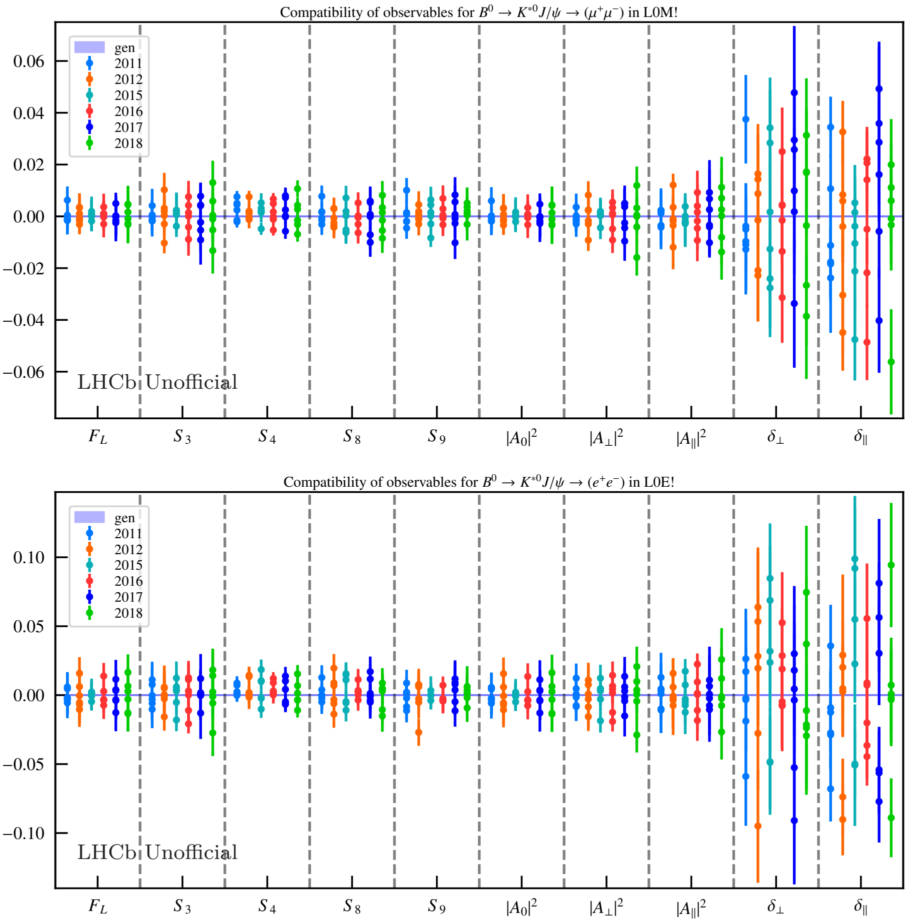


Figure 9.2: Compatibility of the amplitude parameters and angular observables between the generator values and the fit results of the reconstructed simulation of $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ (top) and $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ (bottom) in the trigger category L0L! in all different years. The purple line, labelled as 'gen' correspond to the generator values as obtained from `EvtGen`.

order to allow the comparison with the results in the literature. Since no strong dependence on the $m_{K\pi}$ mass is expected, the acceptance used is still obtained in the region of [795.9, 995.9] MeV/c^2. Two different fit configurations have been studied: the first one is an extension of the mass fits discussed in Sec. 8.3.2 to include the angles and the $m(K\pi)$ mass, the second one corresponds to an amplitude fit in which all sources of backgrounds are reduced as much as possible.

A detailed discussion of the first approach is described in Appendix. C. The main outcome of this study is that it is possible to extend the mass fits including the angular distributions and $m(K\pi)$ mass of the most important backgrounds, providing a reasonable description of the data. When comparing the obtained observables between the muon and electron channels, as it is done in Figure C.4, a good agreement between the P-wave observables is observed, while a significant discrepancy can be noticed on δ_S. Such result is not unexpected since the underlying parametrization of the different distributions of the peaking backgrounds, in particular the partially reconstructed ones, cannot be trusted completely from simulation.

In the second approach, candidates are selected exactly as discussed in Sec. 8.3.2 with two main differences: a tighter MVA cut, corresponding to the values used for the rare modes, and a smaller mass range, corresponding to [5170, 5700] MeV/c^2 and [4900,5800] MeV/c^2 for muons and electrons, respectively. Additionally, to reduce the partially reconstructed component from electron fits, a cut on the reconstructed invariant mass of the B^0 with the J/ψ constraint is applied, requiring $m(K\pi ee)^{J/\psi}_{DTF} > 5150\,\mathrm{MeV}/c^2$.

After these selections, the only background component expected to significantly survive is the combinatorial. Its shape is parametrized using a decreasing exponential function for the mass and Chebyshev polynomials up to third (second) order for $\cos\theta_K$, up to second order for $\cos\theta_\ell$, up to second (first) order for ϕ and up to first (second) order for $m_{K\pi}$ for muons (electrons), assuming factorization between the different variables considered. This approach is similar to the one used in Sec. 10.2 to parametrize the combinatorial component in the rare mode. In fits to $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ candidates, the combinatorial shape is allowed to float freely. In fits to $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ candidates, due to the reduced amount of combinatorial background caused by the additional cuts, the angular and $m(K\pi)$ parametrization of the combinatorial cannot be determined directly from the main fit. Instead, a sample of candidates reconstructed as the lepton flavour violation channel $B^0 \to K^{*0}\mu^\pm e^\mp$, with the selection as aligned as possible to the $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ channel, is used as a proxy for the combinatorial in data.[2]

In the muon fits, the floating parameters correspond to: the real and complex part of the amplitudes $\tilde{A}_{\perp,\parallel,S0}$, the shift and scale of the invariant mass peak, the slope of the exponential distribution, the parameters of the angular combinatorial distributions and the width and centre of Breit-Wigner in $m_{K\pi}$. In the electron fits, due to the low level of combinatorial available, the angular parametrization of the combinatorial is fixed to the fit results obtained in the LFV sample and the shape of the invariant mass is fixed to the values obtained from simulation. All the other parameters are floated as in the muon mode.

The projections of the fits for muons and electrons in the trigger categories L0L! for Run 2 are shown in Figs. 9.3 and 9.4. Similar results have been obtained for the L0I trigger categories and are not shown here for brevity. Overall the fit projections show a good agreement with the data, with four main exceptions:

---

[2] To avoid including any double-semileptonic component in the dataset, the events have been selected above the B^0 nominal mass, requiring $m(K\pi\ell\ell) > 5300\,\mathrm{MeV}/c^2$, and loosening the MVA requirement from 0.97 to 0.3 to increase the statistics available.

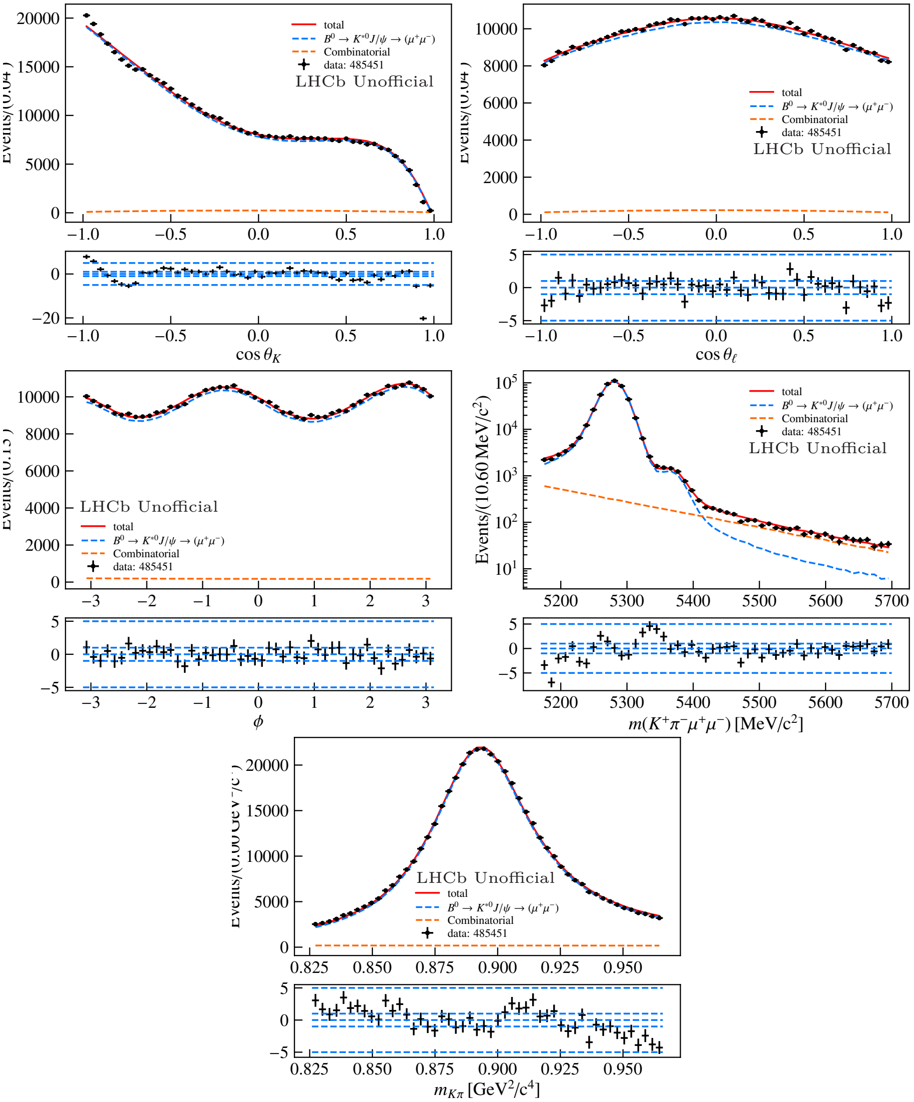


Figure 9.3: Projection of the fit result for the $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ angular fit in the L0L! trigger category for Run 2.

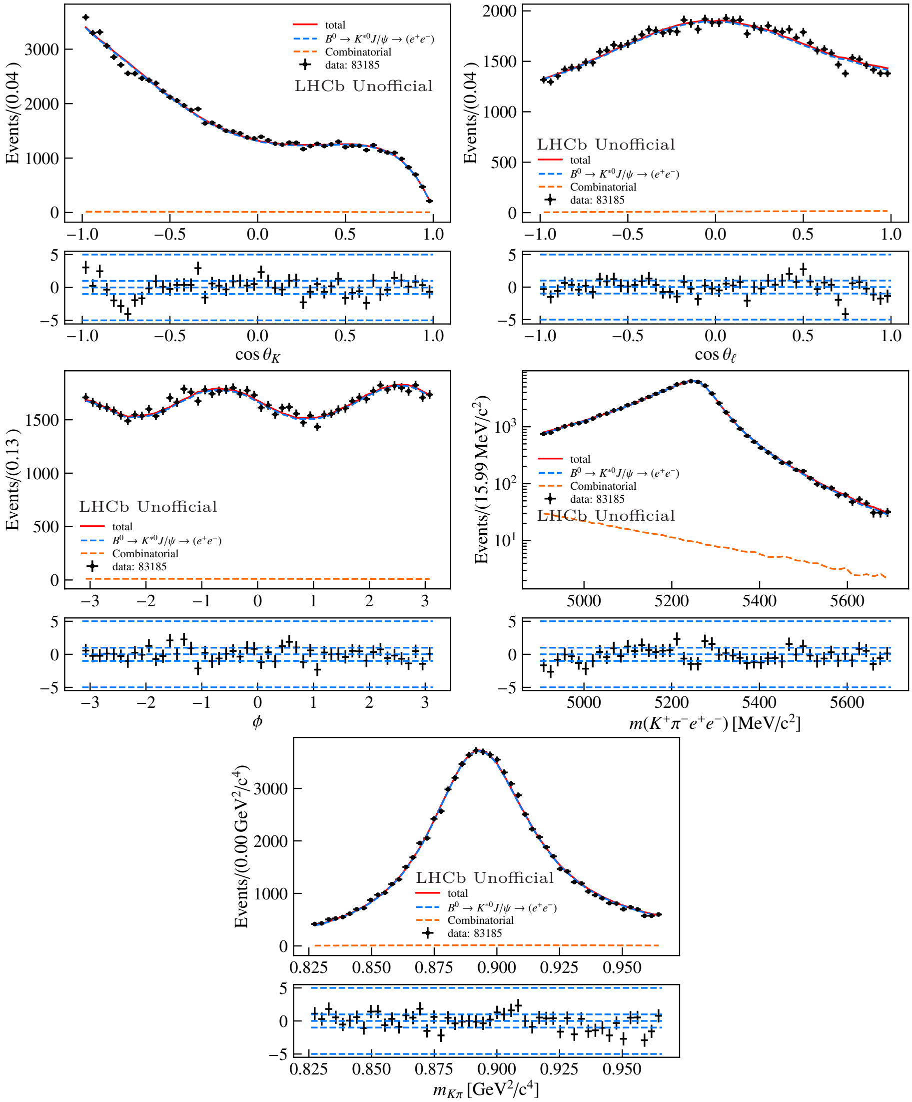


Figure 9.4: Projection of the fit result for the $B^0 \to K^{*0} J/\psi(\to e^+e^-)$ angular fit in the L0L! trigger category for Run 2.

- A large deviation in the region of $\cos\theta_K$ between -1 and 0.5 that is consistent between electrons and muons. This deviation, which is also seen in Ref. [29], is caused by the presence of exotic charmonium states, *e.g.* $Z_c^+(4200)$ and $Z_c^+(4430)$, which decay to the $J/\psi\pi$ pair. While this type of background significantly distorts the low region of the $\cos\theta_K$ distribution, a good description of this contribution goes beyond the scope of this analysis. This is not considered to be a serious problem for the fit validation, as existing muon results are similarly affected and no such component is expected to be present in the fits to the rare modes.
- A large deviation in the region of $\cos\theta_K$ between 0.85 and 1 only in the muons. This deviation is an artefact caused by the choice of truncating the expansion in Legendre polynomials for $\cos\theta_K$ at the sixth order. Preliminary studies suggest that such choice corresponds to a systematic uncertainty between one and two times the statistical uncertainties obtained.
- Small deviations in the invariant mass distributions for both lepton flavours can be seen. This effect is caused by some remaining mismodelling of the signal shapes, visible also in Fig. 8.5, and partially by the contribution of remaining backgrounds that have not been included in the fit (*e.g.* remaining partially reconstructed and peaking background). Since the size of these mismodelling is limited, no major impact is expected in the observables of interest.
- A consistent deviation in the $m_{K\pi}$ invariant mass in the muon channel. Most muons fits are affected by a non-flat trend in the pulls, the effect being more evident when more statistics is available, suggesting that something similar might be happening also in the electrons. A similar behaviour has been observed in Ref. [164] (Figure 9.1) and seems to be linked to the choice of parametrization used for the description of the Breit-Wigner. It has been shown that this deviation can be greatly reduced if a parametrization as the one used in Ref. [29] is used instead.

To meaningfully compare the results obtained from the fits to data with the values obtained in Refs. [29, 165], it is necessary to also estimate the systematic uncertainties of the measurement. The impact of this contribution to the total uncertainty budget is expected to be quite sizeable due to the high statistics available in data for these resonant modes. In the following, the focus is limited on one main source of uncertainty: the uncertainty on the acceptance parametrization due to the limited statistics available in the simulated samples. To estimate this contribution, a set of 100 signal-only toys is generated based on the best fit value obtained from data and corresponding to the channel and trigger category of interest. Each toy generated is then fitted in two different configurations: using the standard parametrization used for the generation or using an alternative acceptance parametrization obtained from bootstrapping (see

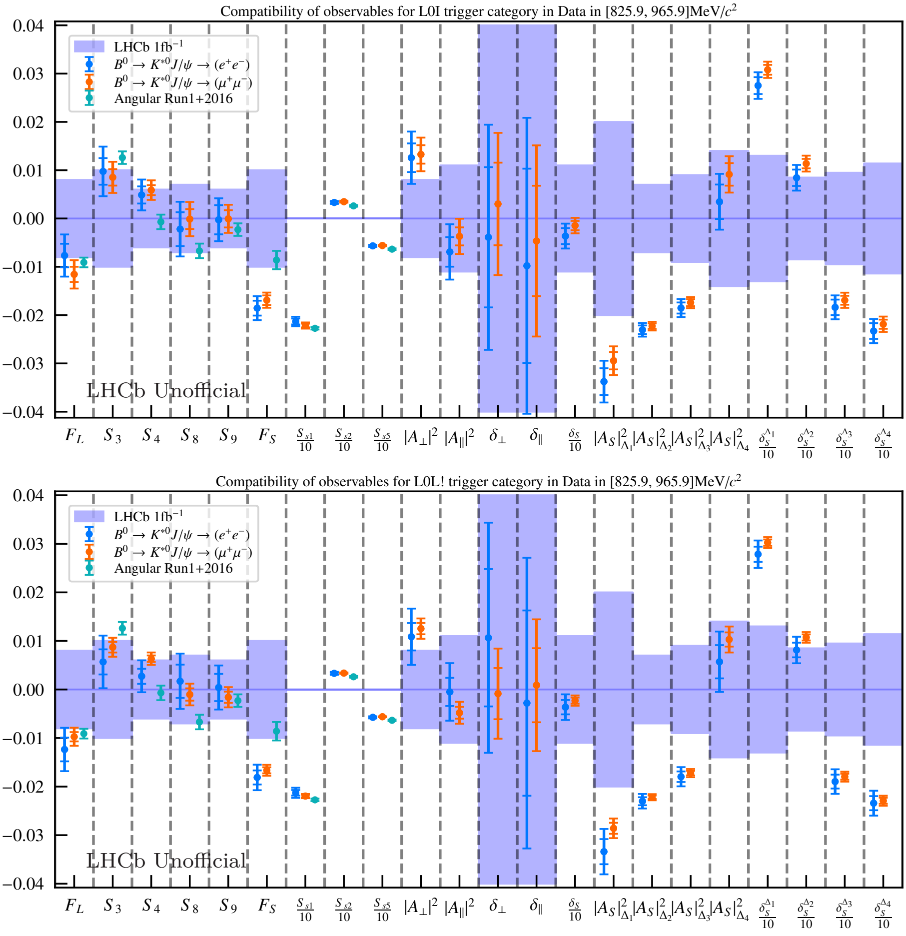


Figure 9.5: Summary of amplitude parameters and the derived angular observables for the simultaneous fit in RUN 1 and RUN 2 of electrons and muons. The fits are performed separately in the trigger categories L0I (top) and L0L! (bottom).

Type	$(\to e^+e^-)$	$\Delta(\sigma)^{e^+e^-}$ LHCb 1fb^{-1}	$\Delta(\sigma)^{e^+e^-}$ LHCb Run 1 +2016	$(\to \mu^+\mu^-)$	$\Delta(\sigma)^{\mu^+\mu^-}$ LHCb 1fb^{-1}	$\Delta(\sigma)^{\mu^+\mu^-}$ LHCb Run 1 +2016	LFU
F_L	0.5643 ± 0.0024 ± 0.0036	-0.84	0.32	0.5604 ± 0.0016 ± 0.0025	-1.36	-0.79	-0.74
F_S	0.0454 ± 0.0015 ± 0.0019	-1.80	-3.19	0.0471 ± 0.0010 ± 0.0012	-1.67	-3.37	0.56
S_3	-0.0032 ± 0.0028 ± 0.0043	0.87	-0.54	-0.0045 ± 0.0017 ± 0.0027	0.81	-1.17	-0.20
S_4	-0.2451 ± 0.0018 ± 0.0027	0.72	1.58	-0.2441 ± 0.0010 ± 0.0017	0.93	2.61	0.27
S_8	-0.0502 ± 0.0035 ± 0.0044	-0.24	0.77	-0.0481 ± 0.0021 ± 0.0029	-0.01	1.71	0.31
S_9	-0.0843 ± 0.0030 ± 0.0033	-0.03	0.44	-0.0841 ± 0.0018 ± 0.0023	-0.01	0.70	0.04
S_{s1}	-0.2132 ± 0.0061 ± 0.0077	–	1.39	-0.2216 ± 0.0039 ± 0.0051	–	0.82	-0.72
S_{s2}	0.0332 ± 0.0020 ± 0.0024	–	2.02	0.0348 ± 0.0012 ± 0.0017	–	3.24	0.42
S_{s5}	-0.0566 ± 0.0015 ± 0.0017	–	2.43	-0.0559 ± 0.0008 ± 0.0009	–	3.59	0.27
δ_S	2.1736 ± 0.0165 ± 0.0198	-0.32	–	2.1954 ± 0.0096 ± 0.0129	-0.13	–	0.72
$\delta_S^{\Delta_1}$	-2.9180 ± 0.0175 ± 0.0213	2.07	–	-2.8853 ± 0.0102 ± 0.0134	2.35	–	1.01
$\delta_S^{\Delta_2}$	2.7443 ± 0.0169 ± 0.0205	0.95	–	2.7736 ± 0.0098 ± 0.0133	1.32	–	0.94
$\delta_S^{\Delta_3}$	1.7561 ± 0.0160 ± 0.0189	-1.88	–	1.7707 ± 0.0092 ± 0.0127	-1.76	–	0.50
$\delta_S^{\Delta_4}$	1.2970 ± 0.0163 ± 0.0191	-2.00	–	1.3111 ± 0.0095 ± 0.0127	-1.90	–	0.47
$\delta_\perp$	2.9361 ± 0.0145 ± 0.0182	-0.08	–	2.9430 ± 0.0085 ± 0.0120	0.07	–	0.25
$\delta_\parallel$	-2.9498 ± 0.0201 ± 0.0231	-0.19	–	-2.9446 ± 0.0114 ± 0.0162	-0.10	–	0.14
$\lvert A_S\rvert^2_{\Delta_1}$	0.0812 ± 0.0028 ± 0.0033	-1.65	–	0.0856 ± 0.0018 ± 0.0024	-1.46	–	0.83
$\lvert A_S\rvert^2_{\Delta_2}$	0.0260 ± 0.0009 ± 0.0011	-3.23	–	0.0267 ± 0.0006 ± 0.0007	-3.16	–	0.45
$\lvert A_S\rvert^2_{\Delta_3}$	0.0335 ± 0.0011 ± 0.0015	-2.02	–	0.0346 ± 0.0007 ± 0.0009	-1.92	–	0.51
$\lvert A_S\rvert^2_{\Delta_4}$	0.1085 ± 0.0036 ± 0.0045	0.23	–	0.1141 ± 0.0023 ± 0.0030	0.63	–	0.82
$\lvert A_\perp\rvert^2$	0.2146 ± 0.0030 ± 0.0045	1.30	–	0.2153 ± 0.0019 ± 0.0029	1.52	–	0.11
$\lvert A_\parallel\rvert^2$	0.2211 ± 0.0031 ± 0.0048	-0.56	–	0.2243 ± 0.0019 ± 0.0031	-0.32	–	0.47

Type	$(\to e^+e^-)$	$\Delta(\sigma)^{e^+e^-}_{\text{LHCb 1fb}^{-1}}$	$\Delta(\sigma)^{e^+e^-}$ LHCb Run 1 +2016	$(\to \mu^+\mu^-)$	$\Delta(\sigma)^{\mu^+\mu^-}$ LHCb 1fb^{-1}	$\Delta(\sigma)^{\mu^+\mu^-}$ LHCb Run 1 +2016	LFU
F_L	0.5596 ± 0.0024 ± 0.0037	-1.35	-0.71	0.5623 ± 0.0010 ± 0.0016	-1.18	-0.30	0.54
F_S	0.0459 ± 0.0014 ± 0.0022	-1.75	-2.93	0.0474 ± 0.0006 ± 0.0009	-1.65	-3.65	0.52
S_3	-0.0073 ± 0.0026 ± 0.0048	0.50	-1.24	-0.0043 ± 0.0011 ± 0.0016	0.85	-1.66	0.52
S_4	-0.2473 ± 0.0016 ± 0.0029	0.40	0.95	-0.2437 ± 0.0007 ± 0.0011	1.03	3.56	1.03
S_8	-0.0463 ± 0.0034 ± 0.0046	0.19	1.42	-0.0490 ± 0.0013 ± 0.0018	-0.14	2.10	-0.44
S_9	-0.0836 ± 0.0028 ± 0.0036	0.05	0.57	-0.0856 ± 0.0012 ± 0.0017	-0.25	0.29	-0.40
S_{s1}	-0.2128 ± 0.0057 ± 0.0087	–	1.34	-0.2195 ± 0.0024 ± 0.0036	–	1.53	-0.59
S_{s2}	0.0333 ± 0.0019 ± 0.0024	–	2.06	0.0337 ± 0.0008 ± 0.0012	–	3.45	0.13
S_{s5}	-0.0574 ± 0.0013 ± 0.0018	–	2.19	-0.0561 ± 0.0005 ± 0.0007	–	3.85	0.53
δ_S	2.1735 ± 0.0147 ± 0.0220	-0.32	–	2.1871 ± 0.0063 ± 0.0085	-0.21	–	0.48
$\delta_S^{\Delta_1}$	-2.9147 ± 0.0156 ± 0.0236	2.09	–	-2.8907 ± 0.0065 ± 0.0092	2.32	–	0.79
$\delta_S^{\Delta_2}$	2.7413 ± 0.0151 ± 0.0228	0.91	–	2.7672 ± 0.0063 ± 0.0089	1.25	–	0.88
$\delta_S^{\Delta_3}$	1.7503 ± 0.0143 ± 0.0209	-1.93	–	1.7605 ± 0.0061 ± 0.0082	-1.88	–	0.37
$\delta_S^{\Delta_4}$	1.2958 ± 0.0145 ± 0.0211	-2.00	–	1.3009 ± 0.0063 ± 0.0083	-2.00	–	0.18
$\delta_\perp$	2.9507 ± 0.0141 ± 0.0190	0.23	–	2.9391 ± 0.0053 ± 0.0076	-0.02	–	-0.45
$\delta_\parallel$	-2.9428 ± 0.0191 ± 0.0231	-0.06	–	-2.9391 ± 0.0076 ± 0.0112	0.02	–	0.11
$\lvert A_S\rvert^2_{\Delta_1}$	0.0816 ± 0.0026 ± 0.0039	-1.63	–	0.0864 ± 0.0011 ± 0.0017	-1.42	–	0.94
$\lvert A_S\rvert^2_{\Delta_2}$	0.0260 ± 0.0008 ± 0.0012	-3.22	–	0.0268 ± 0.0004 ± 0.0005	-3.16	–	0.51
$\lvert A_S\rvert^2_{\Delta_3}$	0.0340 ± 0.0011 ± 0.0016	-1.95	–	0.0348 ± 0.0005 ± 0.0007	-1.91	–	0.35
$\lvert A_S\rvert^2_{\Delta_4}$	0.1107 ± 0.0034 ± 0.0052	0.37	–	0.1153 ± 0.0015 ± 0.0022	0.72	–	0.68
$\lvert A_\perp\rvert^2$	0.2129 ± 0.0028 ± 0.0051	1.10	–	0.2145 ± 0.0012 ± 0.0017	1.52	–	0.27
$\lvert A_\parallel\rvert^2$	0.2275 ± 0.0029 ± 0.0052	-0.04	–	0.2232 ± 0.0012 ± 0.0019	-0.43	–	-0.68

Table 9.6: Overview of the amplitude parameters and the derived angular observables for the simultaneous fits in L0I (top) and L0L! (bottom) Run 1 and Run 2 for $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ and $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$, indicated here as $(\to e^+e^-)$ and $(\to \mu^+\mu^-)$ for brevity. Additionally the compatibility, expressed as a number of standard deviations, of the obtained values with respect to the reference reported in the first column of Table 9.5. The last columns shows instead the compatibility between electron and muon channel.

Sec. 7.2.3). The width of the toy-by-toy difference obtained between the use of the two alternative acceptances is used to estimate the corresponding systematic uncertainty. Overall, the size of this systematic uncertainty is between one and one and half times the statistical uncertainty.

A comparison between the observables extracted from the fits to data and the reference values listed in Table 9.5 is shown in Fig. 9.5, separately for the trigger categories L0I and L0L!. The inner error bar represents the statistical uncertainty, while the outer one corresponds to the sum in quadrature of statistical and systematic uncertainty. The values of the observables for electron and muon channels can also be found explicitly in Table 9.6, together with their compatibility with previous measurements and between themselves. To appropriately compare with the available results, the expected values of $|A_S|^2$ and δ_S are computed in four bins of $m_{K\pi}$: [826, 861] MeV/c^2, [861, 896] MeV/c^2, [896, 931] MeV/c^2 and [931, 966] MeV/c^2, referred for brevity as Δ_i with i from 1 to 4. The procedure used follows the definition in Eq. 9.11.

In general, the compatibility between P- and S-wave observables for electrons and muons in both trigger categories suggests that the current approach does not introduce any difference between the decays that involve the two lepton generations. This is the most important outcome of the test presented here. When comparing with previous measurements, the P-wave observables show a good agreement with respect to the results obtained with 1 fb^{-1} of data taken in 2011, indicated in the first columns in Table 9.5. Some tensions can be noticed for S_4 and S_8 when comparing to the check performed in the angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$ up to the year 2016. A stronger disagreement can be seen when comparing S-wave and S-P interference terms. These deviations are expected to be driven mainly by the difference between the S-wave fraction observed here and in Ref. [29]. A systematic contribution to the choice of the model used for the parametrization of the $m_{K\pi}$ distribution in the S-wave is not included in here, but could greatly reduce the incompatibility with respect to Ref. [29].

Chapter 10

Amplitude fits to $B^0 \to K^{*0}\ell^+\ell^-$ decays

The final goal of this analysis is to perform a measurement of the difference of the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ between muons and electrons. Chapter 5 has described the selections applied to data to isolate $B^0 \to K^{*0}\ell^+\ell^-$ decay candidates in data, while Ch. 7 discussed their impact in the expected number of signal events and their distributions. Additionally, the framework used in the fit has been tested in $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ data candidates in Ch. 9, providing satisfactory results. The following chapter concludes this work discussing the main ingredients of the amplitude fit, together with its main results.

Firstly, Sec. 10.1 extends the parametrization of the signal of Sec. 2.2 to include an S-wave contribution. Then, Sec. 10.2 presents the parametrizations used for the different backgrounds entering the final amplitude fit. As already mentioned in Ch. 4, the observed yield for the signal component is not free to float in the final fit, but rather constrained to the expected value predicted by its differential branching ratio. A detailed description of this constraint is presented in Sec. 10.3. Section 10.4 links the first three sections of this chapter together by describing the full likelihood and the parameters allowed to vary in its maximization. The projections obtained from the best fit results are shown for all the samples considered, while the numerical values of the parameters of interest are kept blind. The chapter ends with the discussion of the statistical properties of the fit (Sec. 10.5), together with an heuristic estimation of the main sources of systematic uncertainties (Sec. 10.6).

10.1 Parametrization of the signal amplitude

The signal component is parametrized starting from the expression defined in Eq. 2.14, where the q^2 dependence is stored implicitly in the I_i angular coefficients as a function of the transversity amplitudes $\mathcal{A}^{L,R}_{0,\perp,\parallel}(q^2)$. The parametrization of the local and non-local hadronic matrix elements follows the parametrization of Appendices A and B, respectively. Additionally, the signal parametrization is supplemented with the inclusion of the invariant mass lineshape $S^\ell(m(K\pi\ell\ell))$, in order to improve the discrimination between signal and backgrounds, and of the effective acceptance $\epsilon_{eff}(\vec{\Omega}, q^2)$, to model the distortion that the selections applied cause on the reconstructed angles and q^2. Considering these contributions, the signal *pdf* for $B^0 \to K^{*0}\ell^+\ell^-$ decays can be written as

$$\mathcal{P}_{sig}(x;\theta) = \mathcal{N}_{norm} \cdot S^\ell(m(K\pi\ell\ell)) \cdot \epsilon_{eff}(\vec{\Omega}, q^2) \cdot \frac{\mathrm{d}^4\Gamma(B^0 \to K^{*0}\ell^+\ell^-)}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}}\,, \quad (10.1)$$

where $\mathcal{N}_{norm}$ corresponds to the analytical normalization of the function. However, as already introduced in Sec. 5.2.5, reconstructed $K^+\pi^-$ candidates can originate not only from the decay of the $K^*(892)^0 \to K^+\pi^-$ resonance, but also from the decay of scalar resonances or from non-resonant decays of the type $B^0 \to K\pi\ell^+\ell^-$.[1] Previous measurements have estimated the contribution of the S-wave component in $m_{K\pi} \in [796, 996]\,\mathrm{MeV}/c^2$ to be around 10% [33]. Even if its size is small, its peculiar distribution in the decay angles and q^2 could have a sizable impact on some variables of interest if not properly treated. For this reason, this contribution is included in the signal parametrization by modifying the differential decay rate of Eq. 2.14 in two ways: with the inclusion of four additional complex amplitudes, $\mathcal{A}^{L,R}_{S0,St}(q^2)$, and the introduction of a parametrization of the $m(K\pi)$ lineshapes for P and S-waves, $A_P(m_{K\pi}, q^2)$ and $A_S(m_{K\pi}, q^2)$, respectively. With the inclusion of the S-wave contribution, the signal *pdf* then becomes

$$\mathcal{P}_{sig}(x;\theta) = \mathcal{N}_{norm} \cdot S^{\ell}(m(K\pi\ell\ell)) \cdot \epsilon_{eff}(\vec{\Omega}, q^2) \cdot \frac{\mathrm{d}^5\Gamma(B^0 \to K^+\pi^-\ell^+\ell^-)}{\mathrm{d}\vec{\Omega}\,\mathrm{d}q^2\,\mathrm{d}m(K\pi)}, \quad (10.2)$$

where the term $\frac{\mathrm{d}^5\Gamma(B^0\to K^+\pi^-\ell^+\ell^-)}{\mathrm{d}\vec{\Omega}\,\mathrm{d}q^2\,\mathrm{d}m(K\pi)}$ refers to the differential decay rate including both P and S-wave contributions, for which the final expression can be found in Eq. 10.16.

10.1.1 Inclusion of the S-wave transversity amplitudes

To include the contribution of decays in which the $K\pi$ system has been produced in an S-wave configuration, it is possible to start considering the resonant decay of a scalar meson, such as $K^*(1430)^0$. The extension of this model to a more generic non-resonant contribution is discussed in Sec. 10.1.2. As described in Ref. [166], four additional transversity amplitudes must be included in the model, $\mathcal{A}^{L,R}_{St}$ and $\mathcal{A}^{L,R}_{St}$, whose definitions read as

$$\begin{aligned}
\mathcal{A}^{L,R}_{S0} &= -N_0 \frac{\sqrt{\lambda_{K^*(1430)^0}}}{M_B\sqrt{q^2}} \Big\{ \Big[(\mathcal{C}_9 - \mathcal{C}_9') \mp (\mathcal{C}_{10} - \mathcal{C}_{10}') \Big] f_+(q^2) \\
&\qquad + \frac{2m_b M_B}{q^2} \Big[(\mathcal{C}_7 - \mathcal{C}_7') f_T(q^2) - 16\pi^2 \frac{M_B}{m_b} \mathcal{H}_{S0}(q^2) \Big] \Big\}, \\
\mathcal{A}_{St} &= -2N_0 \frac{M_B^2 - M^2_{K^*(1430)^0}}{M_B\sqrt{q^2}} (\mathcal{C}_{10} - \mathcal{C}_{10}') f_0(q^2), \qquad (10.3)
\end{aligned}$$

where the normalization factor N_0 is given by

$$N_0 = G_F \alpha_e V_{tb} V^*_{ts} \sqrt{\frac{q^2 \beta_l \sqrt{\lambda_{K^*(1430)^0}}}{3 \cdot 2^{10} \pi^5 M_B}}, \quad (10.4)$$

---

[1] Contribution for higher spin resonances are not considered due to the vanishingly small contribution expected in the $m_{K\pi}$ region of interest.

and $\lambda_{K^*(1430)^0}$ is the kinematic Källén function, defined in Eq. 2.18, after the substitution $M_{K^{*0}} \to M_{K^*(1430)^0}$. Three new form factors are required to parametrize the scalar $B \to [K\pi]_{J=0}$ transition matrix elements, namely f_+, f_T and f_0, whose definitions slightly differ from Ref. [166] and are detailed in Appendix D. The non-local hadronic functions $\mathcal{H}_{S0}(q^2)$, introduced here for coherence with the P-wave parametrization, follow Eq. B.2, while Eq. B.3 is replaced by

$$\hat{\mathcal{H}}_{S0}(z) = \Big[\sum_{k=0}^{K} \alpha_k^{S0} z^k\Big] f_+(q^2)\,. \tag{10.5}$$

Note that the functions $\mathcal{H}_{S0}(z)$, describing the non-local hadronic contribution in the S-wave part, are fixed to zero in the nominal fit configuration (see Sec. 10.4) and will only be considered in the determination of systematic uncertainties on the signal modelling.

10.1.2 Inclusion of the $K\pi$ invariant mass lineshape

To facilitate the discrimination between P- and S-wave in the fit, the differential decay rate can be expanded introducing a parametrization for the hadronic invariant mass $m(K\pi)$. The additional discriminating power comes from the fact that the two contributions have very different distributions, with the P-wave being characterized by a resonant behaviour, with a peak inside the $m(K\pi)$ range of interest, while the S-wave is expected to be almost linear and slowly rising. It is important to point out that the approach described should be regarded as an effective description of the $m(K\pi)$ spectrum, which is motivated by its ability to satisfactorily describe the data in previous analysis [33]. Attempts to improve on this parametrization are discussed in more detail in Appendix E and are not considered in the rest of this work.

A generic decay $B^0 \to K_J^{*0}\ell^+\ell^-$, where K_J^{*0} represents a resonance with angular momentum J that subsequently decays into a kaon and a pion, can be described as a function of $m_{K\pi}$ and q^2 with the following expression

$$\begin{aligned}\tilde{f}_J(m_{K\pi}, q^2) =& \sqrt{p}\cdot B'_{L_B}(p, p_0, d)\cdot\left(\frac{p}{p_0}\right)^{L_B} \\ &\times \sqrt{k}\cdot B'_{L_{K\pi}}(k, k_0, d)\cdot\left(\frac{k}{k_0}\right)^{L_{K\pi}}\cdot f_J(m_{K\pi})\,.\end{aligned} \tag{10.6}$$

In Eq. 10.6, $k\,(p)$ corresponds to the momentum of the $K^+(K_J^{*0})$ in the rest frame of the $K_J^{*0}\,(B)$, the parameters $k_0(p_0)$ are the values of $k(p)$ evaluated at the pole-mass of the mother-particle, B'_L corresponds to the Blatt-Weisskopf barrier or form factors [129], and $L_{K\pi}(L_B)$ correspond the orbital angular momentum between K^+ and $\pi^-(K_J^{*0}$ and the dilepton). The dependence of $k\,(p)$ on $m_{K\pi}$ (q^2) is kept implicit in Eq. 10.6 for simplicity. For the decays of K_J^{*0} into spinless daughters such as kaons and pions, the conservation of total momentum implies that $L_{K\pi} = J$, corresponding to one or zero if the resonance is vector or scalar, respectively. Similarly, under the assumption that in the decays of $B^0 \to K_J^{*0}\ell^+\ell^-$ the dilepton system is dominantly in a vector state, the conservation of angular momentum requires that for scalar resonances the angular momentum between K_J^{*0} and $\ell^+\ell^-$ is one ($L_B = 1$), while it is equal to zero for vector resonances such as $K^*(892)^0$ ($L_B = 0$).

When considering the decay of the $K^*(892)^0$ vector resonance, characterized by a relatively narrow width and located far from the threshold of additional decay channels, the decay amplitude is expected to be effectively described by a relativistic Breit-Wigner function, defined as

$$f_P \equiv f_{BW}^{892}(m_{K\pi}) = \frac{1}{m_{K\pi}^2 - m_{892}^2 - i m_{892}\Gamma_{892}(m_{K\pi})}\,, \tag{10.7}$$

where m_{892}^2 is the pole-mass of the $K^*(892)^0$ resonance and the running width $\Gamma_{892}(m_{K\pi})$ is given by

$$\Gamma_{892}(m_{K\pi}) = \Gamma_{892} B_{L=1}^{\prime 2}(k, k_{892}, d) \left(\frac{k}{k_{892}}\right)^3 \left(\frac{m_{892}}{m_{K\pi}}\right). \tag{10.8}$$

For decays involving the $K\pi$ system in an S-wave state, caused either by the decay of a scalar resonance such as $K^*(1430)^0$ and $K_0^*(800)$ or from a non-resonant $B^0 \to K\pi\ell^+\ell^-$ decay, a sum of Breit-Wigners cannot be used and the choice of a different parametrization should be preferred. In the following, we rely on a model for the S-wave $K\pi$ spectrum that was developed to describe scattering experiments: the so-called "LASS" model [167]. Its functional form is defined as

$$f_S \equiv f_{LASS}(m_{K\pi}) = \left(\frac{1}{\cot\delta_B - i} + e^{2i\delta_B}\frac{1}{\cot\delta_R - i}\right)\,, \tag{10.9}$$

where the two terms $\cot\delta_B$ and $\cot\delta_R$ are defined as

$$\cot\delta_B = \frac{1}{ak} + \frac{rk}{2} \quad \text{and} \quad \cot\delta_R = \frac{m_{1430}^2 - m_{K\pi}^2}{m_{1430}\Gamma_{1430}(m_{K\pi})}\,, \tag{10.10}$$

with m_{1430}^2 corresponding to the pole-mass of the $K^*(1430)^0$ resonance and a and r empirical parameters fixed to the values $a = 1.95\,[\,\text{GeV}/c]^{-1}$ and $r = 1.78\,[\,\text{GeV}/c]^{-1}$ of Ref. [29, 167, 168]. The second term of Eq. 10.9 corresponds to a relativistic Breit-Wigner for the $K_0^*(1430)^0$ resonance, with

$$\Gamma_{1430}(m_{K\pi}) = \Gamma_{1430}\frac{k}{k_{1430}}\frac{m_{1430}}{m_{K\pi}}. \tag{10.11}$$

An important feature of Eq. 10.6 is its dependence on the momentum of the decay products, that results in a non-trivial dependence between $m_{K\pi}$ and q^2. This dependence plays a relevant role when extending the signal parametrization to the $K\pi$ system. For illustration, Figure 10.1 shows a comparison between the modulus square of the functions f_{BW}^{892} and f_{LASS}, together with their interference term for four different relative phases δ_S.

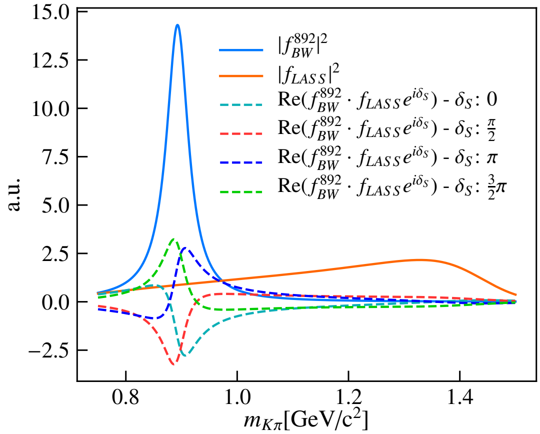


Figure 10.1: Comparison between the $m_{K\pi}$ distributions of the square modulus of f_{BW}^{892} and f_{LASS}, together with the corresponding interference term for the relative phases $\delta_S = \{0, \frac{\pi}{2}, \pi, \frac{3}{2}\pi\}$.

10.1.3 Full differential decay rate

With the introduction of the description of the $K\pi$ invariant mass system, it is possible to define the total differential decay rate. The transversity amplitudes defined in Eq. 2.16 and 10.3, can be modified to accommodate the additional $m(K\pi)$ dependency as follows

$$\begin{aligned} \mathcal{A}^{L,R}_{0,\perp,\parallel,t}(q^2, m_{K\pi}) &= \mathcal{A}^{L,R}_{0,\perp,\parallel,t}(q^2) \cdot A_P(m_{K\pi}, q^2)\,, \\ \mathcal{A}^{L,R}_{S0,St}(q^2, m_{K\pi}) &= \mathcal{A}^{L,R}_{S0,St}(q^2) \cdot A_S(m_{K\pi}, q^2)\,, \end{aligned} \tag{10.12}$$

where

$$\begin{aligned} A_P(m_{K\pi}, q^2) &= \mathcal{N}_P(q^2) \cdot f_P(m_{K\pi}, q^2)\,, \\ A_S(m_{K\pi}, q^2) &= \mathcal{N}_S(q^2) \cdot g_S e^{i\delta_S} \cdot f_S(m_{K\pi}, q^2)\,, \end{aligned} \tag{10.13}$$

while the coefficients g_S and δ_S represent the relative magnitude and phase between the S- and P-wave contributions. The constants $\mathcal{N}_P$ and $\mathcal{N}_S$ are defined requiring that the modulus squared of A_P and A_S are normalized to unity in the entire physical region of $m_{K\pi}$ as a function of q^2

$$\int_0^\infty \left|\mathcal{N}_i(q^2) \cdot f_i(m_{K\pi}, q^2)\right|^2 \mathrm{d}m_{K\pi} = 1 \quad \forall\ q^2\,, \tag{10.14}$$

The above normalization arises from the condition

$$\int_0^\infty \frac{\mathrm{d}^5\Gamma}{\mathrm{d}q^2\,\mathrm{d}\Omega\,\mathrm{d}m_{K\pi}} dm_{K\pi} = \frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}\Omega}\,, \tag{10.15}$$

which is imposed to satisfy the consistency of the P-wave differential decay rate when integrating over the entire physical $m_{K\pi}$ region.

These additional scalar amplitudes can be incorporated in the differential decay rate of Eq. 2.14 in the following way:

$$\begin{aligned}\frac{32\pi}{9}\frac{\mathrm{d}^5\Gamma}{\mathrm{d}q^2\,\mathrm{d}m_{K\pi}\,\mathrm{d}\vec{\Omega}}\bigg|_{\mathrm{P+S-wave}} &= \frac{32\pi}{9}\frac{\mathrm{d}^5\Gamma}{\mathrm{d}q^2\,\mathrm{d}m_{K\pi}\,\mathrm{d}\vec{\Omega}}\bigg|_{\mathrm{P-wave}}\\ &+\left(I_{1c}^S + I_{2c}^S\cos 2\theta_l\right)\\ &+\left(\tilde{I}_{1c} + \tilde{I}_{2c}\cos 2\theta_l\right)\cos\theta_K\\ &+\left(\tilde{I}_4\sin 2\theta_l + \tilde{I}_5\sin\theta_l\right)\sin\theta_K\cos\phi\\ &+\left(\tilde{I}_7\sin\theta_l + \tilde{I}_8\sin 2\theta_l\right)\sin\theta_K\sin\phi\,, \end{aligned} \tag{10.16}$$

where the additional angular coefficients [166] I_i^S are pure S-wave contributions and $\tilde{I}_i$ are S-P wave interference terms. Their definition is:

$$\begin{aligned} I_{1c}^S &= \frac{1}{3}\left\{\left[|\mathcal{A}_{S0}^L|^2 + |\mathcal{A}_{S0}^R|^2\right] + \frac{4m_l^2}{q^2}\left[|\mathcal{A}_{St}|^2 + 2\,\mathcal{R}e(\mathcal{A}_{S0}^L\mathcal{A}_{S0}^{R\;*})\right]\right\},\\ I_{2c}^S &= -\frac{1}{3}\beta_l^2\left[|\mathcal{A}_{S0}^L|^2 + |\mathcal{A}_{S0}^R|^2\right],\\ \tilde{I}_{1c} &= \frac{2}{\sqrt{3}}\,\mathcal{R}e\Big[\mathcal{A}_{S0}^L\mathcal{A}_0^{L\,*} + \mathcal{A}_{S0}^R\mathcal{A}_0^{R*} + \frac{4m_l^2}{q^2}\Big(\mathcal{A}_{S0}^L\mathcal{A}_0^{R*} + \mathcal{A}_0^L\mathcal{A}_{S0}^{R\;*} + \mathcal{A}_{St}\mathcal{A}_t^{\,*}\Big)\Big],\\ \tilde{I}_{2c} &= -\frac{2}{\sqrt{3}}\beta_l^2\,\mathcal{R}e\Big[\mathcal{A}_{S0}^L\mathcal{A}_0^{L\,*} + \mathcal{A}_{S0}^R\mathcal{A}_0^{R*}\Big],\\ \tilde{I}_4 &= -1\cdot\sqrt{\frac{2}{3}}\beta_l^2\,\mathcal{R}e\Big[\mathcal{A}_{S0}^L\mathcal{A}_\parallel^{L\,*} + (L\to R)\Big],\\ \tilde{I}_5 &= \sqrt{\frac{8}{3}}\beta_l^2\,\mathcal{R}e\Big[\mathcal{A}_{S0}^L\mathcal{A}_\perp^{L\,*} - (L\to R)\Big],\\ \tilde{I}_7 &= -1\cdot\sqrt{\frac{8}{3}}\beta_l^2\,\mathcal{I}m\Big[\mathcal{A}_{S0}^L\mathcal{A}_\parallel^{L\,*} - (L\to R)\Big],\\ \tilde{I}_8 &= \sqrt{\frac{2}{3}}\beta_l^2\,\mathcal{I}m\Big[\mathcal{A}_{S0}^L\mathcal{A}_\perp^{L\,*} + (L\to R)\Big], \end{aligned} \tag{10.17}$$

where the dependence of the transversity amplitudes on q^2 and $m_{K\pi}$ is implied for simplicity. In analogy with Eq. 2.15, the -1 in front of $\tilde{I}_{4,7}$ results from the transformation from the theory to the LHCb angular convention [31]. An example of the effect of the inclusion of the S-wave contribution in the full differential decay rate for the central-q^2 region is reported in Figure 10.2.

It is interesting to notice that, when integrating over the angles, the interference terms between P- and S-wave and many of the angular parameters do not contribute to the total differential decay rate, which becomes

$$\frac{\mathrm{d}^2\Gamma}{\mathrm{d}q^2\,\mathrm{d}m_{K\pi}^2} = \left|A_P^2(m_{K\pi})\right|^2\cdot\frac{\mathrm{d}\Gamma^{\mathrm{P}}}{\mathrm{d}q^2} + \left|A_S(m_{K\pi}^2)\right|^2\cdot\frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2}\,, \tag{10.18}$$

with

$$\frac{\mathrm{d}\Gamma^{\mathrm{P}}}{\mathrm{d}q^2} = \frac{3}{4}(2I_{1s} + I_{1c}) - \frac{1}{4}(2I_{2s} + I_{2c}) \tag{10.19}$$

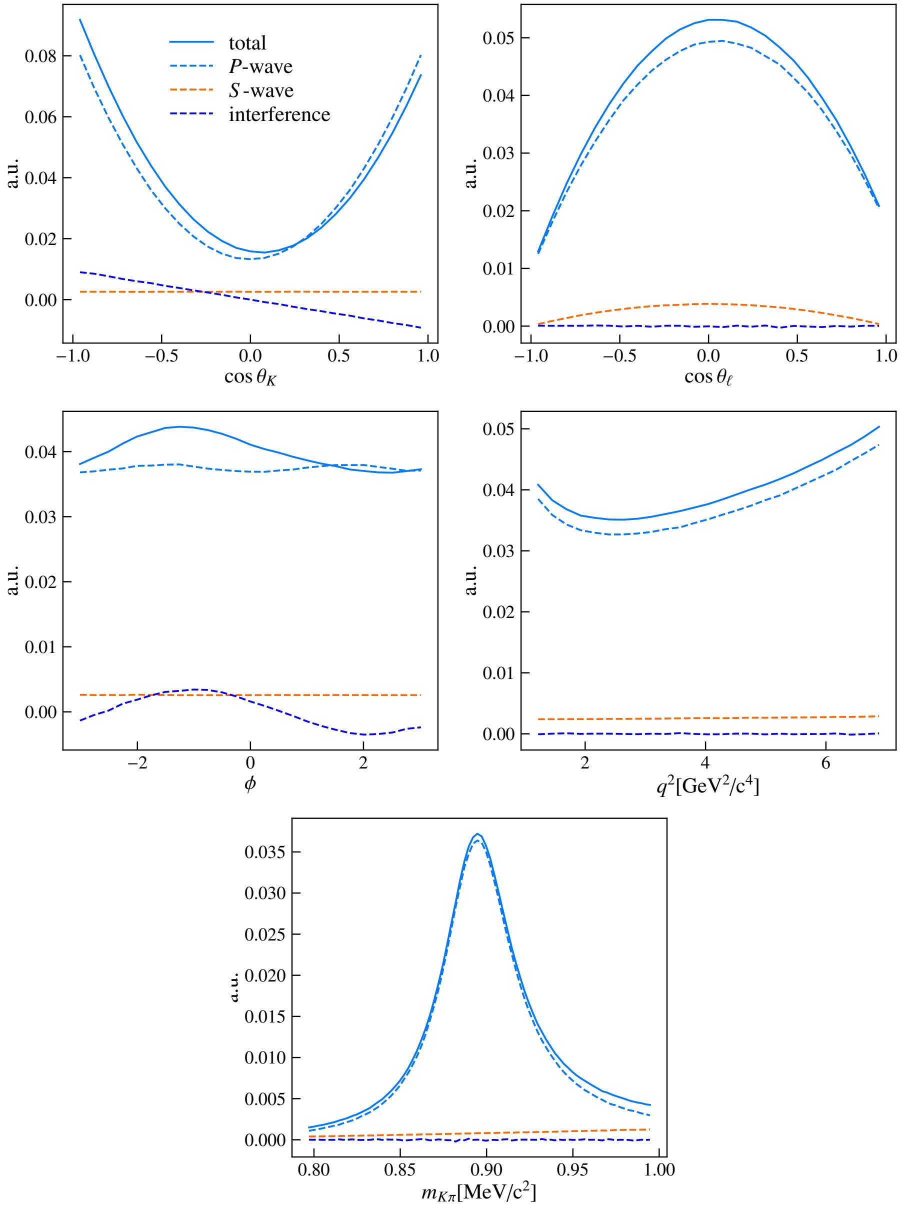


Figure 10.2: Projections of the distribution of $\cos\theta_K$, $\cos\theta_\ell$, ϕ, q^2 and $m_{K\pi}$ for the $B^0 \to K^+\pi^-\mu^+\mu^-$ decays that include P and S-wave contributions that interfere with a relative phase $\delta_S = \pi$. The fraction of S-wave considered, F_S, corresponds to $\approx 7\%$.

and

$$\frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2} = \frac{1}{4} I_{1c}^S - \frac{3}{4} I_{2c}^S \,. \tag{10.20}$$

10.1.3.1 Conventional observables

The conventional observables used to study the decays of $B^0 \to K^{*0}\ell^+\ell^-$ have been presented in Sec. 2.3 and correspond mainly to branching ratios and angular observables binned as a function of q^2. The methodology described in this work allows to obtain a continuos parametrization in these variables allowing to obtain, for the branching ratio, its differential distribution as

$$\mathrm{d}\mathcal{B}(B^0 \to K^{*0}\ell^+\ell^-)/\mathrm{d}q^2 = \frac{\tau_B}{\hbar} \int \mathrm{d}\vec{\Omega}\mathrm{d}m_{K\pi} \frac{\mathrm{d}^5\Gamma[B^0 \to K^{*0}\ell^+\ell^-]}{\mathrm{d}\vec{\Omega}\mathrm{d}q^2\mathrm{d}m_{K\pi}} = \frac{\tau_B}{\hbar} \cdot \frac{\mathrm{d}\Gamma^P}{\mathrm{d}q^2} \tag{10.21}$$

and, for the angular observables, their variation as a function of q^2

$$F_L(q^2) = \frac{\frac{3}{4}I_{1c} - \frac{1}{4}I_{2c}}{\frac{\mathrm{d}\Gamma^P}{\mathrm{d}q^2}} \quad \text{and} \quad S_i(q^2) = \frac{I_i}{\frac{\mathrm{d}\Gamma^P}{\mathrm{d}q^2}} \,. \tag{10.22}$$

Binned values of these observables can then be simply obtained by integrating Eq. 10.21 and Eq. 10.22 in the chosen bin of q^2, Δq^2. In the latter, a separate integration between numerator and denominator must be performed to compare with the available measurements:

$$< S_i > \Big|_{\Delta q^2} = \frac{\int I_i(q^2)\mathrm{d}q^2}{\int \frac{\mathrm{d}\Gamma}{\mathrm{d}q^2}\mathrm{d}q^2} \,. \tag{10.23}$$

Analogously, the S-wave fraction, F_S, can be extracted as function of q^2 and $m_{K\pi}$ with the following expression

$$F_S(m_{K\pi}, q^2) = \frac{\left|A_S(m_{K\pi}, q^2)\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2}(q^2)}{\left|A_P(m_{K\pi}, q^2)\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{P}}}{\mathrm{d}q^2}(q^2) + \left|A_S(m_{K\pi}, q^2)\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2}(q^2)} \,. \tag{10.24}$$

In order to obtain the measured value of the S-wave fraction in a given region of interest, *e.g.* a region corresponding to the bin Δq^2 and $\Delta m_{K\pi}$ in q^2 and $m_{K\pi}$, one must integrate the previous, obtaining

$$< F_S > \Big|_{\substack{\Delta q^2 \\ \Delta m_{K\pi}}} = \frac{\int \left|A_S(m_{K\pi}, q^2)\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2}(q^2)\, \mathrm{d}m_{K\pi}\mathrm{d}q^2}{\int \left(\left|A_P(m_{K\pi},q^2)\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{P}}}{\mathrm{d}q^2}(q^2) + \int \left|A_S(m_{K\pi},q^2)\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2}(q^2)\right)\mathrm{d}q^2\mathrm{d}m_{K\pi}} \,, \tag{10.25}$$

where the double integration runs over q^2 and $m_{K\pi}$. Note that, also in this case, $< F_S >$ is not a direct parameter of the fit, but rather a function of the parameters of the model.

In addition to the conventional flavour specific observables, it is also possible to derive LFU universality tests such as $\Delta S_i = S_i^{\mu} - S_i^{e}$, $Q_i = P_i^{\prime,\mu} - P_i^{\prime,e}$ [56] and $R_{K^{*0}} = \frac{\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)}{\mathcal{B}(B^0 \to K^{*0}e^+e^-)}$, both as a function of q^2 or integrated in a specific bin. The access to these binned observables is an interesting feature of this analysis, since it provides a direct metric to validate the results of the amplitude fit against the published results.

10.2 Parametrization of the backgrounds

The three most important backgrounds that enter the data fits of $B^0 \to K^{*0}\ell^+\ell^-$ decays are combinatorial, double-semileptonic and partially reconstructed decays. Due to the better momentum resolution, only the combinatorial background affects $B^0 \to K^{*0}\mu^+\mu^-$ decays in the range chosen for the mass fit.

10.2.1 Combinatorial background parametrization

The combinatorial background consists of decay candidates built from random particles coming from the fragments of the same *p*-*p* collision. Due to the sheer number of particles produced, it is possible that some of these combinatorial candidates pass the full chain of selections applied. Since they are not coming from a real B^0 meson decay, they do not peak in the $m(K\pi\ell\ell)$ region of the signal and are expected to be distributed according to a decreasing exponential. Similarly, their distributions in the angles, q^2 and $m(K\pi)$ are supposed to be flat or with a mild curvature. For these reasons, the parametrization of this background is performed using Chebyshev polynomials [169] up to order 3 for $\cos\theta_K$, up to order 2 for $\cos\theta_\ell$, ϕ and in q^2, up to order 1 for $m_{K\pi}$. Assuming factorization between the different variables considered, the combinatorial *pdf* can thus be expressed as

$$\mathcal{P}_{comb}(x;\theta) = \mathcal{N}_{comb} \cdot \prod_{y} \Big(\sum c_i C_i(x) \Big) \cdot e^{-\lambda m_{K\pi\ell\ell}} \,, \tag{10.26}$$

where $y \equiv \{\cos\theta_K, \cos\theta_\ell, \phi, q^2, m(K\pi)\}$, $C_i(x)$ denotes Chebyshev polynomials of order i and $\mathcal{N}_{comb}$ is a factor needed to normalize the *pdf* to unity. This parametrization is performed separately for the two runs in the rare muon channel, but kept together in the electron one. A separate parametrization between different trigger categories is however allowed in the second case. These components are freely floating in the final fit and determined directly from data.

The application of a cut on the MVA or q_c^2 can alter the natural exponential distribution of the combinatorial background in the invariant mass. The extent of the distortion is investigated in the lepton flavour-violating sample of $K\pi\mu^{\pm}e^{\mp}$ candidates after the standard selection, where the additional requirement of `ProbNNmu`<0 and `ProbNNe`<0.2 is applied on the reconstructed electrons and muons to reduce their misidentification. As Figure 10.3 shows, the selected events maintain an exponential behaviour in the mass region of [4900,5700] MeV/c^2 regardless of the dilepton invariant mass squared variable (left) and of the different combinatorial MVA thresholds (right) used. For this reason, the exponential parametrization of Eq. 10.26 is expected to be a good representation of the combinatorial in muon and electron data samples.

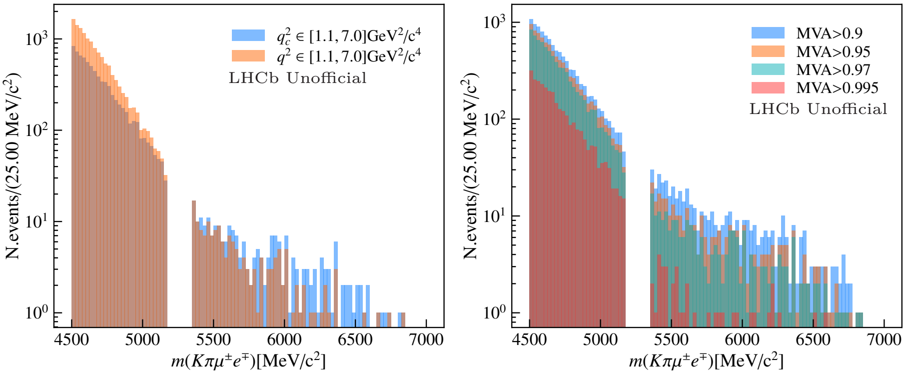


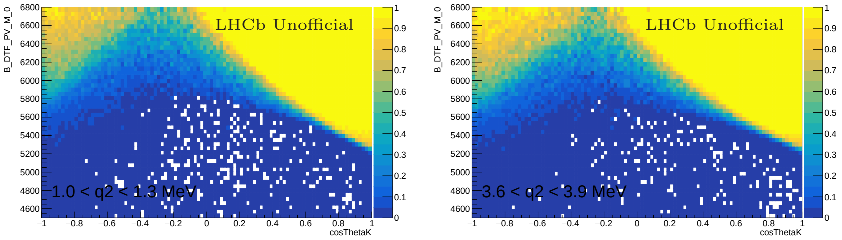
Figure 10.3: Distribution of $m(K\pi\mu^\pm e^\mp)$ mass in the central-q_c^2 region in Run 2 for the nominal working point of MVA $>$ 0.97 (left) and for different MVA thresholds (right).

10.2.1.1 Effect of $B^+ \to K^+\ell^+\ell^-$ veto

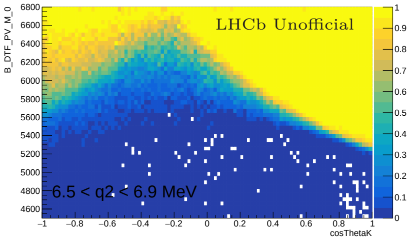
Of all the vetoes applied in Sec. 5.1.2, the one applied to remove a possible $B^+ \to K^+\ell^+\ell^-$ decays contribution has the strongest impact on the parametrization of the fit components. The veto, due to the cut performed on the invariant mass made of three out of the four final state particles, removes a certain volume of the three-dimensional space in the variables $\cos\theta_K$, q^2 and $m(K\pi\ell\ell)$. Any component that has a non-negligible overlap with this region in the fit will be distorted. This effect, being most prominent in the region above 5300 MeV/c^2, is particularly worrying for the parametrization of the combinatorial background, since it breaks the factorization hypothesis.

To resolve this issue, an additional term describing the efficiency of the veto on combinatorial events is introduced in the parametrization, such that the total *pdf* correctly takes into account the affected phase space. A three-dimensional efficiency map is computed in $\cos\theta_K$, q^2 and $m(K\pi\ell^+\ell^-)$ using the lepton flavour violating channel reconstructed as $B^0 \to K^{*0}\mu^\pm e^\mp$. Since the cut represents a purely kinematic effect, the differences in resolution between electron and muons are expected to be negligible, and the efficiency map obtained is expected to work well for the parametrization of the combinatorial components in $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays. The function $\varepsilon^{3D}_{veto}(\cos\theta_K, q^2, m_{K\pi\ell\ell})$ is defined as the ratio between the three-dimensional histogram of events that are removed by the veto over the total, and it is parametrized in the ranges $\cos\theta_K \in [-1, 1]$, $q^2 \in [1, 14]\,\mathrm{GeV}^2/c^4$ and $m(K\pi\ell^+\ell^-) \in [4500, 6800]\,\mathrm{MeV}/c^2$ with 100, 40, 50 bins, respectively. Notice that, for the electron case, q_c^2 is used instead of q^2. Figure

10.4 shows the function ε^{3D}_{veto} as a function of $\cos\theta_K$ and $m(K\pi\ell\ell)$, in three different bins of q^2.



(a) $1 < q^2 < 1.3\,\mathrm{GeV}^2/c^4$ (b) $3.6 < q^2 < 3.9\,\mathrm{GeV}^2/c^4$



(c) $6.5 < q^2 < 6.9\,\mathrm{GeV}^2/c^4$

Figure 10.4: Two-dimensional projection of ε^{3D}_{veto} as a function of $\cos\theta_K$ and $m(K\pi e^{\pm}\mu^{\mp})$ in bins of q^2.

The background *pdf* can then be fully formalized as a six-dimensional function given by

$$\mathcal{P}_{comb}(x;\theta) = \mathcal{N}_{comb}\cdot(1-\varepsilon^{3D}_{veto}(\cos\theta_K, q^2, m_{K\pi\ell\ell}))\cdot\left[\prod_y\left(\sum c_i C_i(y)\right)\cdot e^{-\lambda m_{K\pi\ell\ell}}\right]. \tag{10.27}$$

The validity of this approach is probed in muon data using the upper mass sideband region defined as $m(K\pi\mu\mu)\in[5450, 5700]\,\mathrm{MeV}/c^2$, where all the background coefficients are floated in the fit. In order to allow for enough statistics to resolve the background shape over this reduced range, the MVA cut has been relaxed to MVA > 0.2 for both runs. Figure 10.5 shows the effect of the $B^+\to K^+\ell^+\ell^-$ veto and how the term $(1-\epsilon_{veto})$ introduced in Eq. 10.27 successfully describes the event distribution. For illustration purpose, the unfolded distributions are also plotted together with the data where no veto has been applied. A good agreement is observed in all the distributions when comparing the fit result (solid red) with the vetoed dataset (black). The agreement becomes slightly worse when comparing the unfolded distribution (dashed red) to the dataset where the veto has not been applied. This is expected, since a contribution from $B^+\to K^+\mu^+\mu^-$ decays might be present in data. Even if a small excess can be observed in $\cos\theta_K$ for Run 1, a clear signature of $B^+ \to K^+\mu^+\mu^-$ decays does not seem to be present in the selected data. The current nominal strategy is to keep the veto in our final fit, since it is the safest option and no such check can be performed in the electron mode due to the worse resolution and smaller statistics.

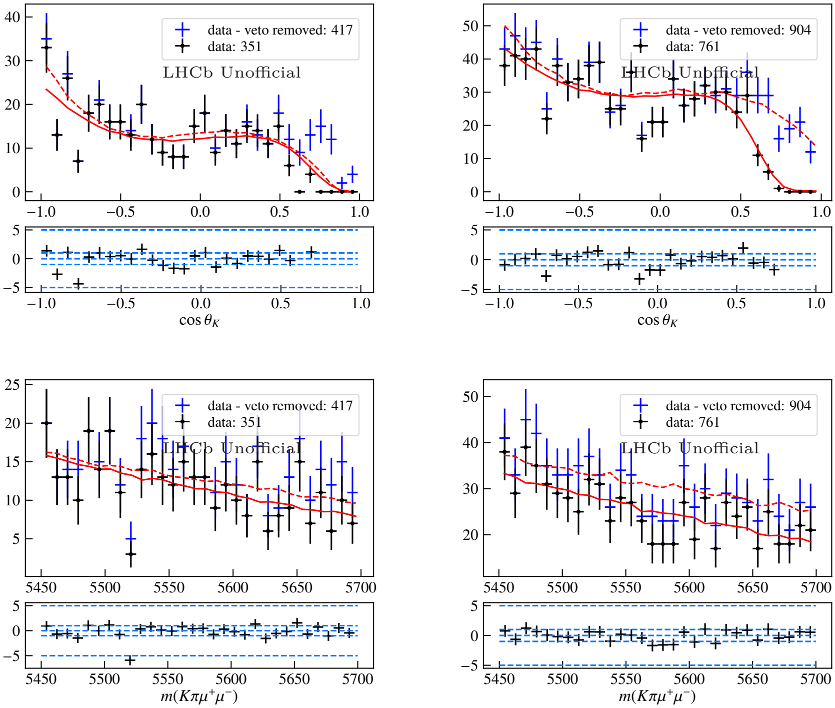


Figure 10.5: Projection of the fit to the combinatorial background events in the upper mass sideband [5450, 5700]MeV/c^2 of central-q^2 for Run 1 (left) and Run 2 (right). The MVA selection has been been loosened as described in the text, while the dashed line illustrate the resulting Chebyshev polynomials once the effect of the veto is unfolded.

10.2.2 Double-semileptonic cascade decays

The double-semileptonic cascade background (DSL) comprises decays of the type $H_b \to H_c(\to H_s\ell^-\bar{\nu}_\ell)\ell^+\nu_\ell$, with a K and a π in the final state. These decays are particularly dangerous because they present the same visible final state particles of the signal and have branching fractions of $\mathcal{O}(10^{-4})$, roughly a thousand times larger than $B^0 \to K^{*0}\ell^+\ell^-$ decays. Due to the energy loss from the missing neutrinos, these backgrounds have a distribution in $K\pi\ell^+\ell^-$ mass similar to that of the combinatorial, but with a slope that dies earlier at the invariant mass of the B^0 due to kinematics. The similarity is, however, not mirrored in the angular distribution of θ_ℓ: as the antilepton from the H_b tends to have higher momentum than the one from the H_c, and θ_ℓ by definition always refers to the angle between the H_b and this antilepton, it will generally take on values close to zero, giving rise to a distinctive, asymmetric $\cos\theta_\ell$ distribution peaked near one. This background, while present in both $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays, in the first case is easily rejected by the choice of mass range, thanks to the better momentum resolution in the reconstruction of the muons. For this reason, this background component is studied in the following only in the context of $B^0 \to K^{*0}e^+e^-$ decay fits. As already mentioned in Sec. 5.2.5, two alternatives have been used to veto this contribution in previous analyses: a cut to the invariant mass of the $K^+\pi^-e^-$ above the $D^-(H_c)$ mass at 1780 MeV/c^2 [128], and a cut on $|\cos\theta_\ell| < 0.8$ [170]. Due to the distortion or reduction of $\cos\theta_\ell$ that these two vetoes imply on the signal angular distribution, none of the two is applied here. Instead, the parametrization and inclusion of this background component in the final fit is preferred.

To investigate the best strategy to model this type of contribution, a sample of simulated $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\bar{\nu})e^+\nu$ decays has been generated and reconstructed within the LHCb software, since this channel is expected to be the single dominant contribution in this type of background. The validation in data of this assumption is of paramount importance. A possible proxy for this background in data is the lepton-flavour-violating channel $B^0 \to K^* e^\pm\mu^\mp$, where other sources of backgrounds such as partially reconstructed or J/ψ leakage are absent, while a double-semileptonic contribution involving muons and electrons is expected. A preliminary comparison between events of $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\bar{\nu})e^+\nu$ and $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)\mu^-[e^-]\bar{\nu})e^+[\mu^+]\nu$ simulated decays, where the square brackets indicate the decay where the electron and the muon are swapped, is performed to verify that the impact of resolution between different lepton generations is negligible. As Figure 10.6 shows, the three simulated decays agree well in all distributions. It is therefore reasonable to assume that the DSL contribution in the LFV data could be a good representation of the one expected in the signal rare mode. Further compatibility checks using data-driven approaches are reported in Appendix F.1.1.

A direct comparison between simulation and data is not trivial, since in the latter multiple background contributions can be present. The $\cos\theta_\ell$ angular distribution for the $K\pi e\mu$ data and $B^0 \to D^-(\to K^{*0}e^-\bar{\nu})e^+\nu$ simulation is depicted in Figure 10.7. The LFV sample has been obtained by applying the same selection as the rare mode, but considering the invariant mass region of $[4500, 5500]\,\text{MeV}/c^2$.[2] While contributions from combinatorial backgrounds are still present in the data (though dominated by DSL decays), an immediate inspection of these distributions clearly shows a significant disagreement between the two samples. In particular, the width of the peaking structure around

---

[2] Note that while the signal region where potentially the LFV channel can appear is investigated, no inspection to the invariant mass is performed. In addition, any impact of this contribution in the angular distributions are expected to be marginal. Therefore, it is safe to assume that the relevant figures for the search for this unobserved channel are preserved.

$\cos\theta_\ell$ close to +1 seems larger for data than for simulation. It has been tried to explain some of these differences by the presence of events from other DSL decays not included in the single-component simulation, such as $B^0 \to D^*(2010)e^-\bar{\nu}$, which may modify the shape and width of the main peak, or even data-driven corrections to the simulation. None of the combinations tried seems to account for the disagreement observed, as shown in Appendix F.1.2. Additionally, the simulation fails to reproduce the smaller bump at -1. Two main hypotheses have been found to explain this feature observed in data: $B_s^0 \to D_s^+(\to K^{*0}e^+\nu)e^-\bar{\nu}$ decays or the combination of a genuine DSL decay where a random K^{*0} from the underlying event is used in the reconstruction. In the first case, the $\cos\theta_\ell$ distribution will peak at -1, opposite to $B^0 \to D^-(\to K^{*0}e^-\bar{\nu})e^+\nu$ decays, due to the fact the decay of the D meson produces a K^{*0} and a positron rather than an electron. In the second case, the $\cos\theta_\ell$ distribution is expected to peak symmetrically at -1 and +1, since a random K^{*0} or $\overline{K}^{*0}$ is picked from the underlying event, and the event is reconstructed as the decay of a B^0 or $\overline{B}^0$, for which the definition of $\cos\theta_\ell$ refers to the electron or the position, respectively. This contribution is also expected to be characterized by a longer exponential tail with respect to the fully reconstructed DSL contribution, due to the combinatorial nature of the K^{*0}.

Since the rich structure seen in the LFV sample is difficult to model from simulation, a data-driven approach to obtain a parametrization for the DSL-like contribution is preferred and discussed in the Sec. 10.2.2.1. For completeness, a parametrization on $B^0 \to D^-(\to K^{*0}e^-\bar{\nu})e^+\nu$ simulated decays is also performed and reported in Appendix F.1.3. This model is used in Sec. 10.5 to estimate the systematic uncertainty linked to the choice of the DSL model. A comparison between the two parametrization is provided in Figure 10.9.

10.2.2.1 Data-driven parametrization from $B^0 \to K^{*0}e^\pm\mu^\mp$ candidates

A data-driven approach has been designed to obtain a double-semileptonic cascade background parametrization from $B^0 \to K^{*0}e^\pm\mu^\mp$ data, aiming at the isolation of this contribution from combinatorial. This is achieved by applying a very tight cut to the combinatorial MVA (MVA > 0.9975) in the lower B^0 invariant mass range of $[4500, 5200]\,\mathrm{MeV}/c^2$ and by further requiring the `ProbNNmu`<0 and `ProbNNe`<0.2 for the reconstructed electron and muon, respectively. This last cut is necessary to avoid a contribution from mis-ID events from real $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ decays. The combinatorial MVA is expected to efficiently select the double-semileptonic contribution in data and reduce the contamination of the combinatorial, due to the more signal-like features of the former with respect to the latter. These cuts are applied on top of the full set of selections. In this sample, a fit to the angular distributions, q_c^2 and $m(K\pi)$ is performed to obtain a parametrization for the DSL contribution. Due to the small mass range considered, a fit to $m(K\pi e\mu)$ is not included; its parametrization is extracted directly from the amplitude fits to the rare mode.

The *pdf* used for the parametrization corresponds to Chebyshev polynomials up to order 3, 2, 2 for $\cos\theta_K$, ϕ and q_c^2, respectively. The $m(K\pi e\mu)$ is described

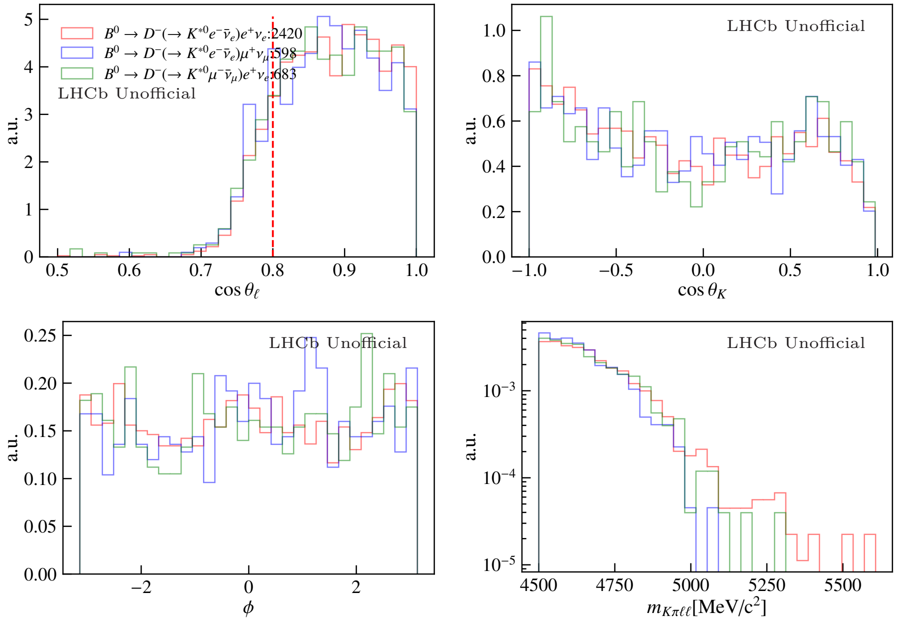


Figure 10.6: Comparison between the invariant mass and angular distributions of DSL simulations for $B^0 \to D^-(\to K^{*0}e^-\bar{\nu}_e)e^+\nu_e$ (red) and $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)\mu^-[e^-]\bar{\nu})e^+[\mu^+]\nu$ (blue or green) decays.

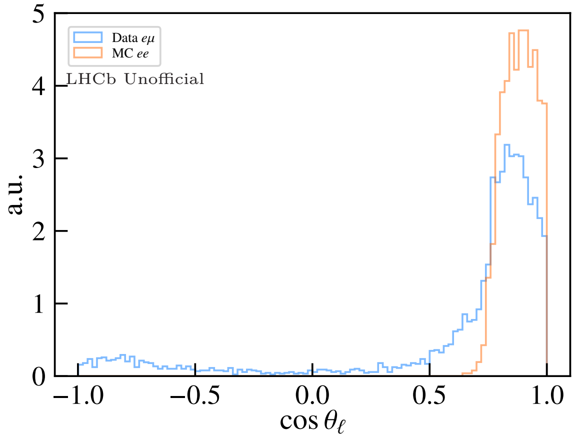


Figure 10.7: Comparison between the $K\pi e\mu$ data and $B^0 \to D^-(\to K^{*0}e^-\bar{\nu})e^+\nu$ simulation in the $\cos\theta_\ell$ distribution.

by a decreasing exponential, $\cos\theta_\ell$ by a KDE lineshape and $m(K\pi)$ by a linear term plus a Gaussian. The drop of the KDE at the boundaries of $\cos\theta_\ell$ has been mitigated by a mirroring of the simulated sample with respect to -1 and 1. The overall parametrization reads as

$$\mathcal{P}_{DSL} = \prod_{x} \Big(\sum c_i C_i(x) \Big) \cdot e^{-\lambda m} \cdot KDE(\cos\theta_\ell) \cdot f_{DSL}(m_{K\pi}) \,, \qquad (10.28)$$

where $x \equiv \{\cos\theta_K\ \phi,\ q^2\}$, $C_i(x)$ denotes Chebyshev polynomials of i^{th} order, m is the variable describing the $m(K\pi e\mu)$ distribution and

$$f_{DSL}(m_{K\pi}) = [f_{\mathcal{G}} \cdot \mathcal{G}(\mu_{K^{*0}}, \sigma_{K^{*0}}) + (1 - f_{\mathcal{G}}) \cdot \mathcal{N}_{linear}(a \cdot m_{K\pi} + b))] \,, \quad (10.29)$$

where $\mathcal{G}$ is a Gaussian *pdf* with mean $\mu_{K^{*0}}$, standard deviation $\sigma_{K^{*0}}$, and a and b are the coefficients of the straight line describing the asymmetric term needed to obtain a satisfactory description of the distribution. The factor $\mathcal{N}_{linear}$ is needed to ensure the normalization of the second part of the *pdf*. Factorization between the different dimensions considered is assumed. Figure 10.8 shows the corresponding fit results to the DSL enriched sample, whose parameters are gathered in Table 10.1. Additionally, Figure 10.9 shows the comparison between the parametrization just obtained, with the one derived from a fit to the simulated sample of $B^0 \to D^-(\to K^{*0}e^-\bar{\nu})e^+\nu$ decays, as described in Appendix F.1.3.

Table 10.1: Parameter values obtained from fits to reconstructed $B^0 \to K^* e^\pm \mu^\mp$ candidates in data enriched with a double-semileptonic contribution.

Year	RUN 1		RUN 2	
Type	L0I	L0E!	L0I	L0E!
$c^K_{1;DSL}$	-0.423 ± 0.123	-0.563 ± 0.103	-0.550 ± 0.078	-0.554 ± 0.074
$c^\phi_{1;DSL}$	-0.009 ± 0.029	-0.021 ± 0.030	0.012 ± 0.027	-0.030 ± 0.024
$c^{q^2}_{1;DSL}$	0.350 ± 0.102	0.626 ± 0.084	0.591 ± 0.067	0.481 ± 0.068
$c^K_{2;DSL}$	0.183 ± 0.096	0.003 ± 0.095	-0.017 ± 0.075	-0.051 ± 0.074
$c^\phi_{2;DSL}$	-0.016 ± 0.008	-0.003 ± 0.009	0.020 ± 0.009	0.003 ± 0.008
$c^{q^2}_{2;DSL}$	0.049 ± 0.094	0.046 ± 0.082	0.097 ± 0.064	0.042 ± 0.066
$c^K_{3;DSL}$	-0.108 ± 0.096	-0.199 ± 0.092	-0.266 ± 0.068	-0.248 ± 0.065
$f^{K\pi}_{\mathcal{G};DSL}$	0.308 ± 0.081	0.484 ± 0.082	0.399 ± 0.049	0.419 ± 0.043
a_{DSL}	0.001 ± 0.031	0.008 ± 0.130	0.000 ± 0.001	0.000 ± 0.000
$\mu^{K^{*0}}$	0.901 ± 0.005	0.902 ± 0.004	0.894 ± 0.002	0.894 ± 0.002
b_{DSL}	0.000 ± 0.003	-0.001 ± 0.007	-0.000 ± 0.000	-0.000 ± 0.000
$\sigma^{K^{*0}}$	0.023 ± 0.006	0.027 ± 0.005	0.020 ± 0.003	0.020 ± 0.002

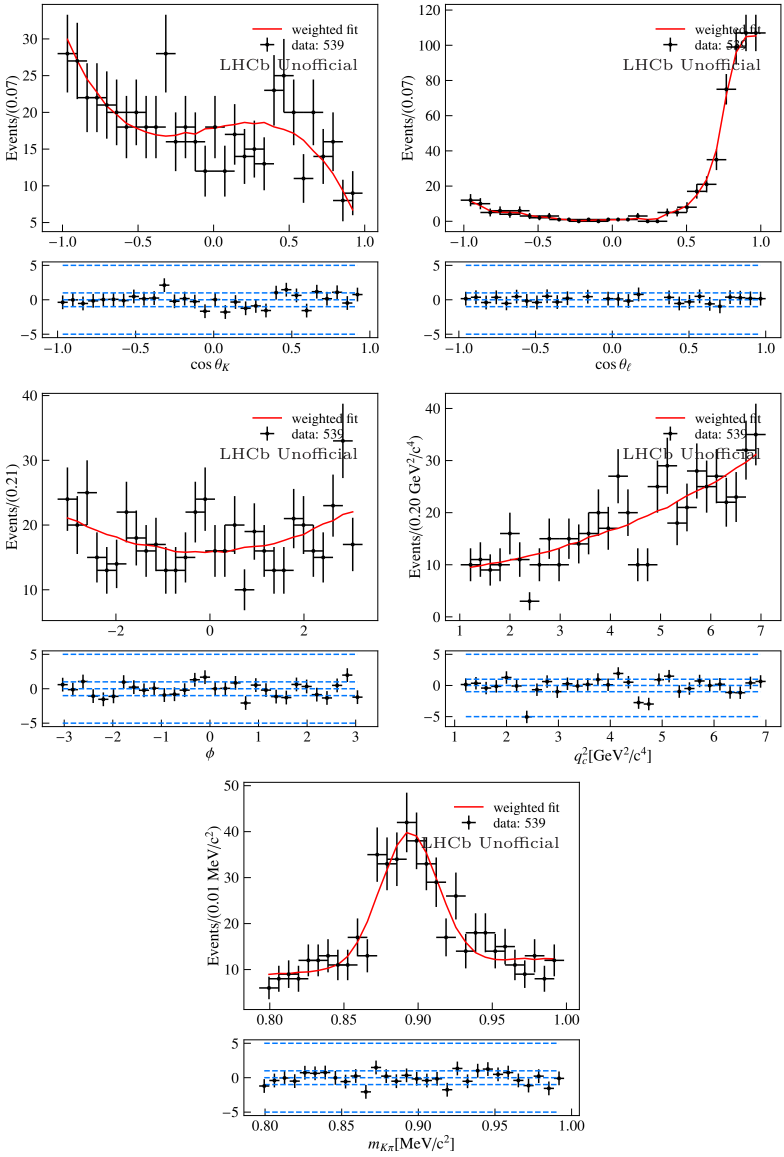


Figure 10.8: Data fits to enriched DSL samples from $B^0 \to K^{*0}e^{\pm}\mu^{\mp}$ decay candidates corresponding to Run 2 conditions in the trigger category L0I.

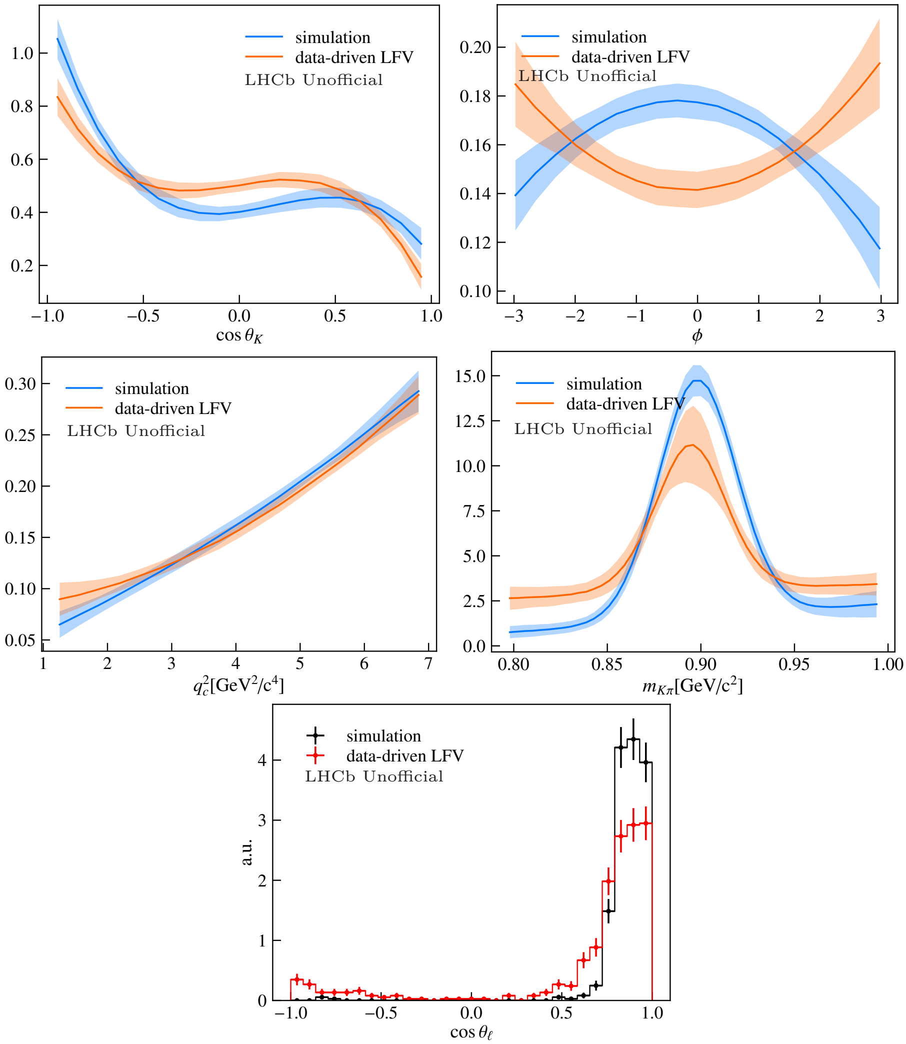


Figure 10.9: Comparison of the two parametrizations available to model the double-semileptonic background corresponding to Run 2 conditions for the trigger category L0I. The bands represent the 1σ uncertainty of the shapes as obtained from the covariance matrix of the fit result.

10.2.3 Partially reconstructed $B^+ \to K^+\pi^+\pi^-e^+e^-$ decays

In addition to double-semileptonic decays, another important source of partially reconstructed background consists of decays featuring heavier kaon resonances, such as $B^+ \to K_1^+(\to K^+\pi^+\pi^-)e^+e^-$ and $B^+ \to K_2(1430)^+(\to K^+\pi^+\pi^-)e^+e^-$, which can be reconstructed as the signal when one or more pions are not considered. Due to the missing pion(s), the shape of this partially reconstructed background in the $m(K\pi ee)$ invariant mass is that of a broad peak with its centre shifted towards the lower mass region. The modelling strategy for this background involves the use of a KDE lineshape for $m(K\pi ee)$, factorized Chebyshev polynomials up to third order for $\cos\theta_K$, second order for the remaining angles and q_c^2, and a Gaussian on top of a linear term for $m(K\pi)$, similarly to what has been done in Eq. 10.29 in the case of the double-semileptonic contribution. The KDE lineshape and the remaining fit parameters are obtained from $B^+ \to K^+\pi^+\pi^-e^+e^-$ simulated decays reconstructed under the $B^0 \to K^{*0}e^+e^-$ mass hypothesis. Since there is no model available that reproduces the full $K\pi\pi$ spectrum observed in data, the simulation has been generated flat in $m(K\pi\pi)$, and a reweighting tool has been used to obtain a data-driven correction. The correction consists of a simultaneous reweighting of $m(K^+\pi^+\pi^-)$, $m(K^+\pi^-)$ and $m(\pi^+\pi^-)$ invariant masses, performed by training a BDT on $B^+ \to K^+\pi^+\pi^-J/\psi(\to \mu^+\mu^-)$ simulated decays and $_s\mathcal{W}eighted$ data. The reweighter is then applied to the generator level information of our $B^+ \to K^+\pi^+\pi^-e^+e^-$ simulation samples, from which the weights can be ported to the reconstructed tuple via a matching procedure. Figure 10.10 shows the impact of the reweighting on the generator level distributions, displaying nicely the expected mass distributions as observed in data. The reweighted distributions of the $m(K^+\pi^-)$ and $m(K^+\pi^-e^+e^-)$ systems after reconstruction can be seen instead in Fig. 10.11. To improve the agreement between data and simulation, a smearing on the momentum of the two electrons has also been performed, accordingly to what has been described in Sec .6.6.

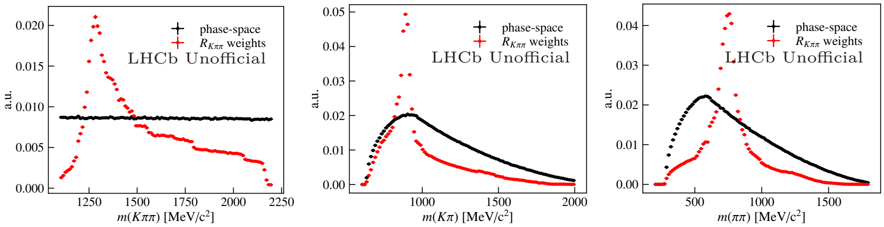


Figure 10.10: Distribution of $m(K^+\pi^+\pi^-)$, $m(K^+\pi^-)$ and $m(\pi^+\pi^-)$ as obtained in the generator level phase space simulation of $B^+ \to K^+\pi^+\pi^-e^+e^-$ before (black) and after (red) the reweighting procedure to reproduce the $K\pi\pi$ spectrum observed in data.

Additionally, two more corrections to the simulation are considered after the reweighting of the $K\pi\pi$ spectrum: tracking and PID corrections, as

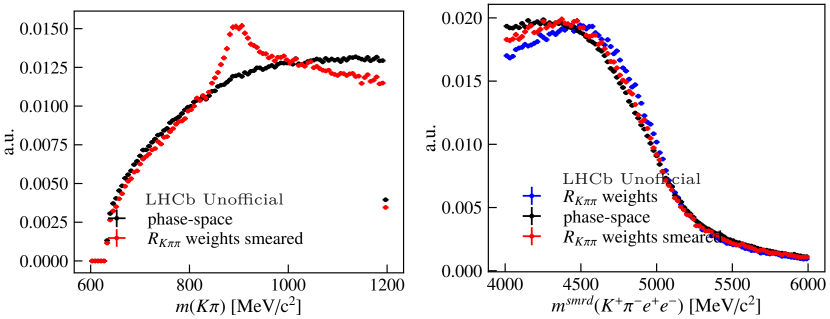


Figure 10.11: Distribution of $m(K^+\pi^-)$ and $m^{smrd}(K^+\pi^-e^+e^-)$ in the reconstructed phase space simulation of $B^+ \to K^+\pi^+\pi^-e^+e^-$. The black and red distributions show the impact of the reweighting on the two reconstructed variables while in blue (only on the right) is highlighted the impact of removing the smearing in the $K^+\pi^-e^+e^-$ invariant mass.

already discussed in Ch. 6. The simulated sample of $B^+ \to K^+\pi^+\pi^-e^+e^-$ decays reweighted, smeared and after the full selection used for the signal rare mode, is then used to obtain a six-dimensional parametrization of the partially reconstructed background in the region of the signal. Figure 10.12 shows these parametrizations for the trigger category L0I in Run 2, while the other fit results are omitted for brevity. Table 10.2 reports the values of the parameters obtained from the fits.

Table 10.2: Parameter values obtained from fits to $B^+ \to K^+\pi^+\pi^-e^+e^-$ in Run 1 and Run 2 for the trigger categories L0I and L0E!

Year	Run 1		Run 2	
Type	L0I	L0E!	L0I	L0E!
$c^K_{1;PR}$	-0.886 ± 0.030	-0.760 ± 0.035	-0.951 ± 0.014	-1.032 ± 0.014
$c^\ell_{1;PR}$	-0.019 ± 0.037	0.129 ± 0.039	-0.035 ± 0.019	0.024 ± 0.019
$c^\phi_{1;PR}$	-0.007 ± 0.012	0.005 ± 0.012	0.007 ± 0.006	0.019 ± 0.006
$c^{q^2}_{1;PR}$	-0.094 ± 0.037	-0.012 ± 0.038	-0.057 ± 0.019	-0.000 ± 0.019
$c^K_{2;PR}$	0.025 ± 0.036	-0.060 ± 0.038	0.142 ± 0.017	0.230 ± 0.016
$c^\ell_{2;PR}$	-0.312 ± 0.040	-0.309 ± 0.043	-0.285 ± 0.020	-0.341 ± 0.021
$c^\phi_{2;PR}$	0.001 ± 0.004	-0.004 ± 0.004	0.001 ± 0.002	0.004 ± 0.002
$c^{q^2}_{2;PR}$	-0.071 ± 0.037	0.114 ± 0.035	-0.047 ± 0.018	0.014 ± 0.019
$c^K_{3;PR}$	-0.094 ± 0.031	0.052 ± 0.035	-0.155 ± 0.014	-0.125 ± 0.014
$f^{K\pi}_{\mathcal{G};PR}$	0.539 ± 0.052	0.476 ± 0.034	0.422 ± 0.013	0.399 ± 0.017
a_{PR}	0.010 ± 0.149	-0.001 ± 0.000	-0.028 ± 0.042	-0.034 ± 0.039
$\mu^{K^{*0}}_{PR}$	0.888 ± 0.002	0.892 ± 0.002	0.890 ± 0.001	0.890 ± 0.001
b_{PR}	-0.002 ± 0.019	0.001 ± 0.001	0.049 ± 0.067	0.053 ± 0.058
$\sigma^{K^{*0}}_{PR}$	0.034 ± 0.003	0.026 ± 0.002	0.023 ± 0.001	0.024 ± 0.001

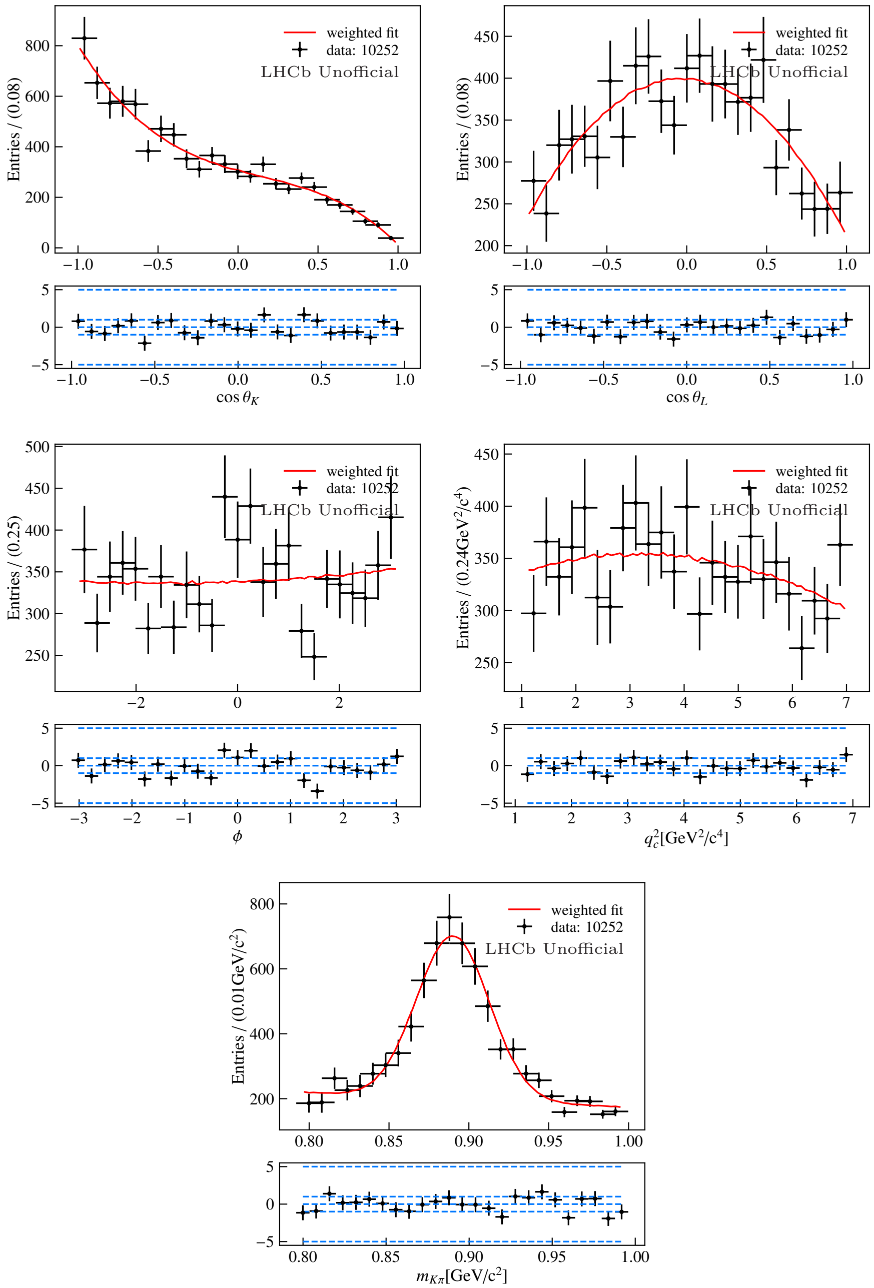


Figure 10.12: Result of the fit to the weighted and smeared $B^+ \to K^+\pi^+\pi^-e^+e^-$ simulation sample in Run 2 for the trigger category L0I.

This reweighting strategy only takes the mass shapes into account and does not explicitly correct for angular distributions. It is known, however, that the spin of both intermediate resonances and missing final state particles can have a significant impact on the shape of partially reconstructed backgrounds. To account for these differences, an alternative parametrization of the partially reconstructed background is obtained using simulated samples that mimic the underlying dynamics of the $B^+ \to K_1^+ e^+ e^-$ decay. While the fit results are detailed in Appendix F.2, a comparison of the two parametrizations available is shown in Figure 10.13 for simulated samples corresponding to Run 2 in the L0I trigger category. As it can be seen, while there is a small difference between the shapes of the two modes in the reconstructed $m(K\pi ee)$ (histogrammed here since it is parametrized with a KDE rather than a fit), a clear difference can be noticed between the two parametrizations in the other variables of interest. The impact of such differences is studied in toys and a systematic uncertainty based on the choice of the model is assigned.

10.3 Constraint on the observed signal yield

Additional stability is provided to the fit when the $B^0 \to K^+\pi^-\ell^+\ell^-$ branching ratio information is used as a constraint on the number of signal events observed. In fact, the total signal yield $\hat{N}_{sig}$ (P+S waves) directly enters the extended likelihood as shown in Eq. 10.35 and can be expressed as

$$\hat{N}_{sig} = N_{J/\psi} \cdot (1 - F_S^{J/\psi}) \cdot \frac{\mathcal{B}(B^0 \to K^+\pi^-\ell^+\ell^-)}{\mathcal{B}(B^0 \to J/\psi K^{*0}) \cdot \mathcal{B}(J/\psi \to \ell^+\ell^-)} \cdot \underbrace{\frac{\epsilon_{tot}^{sig}}{\epsilon_{tot}^{J/\psi}}}_{R_\varepsilon} . \quad (10.30)$$

Notice that the term $\mathcal{B}(K^{*0} \to K^+\pi^-)$ is omitted since it cancels out in the ratio. The terms appearing in the previous equation can be split in two categories: some corresponding to external constraints and some that are a function of the parameters floating in the fit. In the first category we find:

- $N_{J/\psi}$, yield of the corresponding resonant mode obtained from the mass fits described in Sec. 8.3.2, used to normalize the branching ratio of the signal;
- $F_S^{J/\psi}$, the S-wave fraction in the control channel. This could, for consistency, be directly extracted from the amplitude fit in Sec. 9.2.3. However, the difference between using this or an external value is expected to be small, since the uncertainty on this quantity has been shown to be dominated by a systematic uncertainty of 0.01 due to the difficulties in modelling exotic contributions from $B^0 \to Z(4430)^- K^+$ decays [164]. For this reason, the value used corresponds to $F_S = (6 \pm 1)\%$ [164].
- $\mathcal{B}(B^0 \to J/\psi K^{*0})$, the branching ratio of the resonant channel, which is fixed to the value in Ref. [161].

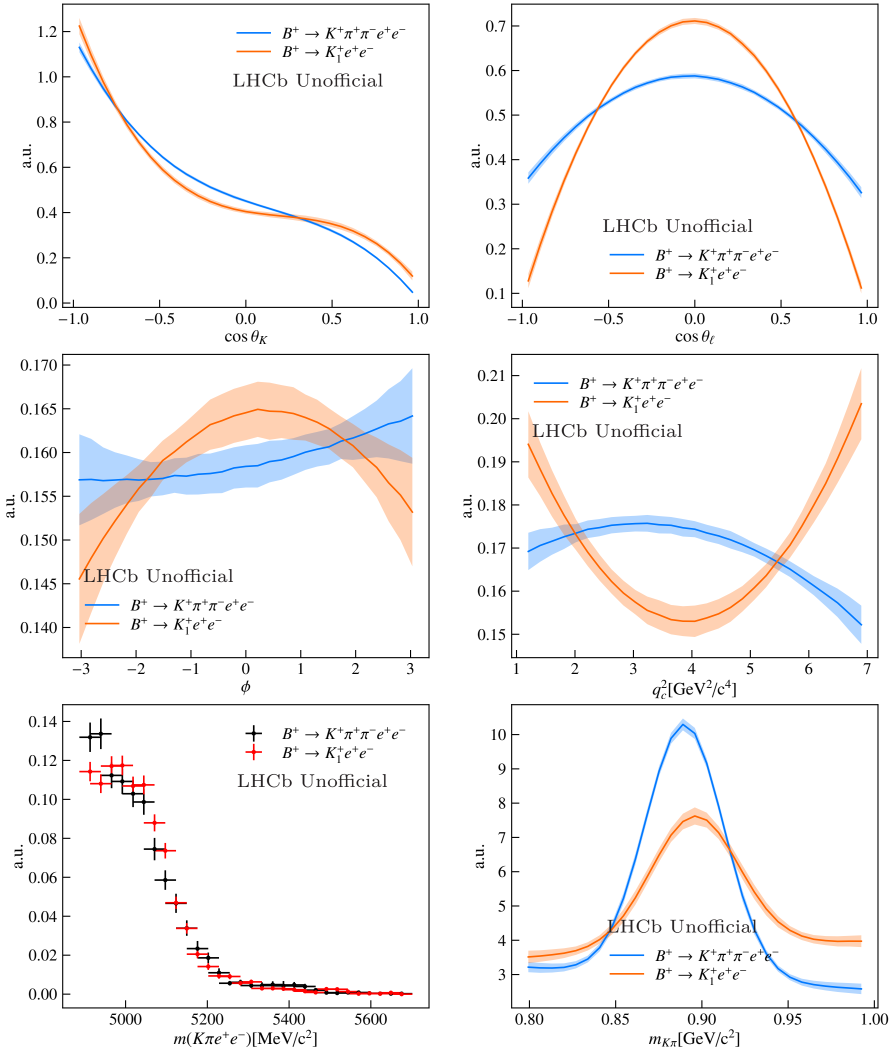


Figure 10.13: Comparison of the two parametrizations available to model the partially reconstructed background corresponding to Run 2 and the trigger category L0I. The bands represent the 1σ uncertainty of the shapes obtained from the covariance matrix of the fit result.

- $\mathcal{B}(J/\psi \to \ell^+\ell^-)$, the branching ratio of the J/ψ decaying in the dileptons, fixed from PDG [129].
- R_ε, the ratio of the total efficiencies between the rare and resonant mode, defined in Sec. 7.1.2.

In the second category, we are left with the term $\mathcal{B}(B^0 \to K^+\pi^-\ell^+\ell^-)$, corresponding to the branching ratio of the rare mode, where both P- and S-wave are considered. This quantity is given by

$$\mathcal{B}(B^0 \to K^+\pi^-\ell^+\ell^-) = \frac{\tau_B}{\hbar}\int_{\substack{\Delta q^2\\ \Delta m_{K\pi}}} \frac{\mathrm{d}^2\Gamma(\vec{\Delta\mathcal{C}};\theta_{nuis})}{\mathrm{d}q^2\mathrm{d}m_{K\pi}}\Big|_{S+P}\mathrm{d}q^2\mathrm{d}m_{K\pi}\,, \tag{10.31}$$

where the integration is performed in Δq^2, a region of q^2 corresponding to central-q^2 and in $\Delta m_{K\pi}$, a region of $m_{K\pi}$ corresponding to an interval of $\pm 100\,\mathrm{MeV}/c^2$ from the nominal mass of the $K^{*0}(892)$ resonance. For clarity, the dependency on the $\vec{\Delta\mathcal{C}} = (\Delta\mathcal{C}_9, \Delta\mathcal{C}_{10})$ and θ_{nuis} (vector indicating the remaining nuisance parameters) is explicit. It is important to notice that the expression in Eq. 10.30 is an approximation of the more rigorous

$$\begin{aligned} N_{sig} =& \frac{N_{J/\psi}\cdot(1-F_S^{J/\psi})}{\mathcal{B}(B^0 \to J/\psi K^{*0})\cdot\mathcal{B}(J/\psi \to \ell^+\ell^-)\cdot\epsilon_{tot}^{J/\psi}} \\ &\times\frac{\tau_B}{\hbar}\int \mathrm{d}x\;\epsilon_{eff}^{sig}(x)\cdot\frac{\mathrm{d}^5\Gamma(\vec{\Delta\mathcal{C}};\theta_{nuis})}{\mathrm{d}x}\Big|_{S+P}\,, \end{aligned} \tag{10.32}$$

where $x \equiv \{\cos\theta_K, \cos\theta_\ell, \phi, q^2, m_{K\pi}\}$ and $\epsilon_{eff}^{sig}(x)$ corresponds to the effective acceptance discussed in Eq. 7.9, properly normalized to represent the differential absolute efficiency. The two Eqs. 10.30 and 10.32 are identical only under the hypothesis that the underlying physics distribution in data is identical to the one generated in the simulation, or expressed differently

$$\epsilon_{tot}^{sig} \overset{!}{=} \int \mathrm{d}x\;\epsilon_{eff}^{sig}(x)\cdot\frac{\mathrm{d}^5\Gamma(\vec{\Delta\mathcal{C}};\theta_{nuis})}{\mathrm{d}x}\Big|_{S+P}\Bigg/\int \mathrm{d}x\frac{\mathrm{d}^5\Gamma(\vec{\Delta\mathcal{C}};\theta_{nuis})}{\mathrm{d}x}\Big|_{S+P}\,. \tag{10.33}$$

In the following, to simplify an already complicated likelihood, we decided to perform the fit using as a constraint the simplified relationship in Eq. 10.30. The impact of such assumption is considered as a source of systematics in Sec. 10.6.

10.4 Amplitude fits to data

An extended maximum likelihood fit is performed in data, simultaneously in the electron and muon channel, split in six subsamples corresponding to: the region of central-q^2 and central-q^2_c for $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays, respectively; two different trigger categories for $B^0 \to K^{*0}e^+e^-$ decays (L0I and L0E!), while L0I and L0M! are merged for $B^0 \to K^{*0}\mu^+\mu^-$ decays; and two different runs, corresponding to Run 1 and Run 2. The invariant mass regions chosen are $m(K\pi ee) \in [4900, 5700]\,\mathrm{MeV}/c^2$ and $m(K\pi\mu\mu) \in [5170, 5700]\,\mathrm{MeV}/c^2$, while the MVA response is required to be larger than 0.97 and 0.8, respectively, as suggested by the optimization process in Sec 5.2.6. The $m(K\pi)$ region is chosen to be $[796{,}996]\,\mathrm{MeV}/c^2$.

The full likelihood can be expressed as

$$\mathcal{L}_{tot}(\theta) = \mathcal{L}_{\mathrm{constr}} \cdot \prod_{r\in\{\textsc{Run}\,1,\textsc{Run}\,2\}} \mathcal{L}^{\mu}_{r}(\theta) \cdot \prod_{\substack{r\in\{\textsc{Run}\,1,\textsc{Run}\,2\}\\ t\in\{\mathrm{L0I},\mathrm{L0E!}\}}} \mathcal{L}^{e}_{rt}(\theta_{nuis})\,, \tag{10.34}$$

with the single likelihood defined as

$$\mathcal{L}^{\ell}_{rt}(\theta) = \frac{e^{-\hat{N}_{tot}} \cdot \hat{N}_{tot}^{N_{obs}}}{N_{obs}!} \cdot \prod_{i}^{N_{obs}} \left[\frac{\hat{N}_{sig}(\theta)}{\hat{N}_{tot}} \mathcal{P}_{sig}(x;\theta) + \sum_{k} \frac{\hat{N}_k}{\hat{N}_{tot}} \mathcal{P}_k(x;\theta) \right], \tag{10.35}$$

where the *pdf*s are parametrized as a function of $x \equiv \{\vec{\Omega}, q^2, m(K\pi), m(K\pi\ell\ell)\}$; $\mathcal{P}_{sig}(x;\theta)$ is the full signal *pdf*, defined in Sec. 10.1.3; $\mathcal{P}_k(x;\theta)$ is the *pdf* corresponding to the k^{th} background component included in the fit, whose possible parametrizations were described in Sec. 10.2; N_{obs} is the total number of candidates observed in the subsample considered; $\hat{N}_{sig}$ corresponds to the estimator for the number of signal events observed, expressed as a function of the Wilson coefficients, as discussed in Eq. 10.30, and $\hat{N}_k$ corresponds to the estimator for the number of observed events for the k^{th} background component. The likelihood term $\mathcal{L}_{\mathrm{constr}}$ describes instead the effect of the multidimensional Gaussian constraints applied on the parameters used to describe the form factors and CKM parameters, as explicitly shown in the following

$$\begin{aligned}\mathcal{L}_{\mathrm{constr}} =& \frac{1}{\sqrt{(2\pi)^{19}\mathrm{det}|C|}} e^{-\frac{1}{2}(\vec{\phi}-\vec{\phi}^{\mathrm{constr}})^T C^{-1}(\vec{\phi}-\vec{\phi}^{\mathrm{constr}})} \\ & \times \prod_j \frac{1}{\sqrt{2\pi}\sigma_j^{\mathrm{constr}}} e^{-\frac{1}{2}\left(\frac{\eta_j-\eta_j^{\mathrm{constr}}}{\sigma_j^{\mathrm{constr}}}\right)^2}\,,\end{aligned} \tag{10.36}$$

where C represents the covariance matrix between the different form factor parameters ϕ, ϕ^{constr} corresponds to the central values of the 19 form factor parameters considered, σ_j^{constr} is the standard deviation of the j^{th} CKM parameter η_j, which is centred with a Gaussian distribution to is expected value η_j^{constr}.

The optimization is performed with respect to the fit parameters θ, containing both the parameters of interest $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, and nuisance parameters, θ_{nuis}. When considering the signal component, the nuisance parameters of the fit correspond to:

- The effective Wilson coefficients $\tilde{\mathcal{C}}^e_9$, $\tilde{\mathcal{C}}^e_{10}$, whose estimation could be polluted by a mismodelling of the charm-loop contribution. Since no contributions from the complex part of the WC is considered here, all WCs are assumed to be real.

- The Wilson coefficient $\mathcal{C}_7$, whose value is fixed to the SM ($\mathcal{C}_7 = -0.34$ [23]) due to the strong constraint obtained from radiative B meson decays [127].

- The chirality-flipped Wilson coefficients $\mathcal{C}_i^{e\prime}$ and $\Delta\mathcal{C}_i^{\prime}$, which are set to zero and not floated the fit.

- The Wolfenstein parameters $\{\alpha_m^{CKM}\} \equiv \{\lambda, A, \bar{\rho}, \bar{\eta}\}$, used to parametrize the CKM elements that enter the amplitudes. They are Gaussian constrained to the values obtained from a SM fit of the unitarity triangle [171].

- The P-wave form-factors parameters $\{\alpha_l^{(FF)}\} \equiv \{V, A_{0,\ldots,2}, T_{1,\ldots,3}\}$ (see Appendix A). They are Gaussian constrained from a combined fit to Light-Cone Sum Rules and lattice QCD results as determined in Ref. [34].

- The P-wave non-local hadronic parameters (real and imaginary parts) $\{\alpha_k^{(\lambda)}\}$ for each polarization $\lambda = \perp, \|, 0$ at a given order of the analytical expansion z^k (see Appendix B). They are free to float in the fit and are expected to describe the charm-loop contribution to the signal amplitude. In the following, an expansion up to z^2 is considered as the nominal configuration.

- The relative magnitude (g_S) and phase (δ_S) between the P- and S-wave in the $K\pi$ system. Notice that these terms are needed to account for the normalization of the S-wave with respect to the P-wave, since an analytical normalization in $m_{K\pi} \in (-\infty, \infty)$ is not allowed for the expression in Eq. 10.9.

- The shift and scale factors of data with respect to the simulation for the muon channel, s_σ^μ, $\Delta\mu^\mu$. The shift and the scaling for the electron channels are obtained from the resonant modes and fixed in the final fit.

This corresponds to a total of 47 nuisance parameters, assuming that the non-local hadronic model can be curtailed at z^2. In addition to the above-mentioned nuisance parameters floated in the fit, the following set of parameters (all belonging to the S-wave contribution) are fixed in the nominal fit configuration.

- The S-wave form factor parameters $\{\alpha_s^{(FF)}\} \equiv \{f_+, f_T, f_0\}$. For more information see Appendix. D.

- The S-wave non-local hadronic parameters $\{\alpha_k^{S0}\}$ are fixed to zero, *i.e.* the contribution of non-local hadronic effects is ignored in the S-wave amplitudes.

When including the parametrization of the background contributions in the full fit, the number of total nuisance parameters grows to 109 and includes:

- The yield and the shape parameters for the combinatorial background in $B^0 \to K^{*0}\mu^+\mu^-$ decays, corresponding to 24 free parameters when both runs are considered.

- The yield and the shape parameters for the combinatorial background in $B^0 \to K^{*0} e^+ e^-$ decays, corresponding to 26 free parameters when both runs and trigger categories are considered. To improve the stability of the fit, the shapes have been allowed to be different between trigger categories, but are shared between different runs. This has been motivated by a study performed on the same-sign sample and described in Appendix F.3.
- The yields of the double-semileptonic and partially reconstructed decays, whose full parametrization, described in Sec. 10.2.2 and 10.2.3, respectively, is fully fixed from simulation or control mode data. The only exception are the exponential slopes of the DSL components, floated in data. Systematic uncertainties arising from the choice of these models will be discussed in Sec. 10.6.

A summary of the parameters floated in the fit can be found in Table 10.3. To be sure to find the global minimum, each fit is repeated $\mathcal{O}(20)$ times with randomized starting parameters. All the parameters, which have a Gaussian constraint in the fit, *e.g.* form factors and CKM parameters, are initialized uniformly in a [$+5\sigma$, -5σ] range. For all unconstrained parameters, a sufficiently large starting range is chosen, namely [2, 6] and [-6, -2] for $\tilde{\mathcal{C}}_9^e$ and $\tilde{\mathcal{C}}_{10}^e$, respectively, and [-2, 1] for ΔC_9 and [-1, 2] for ΔC_{10}. The solution with the smallest negative log-likelihood is taken as the default result.

The results of the amplitude fit to $B^0 \to K^{*0} \ell^+ \ell^-$ candidates in data are shown, split for the different runs and trigger categories, in Appendix G. For simplicity, Figures 10.14, 10.15, 10.16 show the fit result when the two runs are merged together. The projections of the fit results are overlaid on the selected data for the merged datasets. As can be seen, a good agreement is observed in all projections. Some tensions can be noticed in the regions dominated by the DSL background, but their impact is expected to be accommodated by the systematic uncertainty associated to the modelling of this background component. The values of the observables of interest are kept blinded.

Having obtained the maximum likelihood estimate of the parameters of interest, $\hat{\theta}$, we then proceed to the determination of the corresponding confidence intervals. In most practical applications, if the number of measurements is big enough, the negative log-likelihood is, with good approximation, a multidimensional parabola in the fit parameters θ. The covariance matrix C_{ij} can, in this approximation, be obtained from the inverse of the second-order partial derivatives of the negative log-likelihood

$$C_{ij}^{-1} = -\frac{\partial \ln \mathcal{L}_{tot}(x;\theta)}{\partial\theta_i \partial\theta_j}\,. \tag{10.37}$$

An alternative possibility is to determine the interval for which the negative log-likelihood increases of 0.5 with respect to its minimum. It can be shown that this method leads to identical results as Eq. 10.37 in the case of a parabolic function but, for a more generic shape, it allows to better follow the behaviour of the likelihood allowing for asymmetric uncertainties to be

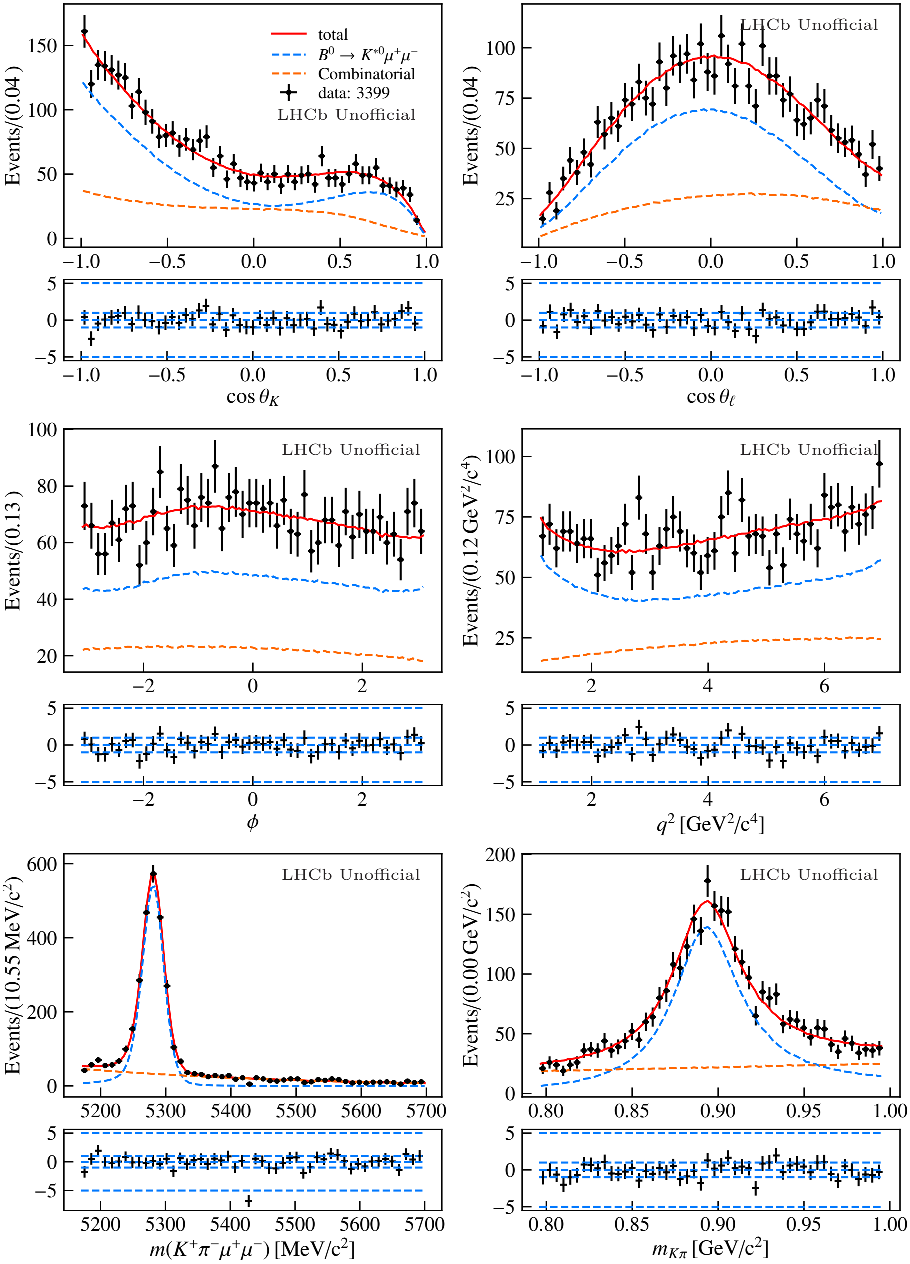


Figure 10.14: Projection of the fit result for the $B^0 \to K^{*0}\mu^+\mu^-$ amplitude fit in central-q^2. A combination of the RUN 1 and RUN 2 datasets and fit results is shown.

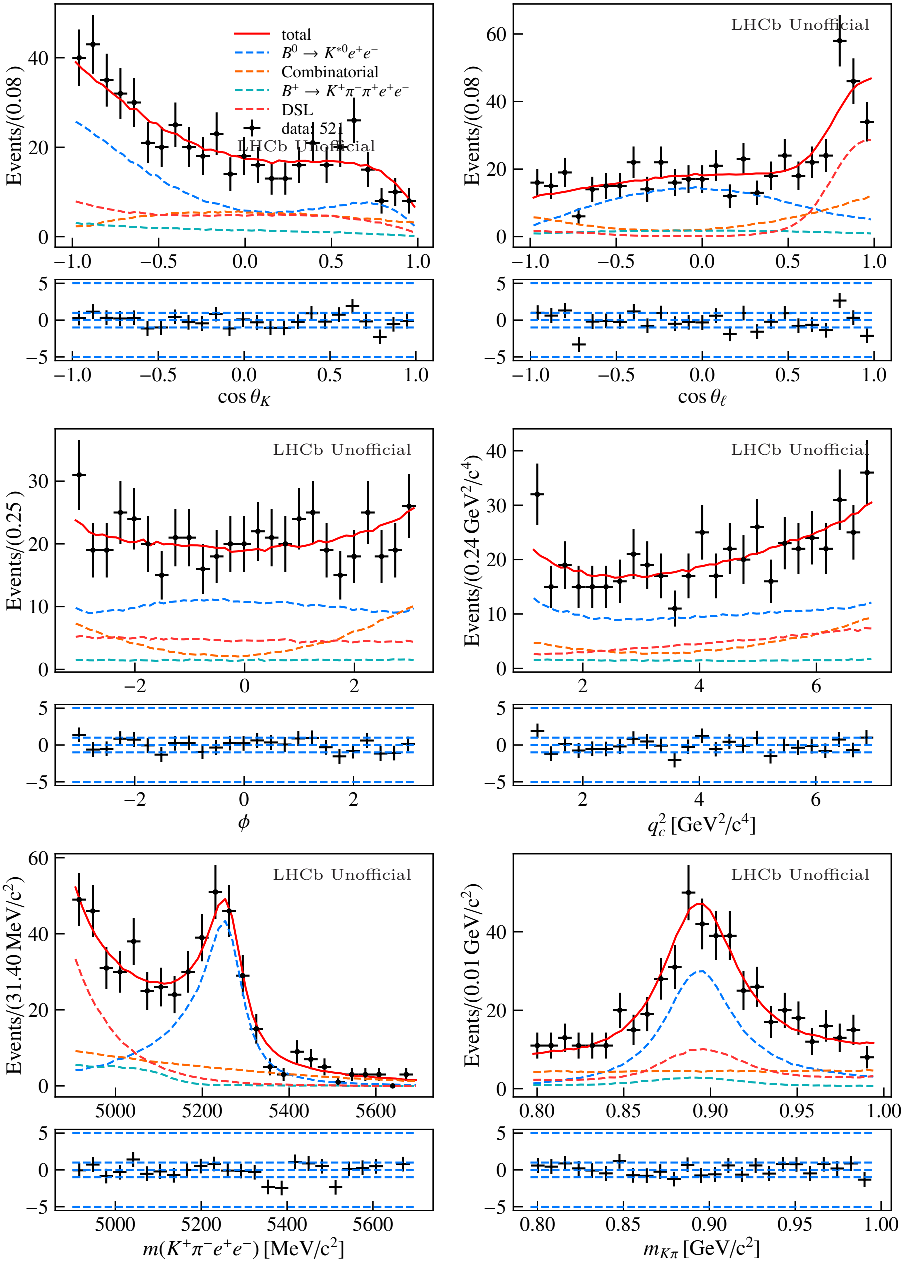


Figure 10.15: Projection of the fit result for the $B^0 \to K^{*0}e^+e^-$ amplitude fit for the trigger category L0E! in central-q_c^2. A combination of the Run 1 and Run 2 datasets and fit results is shown.

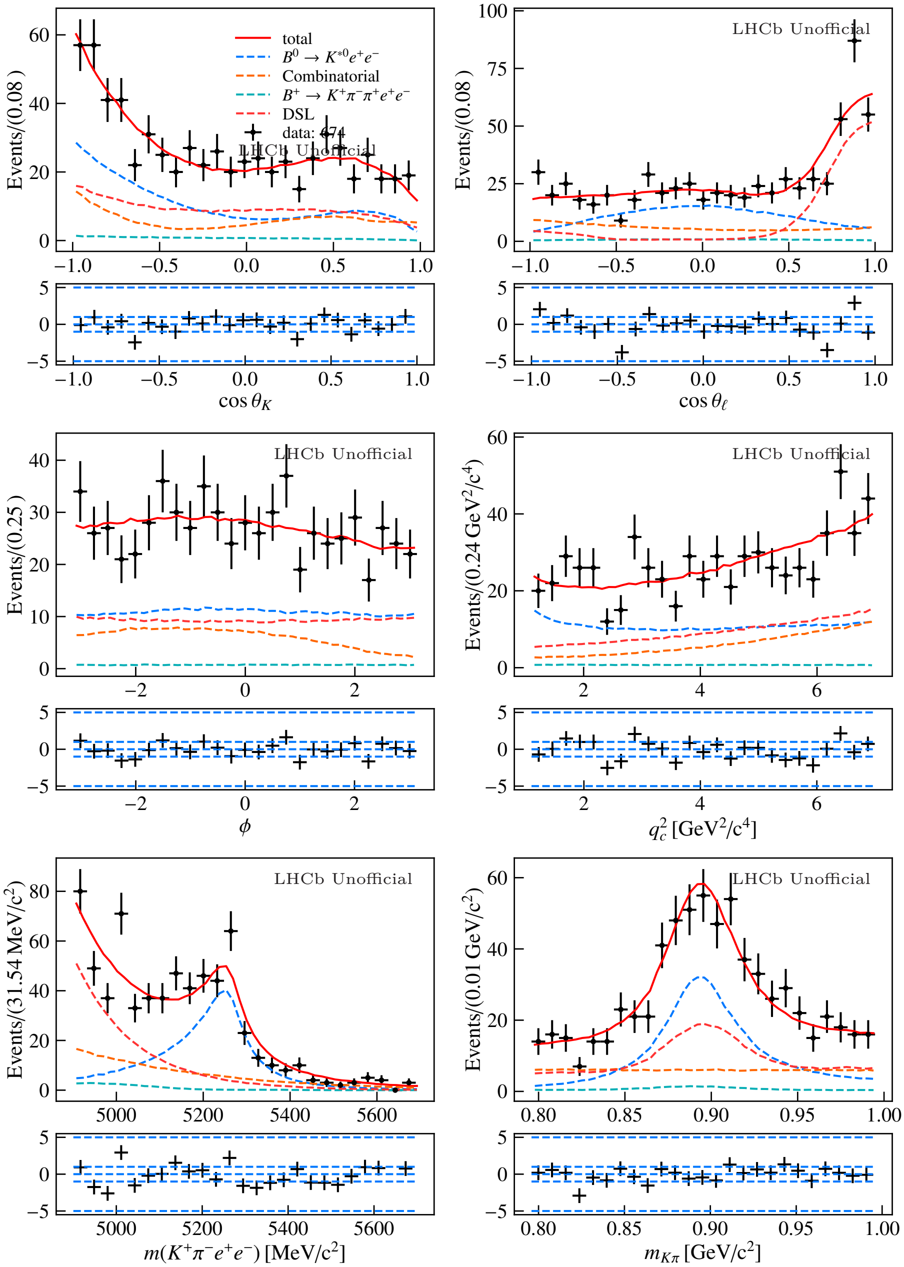


Figure 10.16: Projection of the fit result for the $B^0 \to K^{*0}e^+e^-$ amplitude fit for the trigger category L0I in central-q_c^2. A combination of the Run 1 and Run 2 datasets and fit results is shown.

Table 10.3: Summary of the parameters floated in the amplitude data fits to $B^0 \to K^{*0}\ell^+\ell^-$ decays.

	Parameters
Floating	• s^{μ}_{σ}, $\Delta\mu^{\mu}$ for each run • $\mathcal{N}^{\mu}_{comb}$ for each run, • $\mathcal{N}^{e}_{comb}$, $\mathcal{N}_{B\to K\pi\pi e^+e^-}$, $\mathcal{N}_{DSL}$ for each run and trigger category • $\tilde{\mathcal{C}}^{e}_{9}$, $\tilde{\mathcal{C}}^{e}_{10}$, ΔC_9, ΔC_{10}, • $\alpha^{(\lambda)}_{k}$, where k is the order of the z-expansion used • g_S, δ_S • λ^{μ}_{comb}, $c^{K;\mu}_{l,comb}$, $c^{\ell;\mu}_{l,comb}$, $c^{\phi;\mu}_{l,comb}$, $c^{q^2;\mu}_{l,comb}$, $c^{m_{K\pi};\mu}_{l,comb}$ for each run, • $c^{K;e}_{l,comb}$, $c^{\ell;e}_{l,comb}$, $c^{\phi;e}_{l,comb}$, $c^{q^2;e}_{l,comb}$, $c^{m_{K\pi};e}_{l,comb}$ for each trigger category, • λ^{e}_{comb}, λ^{e}_{DSL}, for each run and trigger category
Gaussian constr.	• λ, A, $\bar{\rho}$, $\bar{\eta}$ • $V \equiv \{V^0, V^1, V^2\}$ • $A_0 \equiv \{A_0^0, A_0^1, A_0^2\}$, $A_1 \equiv \{A_1^0, A_1^1, A_1^2\}$, $A_2 \equiv \{A_2^1, A_2^2\}$, • $T_1 \equiv \{T_1^0, T_1^1, T_1^2\}$, $T_2 \equiv \{T_2^1 T_2^2\}$, $T_3 \equiv \{T_3^0, T_3^1, T_3^2\}$

computed. A multidimensional scan of the likelihood can be very computationally expensive, since it additionally provides information not only on the uncertainties of the parameters of interest $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, but also on all the nuisance parameters, including their correlation. An alternative approach is to obtain a confidence interval for the parameters of interest performing a "profiling" of the likelihood [172]. This procedure consists in repeating the minimization of the negative log-likelihood with respect to nuisance parameters, while the parameters of interest are kept fixed. Mathematically, this corresponds to obtaining the 1D or 2D profile functions $\Delta\mathrm{NLL}(\Delta\mathcal{C}_{9,10})$ and $\Delta\mathrm{NLL}(\Delta\mathcal{C}_9, \Delta\mathcal{C}_{10})$, which correspond to:

$$\Delta\mathrm{NLL}(\Delta\mathcal{C}_{9,10}) = \min_{\theta_{nuis},\Delta\mathcal{C}_{10,9}} \Big[-\ln\Big(\frac{\mathcal{L}_{tot}(\Delta\mathcal{C}_{9,10}; \Delta\mathcal{C}_{10,9}, \theta_{nuis}}{\mathcal{L}_{tot}(\hat{\Delta\mathcal{C}_9}, \hat{\Delta\mathcal{C}_{10}}, \hat{\theta}_{nuis})} \Big) \Big] \tag{10.38}$$

and

$$\Delta\mathrm{NLL}(\Delta\mathcal{C}_9, \Delta\mathcal{C}_{10}) = \min_{\theta_{nuis}} \Big[-\ln\Big(\frac{\mathcal{L}_{tot}(\Delta\mathcal{C}_9, \Delta\mathcal{C}_{10}; \theta_{nuis}}{\mathcal{L}_{tot}(\Delta\hat{\mathcal{C}}_9, \Delta\hat{\mathcal{C}}_{10}, \hat{\theta}_{nuis})}\Big)\Big]. \tag{10.39}$$

It is important to notice that this function is not a negative log-likelihood function itself, but rather corresponds to a test statistics known as *likelihood-ratio*[172]. Under the assumption that the hypotheses of Wilk's theorem [173] hold, it is possible to build the confidence intervals for the variables of interest, exploiting the fact that this statistical test, when multiplied by two, asymptotically corresponds to a χ^2 distribution with a number of degrees of freedom equal to the number of parameters of interest. In practice, when performing a one-dimensional profiling, the uncertainty for the parameter of interest considered corresponds to the distance from the minimum of the profile function to the points where 2ΔNLL has increased of one unit. For a two-dimensional profiling, the confidence interval corresponds instead to the curve where the function 2ΔNLL has increased of 2.296 with respect to its minimum. The values corresponding to the 2σ and 3σ confidence intervals are reported in Table 10.4 for likelihoods profiled in 1D and 2D.

Table 10.4: Difference of negative log-likelihood used to determine the 1, 2, 3 σ contours for onw and two-dimensional projections of WCs. Values are obtained from the inverse CDF of a χ^2 distribution with one and two degrees of freedom.

σ	Coverage	2ΔNLL	
		1D	2D
1	68.27%	1	2.296
2	95.45%	4	6.18
3	99.73%	9	11.83

The 1D likelihood profiles obtained in data for $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, together with the corresponding confidence intervals, are shown in Figure 10.17 and reported in Table 10.5. Overlaid to the points is shown, in blue, the sixth-order polynomial function used to approximate the likelihood profile. The possibility to limit the parametrization to a parabola has been considered, but was found to provide an unsatisfactory description of the profiled points due their asymmetric behaviour around the minimum. Alternatively, the possibility of parametrizing the points below and above the minimum with two separate parabolas was investigated. As the red and cyan curves in Figure 10.17 show, a good agreement with the profiled points is observed using this parametrization. This is a nice feature, since as long as the half-parabolic behaviour holds, the left and right confidence intervals can be computed as multiples of the "1σ interval" as for a normal Gaussian. The likelihood profile obtained in two-dimensions is shown instead in Figure 10.18, together with third-order two-dimensional polynomial used to approximate it and the confidence interval obtained.

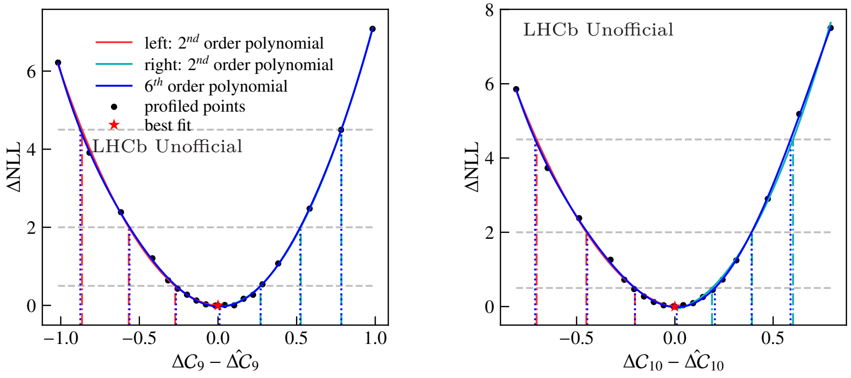


Figure 10.17: One-dimensional profile likelihood in the parameters $\Delta\mathcal{C}_9$ (left) and $\Delta\mathcal{C}_{10}$ (right), centred at the corresponding blinded best fit value (red point). The black points correspond to the points used in the likelihood profile, while the blue line is the fit result obtained with a 6^{th} order polynomial to approximate the likelihood shape and obtain an estimate of the 1, 2, 3σ confidence intervals. The red and cyan lines represent instead a parabolic fit to the profiled points that stay to the left and right of the minimum, respectively.

Table 10.5: Confidence intervals for $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ centred around the blinded best fit value as obtained from a one-dimensional profiling of the likelihood in data.

σ	$\Delta\mathcal{C}_9$		$\Delta\mathcal{C}_{10}$	
	left	right	left	right
1	-0.257	0.262	-0.212	0.194
2	-0.553	0.513	-0.455	0.381
3	-0.866	0.774	-0.721	0.579
avg.	0.273		0.217	

Even though the profiling of the likelihood improves the statistical properties of the fit with respect to the parabolic approximation, the exact coverage of the uncertainties is not ensured and needs to be verified in toys. The coverage of the uncertainties obtained with this procedure is discussed in more detail in Sec. 10.5.

10.5 Statistical properties of the fit

The statistical properties of the fit parameters of interest and the confidence intervals obtained from the likelihood profile of Sec. 10.4 can be studied performing

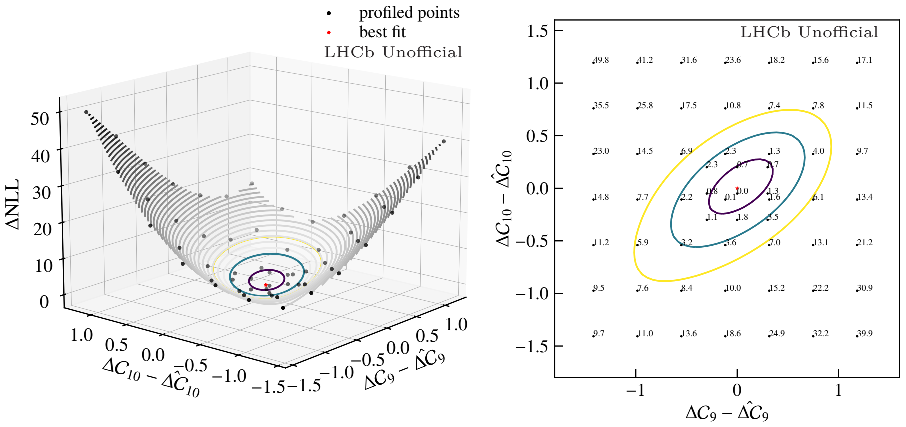


Figure 10.18: Two-dimensional profile likelihood in the parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, centred at the corresponding blinded best fit value (red point). The black points correspond to the points used to profile the likelihood, while the gray lines correspond to a a 3^{rd} order two-dimensional polynomial used to approximate the likelihood shape and obtain an estimate of the 1, 2, 3σ confidence intervals, shown by the indigo, teal and yellow contours. On the left, it is shown a three-dimensional representation of the likelihood, while, on the right, is shown the projection on the $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ plane. Next to each profiled point is reported the value of ΔNLL obtained from the fit.

a set of toy experiments. A series of approximately 500 pseudoexperiments are generated based on the full signal and background models considered in the full data fit, including the separation of different runs and trigger categories (only for electrons). This configuration corresponds exactly to the nominal model used to fit the data. The generation is performed by setting the central values of all fit parameters, *e.g.* Wilson coefficients, yields, background parameters, to the result obtained from the blind fit to data. The only exception are the form factors and CKM parameters, that are set to their corresponding central values from LCSM/lattice-QCD and `UTFitter` predictions (see Sec. 10.4). This different treatment is justified by the fact that these parameters are not floated freely in the final fit, but rather Gaussian constrained to their predicted value. While this caused no complication when performing the fit on data, care must be taken when including these constraints in an ensemble of toy experiments. In particular, the values of the fit parameters used in the generation of the ensembles have to be kept always the same for a given set of pseudoexperiments, while the mean of the external Gaussian constraints applied to a certain subset of parameters has to be sampled based on a normal distribution centred at the corresponding external value and with a standard deviation equal to its uncertainty [174]. As a consequence, each toy will have Gaussian constraints applied with different central values.

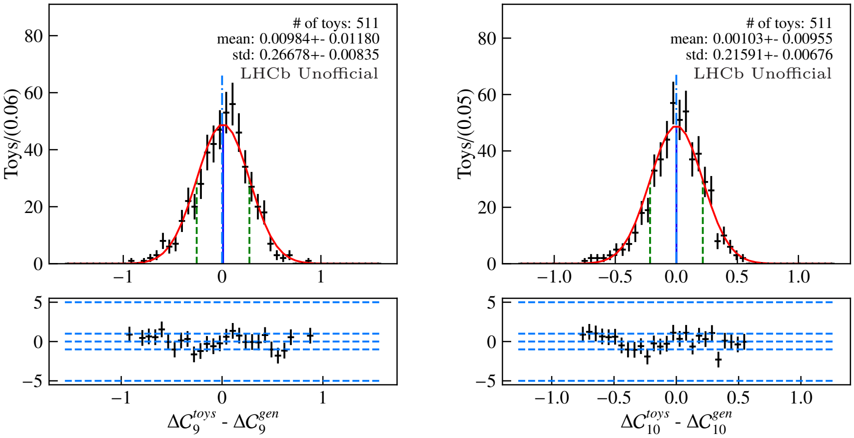


Figure 10.19: One-dimensional fit parameter distributions for $\Delta\mathcal{C}_9$ (left) and $\Delta\mathcal{C}_{10}$ (right) in pseudoexperiments generated from the (blind) fit results. All the distributions are centred at the values used for the generation of the toy ensemble.

Table 10.6: Bias, expected sensitivity and correlation of the difference of Wilson coefficients between muons and electrons.

	Bias	Stat. uncertainty	2D correlation
$\Delta\mathcal{C}_9$	0.010 ± 0.012	0.267 ± 0.008	0.584
$\Delta\mathcal{C}_{10}$	0.001 ± 0.010	0.216 ± 0.007	

After the generation, the set of pseudoexperiments is fitted with the nominal model used in the analysis. Figure 10.19 shows the distributions of the difference of Wilson coefficients as obtained from fits to these ensembles. The width of the obtained distributions is associated to the expected statistical uncertainty, while a bias parameter can be defined as the difference between the expected value of the distribution and the generated value. The expected statistical uncertainty and bias are summarized in Table 10.6. No bias is observed for $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ with the current statistics. Additionally, the statistical uncertainty obtained shows a remarkable agreement with the average 1σ confidence intervals of Table 10.5. Figure 10.20 shows the two-dimensional distributions for the Wilson coefficients of electrons and muons and their difference. As can be seen, there is a correlation between the fit parameters of interest $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, whose value is reported in Table 10.6. When estimating the compatibility with the SM, it is fundamental to take this correlation into account.

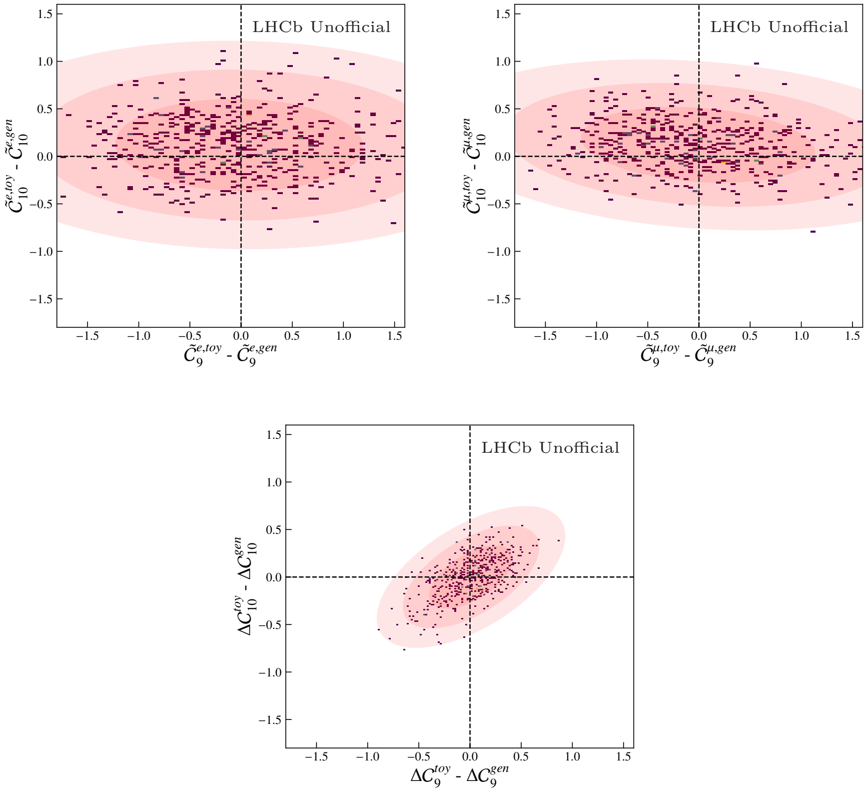


Figure 10.20: Two-dimensional fit parameter distributions for $(\tilde{\mathcal{C}}_9^e, \tilde{\mathcal{C}}_{10}^e)$, $(\tilde{\mathcal{C}}_9^\mu, \tilde{\mathcal{C}}_{10}^\mu)$ and $(\Delta\mathcal{C}_9, \Delta\mathcal{C}_{10})$ for pseudoexperiments generated from the (blind) fit results on the top, middle and bottom row, respectively. All the distributions are centred at the values used for the generation of the toy ensemble.

Toy experiments also allow to study the statistical *coverage* of the uncertainties estimated profiling the likelihood. The *coverage* corresponds to percentage of toys in which the true value of a parameter of interest θ is contained in the quoted confidence interval $[\hat{\theta} - \delta\theta^-,\ \hat{\theta} + \delta\theta^+]$, where $\hat{\theta}$ and $\delta\theta^\pm$ correspond to the best fit value and estimated asymmetric uncertainty in each toy experiment. Intervals for which the probability of containing the true value are compatible with the standard 68.27% are said to *cover*, and directly correspond to a confidence interval for the true value of the parameter of interest.

The coverage is studied for an ensemble of toy experiments, by refitting each of the pseudoexperiments with an alternative model where the values of $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ ($\Delta\mathcal{C}_9$ or $\Delta\mathcal{C}_{10}$ for a 1D profile) are fixed to their generation values. The difference between the negative log-likelihood of the alternative and standard fit configuration corresponds to the profile functions defined in Eq. 10.39 (10.38), computed at the generated values of the parameters of interest, $\Delta\text{NLL}^{gen} \equiv \Delta\text{NLL}(\Delta\mathcal{C}_9{}^{gen}, \Delta\mathcal{C}_{10}{}^{gen})$ $(\Delta\text{NLL}^{gen} \equiv \Delta\text{NLL}(\Delta\mathcal{C}_i^{gen}))$. If the uncertainties estimated using the profiling of the likelihood are covered, the distributions of $2\Delta\text{NLL}^{gen}$ should follow a χ^2 distribution with a number of degrees of freedom equal to the number of parameters that have been fixed to perform the profiling: the number of degrees of freedom is 2 and 1 for the 2D and 1D likelihood profiles, respectively. Figures 10.21 and 10.22 show the distribution of $2\Delta\text{NLL}^{gen}$ for the 2D and the 1D profiling of the difference of Wilson coefficients. The grey dashed lines correspond the 1-2-3 standard deviation intervals and the number reported next to them are the fraction of toys, in percentage, that fall below that value of χ^2. Table 10.7 reports these percentages for the three likelihood profiles performed and compares them with the expected percentages in case of perfect coverage. The coverage observed is compatible with the expected reference values, demonstrating the validity of the use of the likelihood scan to estimate the uncertainty of the measurement.

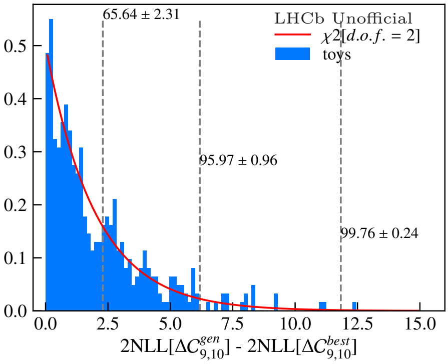


Figure 10.21: Distribution of $2\Delta\text{NLL}^{gen}$ for the 2D-likelihood profile in $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, shown in blue. Superimposed, in red, the value of the χ^2 distribution with two degrees of freedom integrated in the bin. The grey dashed lines correspond the the 1-2-3 standard deviation intervals defined in Table 10.4.

Table 10.7: Coverage of the 1, 2, 3 σ confidence intervals determined from pseudoexperiments.

Interval	Coverage			Reference
	$\Delta\mathcal{C}_9$	$\Delta\mathcal{C}_{10}$	$(\Delta\mathcal{C}_9,\ \Delta\mathcal{C}_{10})$	
1 σ	67.74 ± 2.53	63.37 ± 2.60	65.64 ± 2.31	68.27%
2 σ	96.19 ± 1.04	92.15 ± 1.45	95.97 ± 0.96	95.45%
3 σ	99.71 ± 0.29	99.71 ± 0.29	99.76 ± 0.24	99.73%

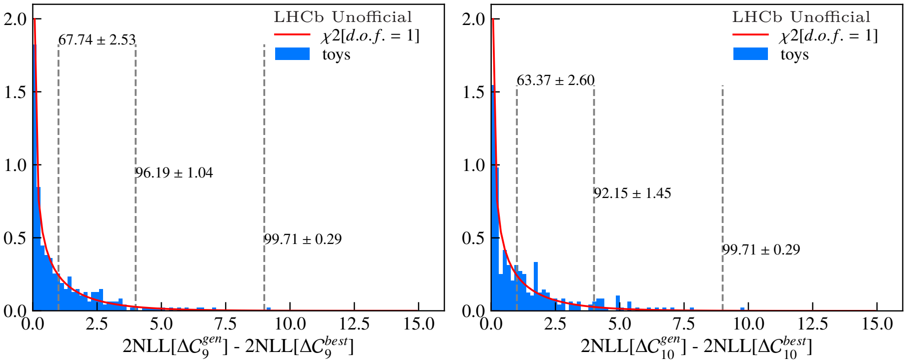


Figure 10.22: Distribution of $2\Delta\mathrm{NLL}^{gen}$ for the 1D-likelihood profile in $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, shown in blue. Superimposed, in red, the value of the χ^2 distribution with one degree of freedom integrated in the bin. The grey dashed lines correspond the the 1-2-3 standard deviation intervals.

10.6 Systematics

The uncertainty obtained from the likelihood profile described in Sec. 10.4 represents only part of the uncertainty budget of the final measurement. The several assumptions made during the analysis described can have a sizable impact on the measurement, and therefore need to be assessed properly and included in the final uncertainty. These contributions, usually referred to as *systematic uncertainties*, are estimated with simulated toys experiments, similarly to the ones described in Sec. 10.5. Currently, the estimation of the systematic uncertainties of the measurement is still ongoing. In the following, a discussion of the contributions that are expected to dominate the systematic uncertainty budget is presented, and an heuristic estimation of the total uncertainty anticipated is provided.

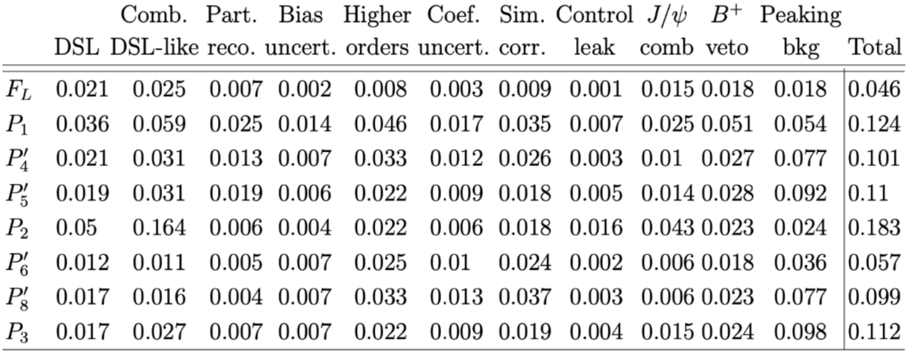
The measurement proposed in this work follows quite closely the strategies used in the ongoing measurements of the R_{K^*} branching fraction ratio [128] and angular analysis of $B^0 \to K^{*0}e^+e^-$ decays [140], with which it shares a sizable fraction of the samples used for data and simulation. It is therefore reasonable to expect that the most important sources of systematic uncertainties observed in these two analyses will also constitute the most relevant part of the systematic error budget of this measurement. Tables 10.8 and 10.9 report a preliminary estimation of the systematic uncertainties for R_{K^*} and the angular observables of $B^0 \to K^{*0}e^+e^-$ decays, respectively.

As it can be seen, the dominant contribution to the uncertainty comes, in both cases, from the contribution of misidentified backgrounds that peak below the signal region (“Misidentified backgrounds” and “Peaking bkg”). At first, the impact of neglecting these contributions in the fits has been studied by both analysis repeating the fits to data for a set of different electron PID

Table 10.8: Individual sources of systematic uncertainties on the R_{K^*} measurements in the central-q^2 (R_{K^*}) region [128]. All values are given in percent relative to the central value. These values are indicative and are computed as weighted averages of systematic variations determined in each data-taking period and trigger category.

Source	central-q^2 R_{K^*}
Form factors	0.76
q^2 smearing	0.31
Particle identification	0.12
Kinematics and multiplicity	0.52
Trigger	0.13
Stability of $r^{K}_{J/\psi}$ and $r^{K^*}_{J/\psi}$	0.47
J/ψ fit model	0.40
Fixed fit parameters	0.16
Combinatorial shape	0.38
Specific backgrounds	0.51
Misidentified backgrounds	2.29
m_{corr} mass	0.33
Total	2.52

Table 10.9: Individual sources of systematic uncertainties on the measurement of the angular observables of $B^0 \to K^{*0}e^+e^-$ decays [140].



	DSL	Comb. DSL-like	Part. reco.	Bias uncert.	Higher orders	Coef. uncert.	Sim. corr.	Control leak	J/ψ comb	B^+ veto	Peaking bkg	Total
F_L	0.021	0.025	0.007	0.002	0.008	0.003	0.009	0.001	0.015	0.018	0.018	0.046
P_1	0.036	0.059	0.025	0.014	0.046	0.017	0.035	0.007	0.025	0.051	0.054	0.124
P_4'	0.021	0.031	0.013	0.007	0.033	0.012	0.026	0.003	0.01	0.027	0.077	0.101
P_5'	0.019	0.031	0.019	0.006	0.022	0.009	0.018	0.005	0.014	0.028	0.092	0.11
P_2	0.05	0.164	0.006	0.004	0.022	0.006	0.018	0.016	0.043	0.023	0.024	0.183
P_6'	0.012	0.011	0.005	0.007	0.025	0.01	0.024	0.002	0.006	0.018	0.036	0.057
P_8'	0.017	0.016	0.004	0.007	0.033	0.013	0.037	0.003	0.006	0.023	0.077	0.099
P_3	0.017	0.027	0.007	0.007	0.022	0.009	0.019	0.004	0.015	0.024	0.098	0.112

requirements. Since tightening the selections should reduce the size of the misidentified backgrounds in data, a saturation effect is expected in the values of the observables of interest obtained. With respect to the nominal PID cut used, which is common between the analyses considered here, a variation roughly equivalent to the statistical uncertainty was found before reaching saturation. Since this behaviour was noticed in both measurements, it is expected that a PID scan of this type would show similar deviations also in this work. While the angular analysis of $B^0 \to K^{*0}e^+e^-$ decays has currently estimated a conservative systematic uncertainty based on the maximum deviation observed, in the measurement of R_{K^*} it was chosen to include these misidentified backgrounds in the final fit using a data-driven estimation of their contribution. The statistical uncertainty associated with this procedure was found to be roughly a third of the statistical uncertainty. A study similar to the one performed in Ref. [128] is currently not available in this analysis, however, a systematic of similar relative size is expected if the approach used can be extended to the multidimensional fit of Chapter 10. The corresponding systematic contribution would then be of the order of 0.09 and 0.07 for $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, respectively.

The choice of the model used in the parametrization of the specific backgrounds considered in the final fit, *i.e.* double-semileptonic and partially reconstructed contributions, is also an important source of systematic uncertainty (see "Specific backgrounds"[3] and "DSL", "Comb DSL-like" and "Part. reco."). As already discussed in Sec. 10.2.2, the nominal parametrization of the double-semileptonic component is obtained from a direct fit to a control region of the LFV sample. In this configuration, the selected sample is expected to be a good description of the mixture of DSL components that are present in the $B^0 \to K^{*0}e^+e^-$ candidates in data. This parametrization is preferred to the use of a mixture of simulated samples due to its data-driven nature. The systematic uncertainty associated with this choice is obtained by performing 300 pseudoexperiments, generated in the standard configuration presented in Sec. 10.2.2, and fitted, firstly, with the matching DSL model and, afterwards, with the model extracted from simulation, as discussed in Appendix F.1.3. A similar procedure has been used to obtain the size of the systematic uncertainty associated with the choice of the partially reconstructed background. A set of 300 pseudoexperiments has been generated in the standard configuration and fitted firstly with the matching partially reconstructed model and afterwards with the model extracted from $B^+ \to K_1^+ \to (K^+\pi^+\pi^-)e^+e^-$ simulation, as discussed in Appendix F.2. The mean and the standard deviation of the toy-by-toy difference associated with these two contributions is summarized in Table 10.10. The systematic uncertainty due to the parametrization of the DSL (partially reconstructed) background is obtained as the sum in quadrature of the mean and standard deviation of the difference observed in toys, and corresponds to 0.075 and 0.056 (0.019 and 0.027) for $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, respectively.

Only other three contributions are expected to have a size comparable to

---

[3] Only the partially reconstructed contribution is considered in Ref. [128], since the DSL component is vetoed.

Table 10.10: Mean and standard deviation of the toy-by-toy difference in $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ for the preliminary systematic studies in the modelling of DSL and partially backgrounds reconstructed.

	DSL		Part. Reco.	
	$<\Delta>$	$\sigma(\Delta)$	$<\Delta>$	$\sigma(\Delta)$
$\Delta\mathcal{C}_9$	0.052 ± 0.004	0.054 ± 0.003	-0.012 ± 0.001	0.015 ± 0.001
$\Delta\mathcal{C}_{10}$	-0.027± 0.004	0.050 ± 0.003	-0.021 ± 0.001	0.017 ± 0.001

the previous: the systematic due to the model used in the generation of the simulated samples ("Form factors"), the systematic linked to the correction of the simulation ("Kinematics and multiplicity", "Stability of $r_{J/\psi}^{K^*}$" and "Sim. Corr."), and the systematic due to the choice of the order of the Chebyshev polynomials used to parametrize the acceptance ("Higher orders"). Assuming for all three contributions uncertainties similar to ones observed for the DSL (partially reconstructed) contribution, the overall systematic uncertainty would correspond to roughly 0.18 and 0.14 (0.12 and 0.1) for both $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, respectively. When these uncertainties are combined in quadrature with the statistical uncertainties reported in Table 10.6, the overall reduction in sensitivity is between 10 and 20% with respect to the statistical uncertainty alone.

Chapter 11

Conclusions and future prospects

The work presented in this dissertation has described the strategy used for the first direct determination of the observables that encode potential non-equal couplings between muons and electrons, obtained from an amplitude analysis of the decays of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$.

The dataset analyzed corresponds to $9\,\mathrm{fb}^{-1}$ collected by the LHCb experiment between the years 2011 and 2018. In the region of dilepton mass square between 1.1 and $7\,\mathrm{GeV}/c^2$, it was possible to isolate roughly 3400 and 1200 candidates of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays. A simultaneous amplitude analysis of these candidates has been performed by means of an unbinned extended maximum likelihood fit, where all the most important experimental effects have been considered, *e.g.* different background contributions and efficiency effects, and the observed yield in the signal modes was included as an additional constraint on the observables $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$.

The central value of these two observables is currently blinded, since a complete estimation of all the most relevant systematic uncertainties is still ongoing. The most important result of this work corresponds therefore to the estimation of the sensitivity of the measurement, which is obtained directly from a likelihood profile on data. This estimation has been confirmed by realistic toys, where the coverage of the confidence intervals obtained from likelihood scans was verified. Additionally, an heuristic estimation based on the preliminary work of Refs. [128, 140] has shown that the systematic uncertainty is expected to be between 50 and 70% of the current statistical uncertainty, suggesting that at the end of RUN 2 this measurement will still be statistically limited. This analysis naturally includes all the information that can be extracted from the measurements of the ratio of branching fractions and difference of angular observables in $B^0 \to K^{*0}\ell^+\ell^-$ decays and, thanks to its unbinned description of the dilepton invariant mass squared, is expected to have better sensitivity to the observables $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$. A comparison of the expected sensitivity on the difference of Wilson coefficients for the binned measurements of R_{K^*} and Q_5 at the end of RUN 2 is shown in Figure 11.1, together with their combination. In red is shown the statistical sensitivity obtained from this analysis. As it can be seen, the 1σ contour is well-within the confidence interval spanned by the combination of the two binned observables, providing the most sensitive measurement of $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ derived from $B^0 \to K^{*0}\ell^+\ell^-$ decays.

During the last years, the LHCb spectrometer has undergone an upgrade, which allowed it to start its RUN 3 data-taking in the summer of 2022. This is just the first step of the LHCb Upgrade program, which is expected to run for more than 10 years and collect a minimum of $300\,\mathrm{fb}^{-1}$ at the end of its life. Ignoring the advantages introduced by the implementation of a fully software based trigger

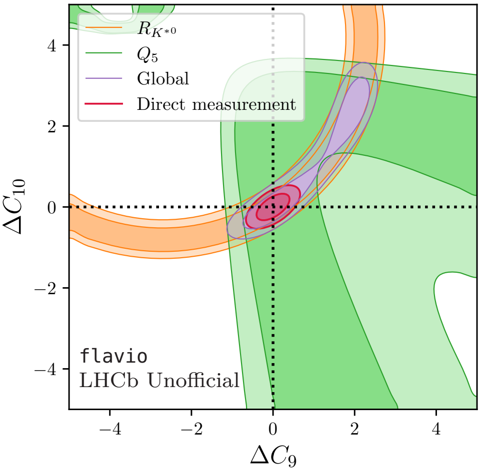


Figure 11.1: The confidence interval obtained for $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ in this work is compared to the expected combination of the binned LFU observables R_{K^*} and Q_5 for the same expected statistics. The separate contribution of the two observables is also shown. The confidence intervals for these two are obtained with the software package `flavio` [175] assuming a SM scenario. The contours shown correspond to the 1σ and 2σ confidence intervals.

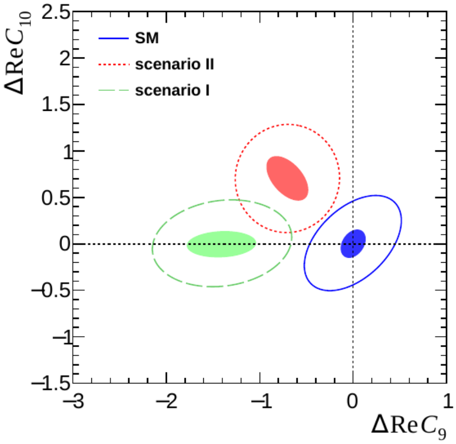


Figure 11.2: Constraints on the $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ from electron and muon modes with the Run 3 and Upgrade II datasets. The 3σ regions for the Run 3 data sample are shown for the SM (solid blue), a vector-axial-vector new physics contribution corresponding to $\Delta\mathcal{C}_9$ = -$\Delta\mathcal{C}_{10}$ = -0.7 (red dotted) and for a purely vector new physics contribution corresponding to $\Delta\mathcal{C}_9$ = -1.4 (green dashed). The shaded regions denote the corresponding constraints for the Upgrade II dataset. More details are provided in Ref. [176]

and potential improvements of the electromagnetic calorimeter, a minimum of 20000 $B^0 \to K^{*0}e^+e^-$ decay events are expected to be collected during the LHCb Upgrade [176]. This endeavour will allow to reduce the statistical uncertainty on the difference in Wilson coefficients of at least a factor four. Such a sensitivity would allow to use these decays, not only to establish the existence of NP if present, but also to discriminate between different NP models, as shown in Figure 11.2. The systematic uncertainties affecting the measurement, that if unchanged would be larger than the statistical uncertainty, are also expected to be similarly reduced with the additional statistics. This measurement is therefore expected to be statistically limited even at the end of the LHCb Upgrade.

Chapter 12

Afterword

After the submission of this work for review, the LHCb collaboration published an update on the R_K and R_{K^*} measurements, corresponding to the completion of the work mentioned in Sec. 10.6. These results, obtained in two regions of q^2 corresponding to $q^2 \in [0.1, 1.1]\,\text{GeV}^2/c^4$ (low-q^2) and $q^2 \in [1.1, 6]\,\text{GeV}^2/c^4$ (central-q^2), supersede the previous work presented in Refs. [52, 60] and correspond to the most precise determination of these quantities up to date. These measurements, reported in Figure 12.1, show a remarkable compatibility with the SM prediction. The difference between the central values published in Refs. [52, 60] is explained by the presence in the final fit of misidentified background events that were considered negligible in previous iterations of the analyses, but that actually resulted to be sizable after a data-driven determination of their contribution was carried out [177].

Even if the hints of lepton flavour non-universality in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays has disappeared, it is important to remember that the deviations observed in the measurement of branching ratios and angular observables of $b \to s\,\ell^+\ell^-$ transitions still hold and are not affected in the same way by the contribution of misidentified backgrounds. For this reason additional work is needed to understand if the tensions observed are of NP origin or a simple underestimation of the systematic uncertainty in the corresponding theory predictions. Additionally, these updated results do not reduce the importance of the work presented in this thesis, but if anything, they increase it. From one side, the improvement in the sensitivity for the couplings $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ will help

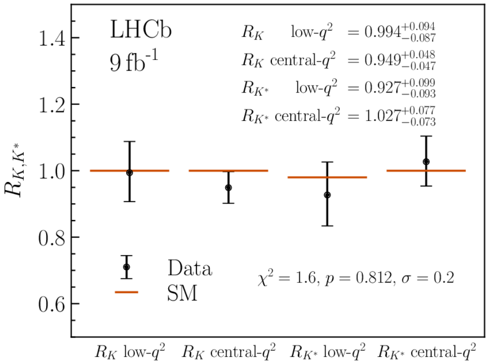


Figure 12.1: Measured values of LFU observables in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays and their overall compatibility with the SM [177].

to further constrain the phase space available for any NP model to come. On the other side, it is fundamental to test lepton flavour universality in different aspects of the same decays, not only studying branching ratios but also their angular distributions. These two aspects are not only sensitive to different NP scenarios but also to different systematic uncertainties. The verification that a coherent picture emerges from these two aspects represents a fundamental check of the measurement performed. The analysis presented here is, at the moment, the only one able provide such a check.

Appendices

Appendix A

P-wave form factors in the narrow-width approximation

To express the $\overline{B}^0 \to K\pi$ matrix elements in terms of the $\overline{B}^0 \to \overline{K}^{*0}$ form factors it is common to assume that the $\overline{K}^{*0}$ decays on-shell. Under this approximation, often referred to as *narrow-width approximation* (NWA), the form factors can be accessed at low q^2 using quark-hadron-duality arguments through QCD Light-Cone Sum Rules [178, 179] and at high q^2 from first principles using Lattice QCD [35, 180]. Due to the complementarity of the two approaches, it is possible to perform a combined fit at the predictions obtained in these two regimes and obtain an effective parametrization that provides a consistent set of form factors over the full q^2 range [34]. It is therefore possible to re-express the form factors introduced in Eq. 2.16 as:

$$\begin{aligned}
\mathcal{F}_\perp(q^2) &= \frac{\sqrt{2\lambda}}{M_B(M_B+M_{K^{*0}})}V(q^2)\,,\\
\mathcal{F}_\parallel(q^2) &= \frac{\sqrt{2}(M_B+M_{K^{*0}})}{M_B}A_1(q^2)\,,\\
\mathcal{F}_0(q^2) &= 8\frac{M_{K^{*0}}}{\sqrt{2}}A_{12}(q^2)\,,\\
\mathcal{F}^T_\perp(q^2) &= \frac{\sqrt{2\lambda}}{M_B^2}T_1(q^2)\,,\\
\mathcal{F}^T_\parallel(q^2) &= \frac{\sqrt{2}(M_B^2-M_{K^{*0}}^2)}{M_B^2}T_2(q^2)\,,\\
\mathcal{F}^T_0(q^2) &= \frac{4M_{K^{*0}}\sqrt{q^2}}{(M_B+M_{K^{*0}})M_B}T_{23}(q^2)\,,\\
\mathcal{F}_t(q^2) &= \frac{\sqrt{\lambda}}{\sqrt{q^2}M_B}A_0(q^2)\,.
\end{aligned} \tag{A.1}$$

A detailed description of the q^2 functions V, $A_{0,1,12}$, $T_{1,2,23}$ is provided in Appendix A.1. The only difference between the presented choice of form factors and the basis commonly used in the literature (see Refs. [136, 181]) corresponds to the definition of A_{12} and T_{23}:

$$A_{12} = \frac{(M_B+M_{K^{*0}})^2((M_B^2-M_{K^{*0}}-q^2)A_1-\lambda A_2)}{16M_BM_{K^{*0}}^2(M_B+M_{K^{*0}})} \tag{A.2}$$

$$= \frac{q^2/2(M_B^2+3M_{K^{*0}}^2-q^2)A_1+\lambda M_{K^{*0}}/(M_B+M_{K^{*0}})A_3}{8M_BM_{K^{*0}}^2(M_B-M_{K^{*0}})} \tag{A.3}$$

$$T_{23} = \frac{(M_B^2 - M_{K^{*0}}^2)(M_B^2 + 3M_{K^{*0}}^2 - q^2)T_2 - \lambda T_3}{8M_B M_{K^{*0}}^2 (M_B - M_{K^{*0}})} \,. \tag{A.4}$$

A.1 Form factors from LCRS and LQCD fits

A series of fits is performed on sets of pseudo-data generated with correlated theoretical uncertainties corresponding to the values of the standard form factors at three q^2 values both at low and high q^2, using the parametrizations provided by LCSR [34] and lattice QCD [35, 182] respectively. The fits are performed using a simplified series expansion (SSE), based on the rapidly converging series in the parameters

$$z(t) \equiv \frac{\sqrt{t_+ - t} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - t} + \sqrt{t_+ - t_0}} \,, \tag{A.5}$$

where $t_\pm = (M_B \pm M_{K^{*0}})$ and $t_0 = t_+(1 - \sqrt{1 - \frac{t_-}{t_+}})$. It is then possible to express $F_i \in \{V, A_{0,1,12}, T_{1,2,23}\}$ as

$$F_i(q^2) = P_i(q^2) \sum_k \alpha_k^i [z(q^2) - z(0)]^k \,, \tag{A.6}$$

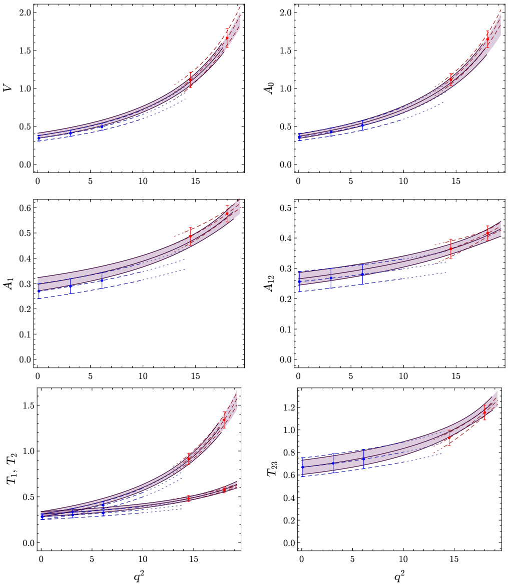


Figure A.1: Combined LCSR and lattice fit to $V, A_{0,1,12}, T_{1,2,23}$, where lattice data points are indicated in red, LCSR points in blue, the gray solid band shows the combined three-parameter fit and the red dashed band the two-parameter lattice fit from Ref. [182]. Plots reproduced from Ref [34].

where $P_i(q^2) = (1 - q^2/M^2_{K^{*0}})^{-1}$ is a simple pole corresponding to the first resonance in the spectrum. For this application the expansion is truncated after the quadratic term in z, corresponding for a maximum of three fit parameters $\alpha^i_{0,1,2}$ for each of the seven F_i. The fit results, as obtained in Ref. [34], are summarized Fig. A.1, while Table A.2 reports the central values of the expansion coefficients obtained.

	$B \to K^*$
$a_0^{A_0}$	0.37 ± 0.03
$a_1^{A_0}$	-1.37 ± 0.26
$a_2^{A_0}$	0.13 ± 1.63
$a_0^{A_1}$	0.30 ± 0.03
$a_1^{A_1}$	0.39 ± 0.19
$a_2^{A_1}$	1.19 ± 1.03
$a_0^{A_{12}}$	0.27 ± 0.02
$a_1^{A_{12}}$	0.53 ± 0.13
$a_2^{A_{12}}$	0.48 ± 0.66
a_0^{V}	0.38 ± 0.03
a_1^{V}	-1.17 ± 0.26
a_2^{V}	2.42 ± 1.53
$a_0^{T_1}$	0.31 ± 0.03
$a_1^{T_1}$	-1.01 ± 0.19
$a_2^{T_1}$	1.53 ± 1.64
$a_0^{T_2}$	0.31 ± 0.03
$a_1^{T_2}$	0.50 ± 0.17
$a_2^{T_2}$	1.61 ± 0.80
$a_0^{T_{23}}$	0.67 ± 0.06
$a_1^{T_{23}}$	1.32 ± 0.22
$a_2^{T_{23}}$	3.82 ± 2.20

Figure A.2: Fit results for the SSE expansion coefficients in the combined LCSR and lattice fit as obtained in Ref. [34].

Appendix B

Long distance effects from analyticity

Non-local matrix elements arising from the calculation of the time-ordered product between four-quark operators, *i.e.* $\mathcal{O}_{1,2}$, and electromagnetic currents are particularly difficult to determine reliably from first principles and have been the focus of much attention over the last two decades. Different approaches have been suggested to extract these contributions from data-driven analysis either using empirical parametrization [122, 183, 184] or exploiting the analytical properties of its structure [23, 124, 125, 185]. In the following, we will rely on the parametrization presented in Ref. [23] and briefly summarize here.

The non-local hadronic functions $\mathcal{H}_\lambda(q^2)$ can be expressed in terms of a "conformal" variable z [186, 187]

$$z(q^2) \equiv \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}\,, \quad t_+ = (2M_D)^2\,, \quad t_0 = t_+ - \sqrt{t_+(t_+ - M^2_{\psi(2S)})}\,, \tag{B.1}$$

where

- $2M_D$ corresponds to the threshold at which is possible to produce a two-particle intermediate states (branch cut) through $B^0 \to K^{*0}[\bar{D}D](\to \ell^+\ell^-)$ decays,
- $M_{\psi(2S)}$ is the mass of the $\psi(2S)$ resonance that, together with the J/ψ, contributes to the decay as a one-particle intermediate state (pole).

This choice enables to map the $c\bar{c}$ branch cut in the unit circumference $|z| = 1$ and the region of q^2 between -7 GeV^2 and $M^2_{\psi(2S)}$ in $|z| < 0.52$. As a consequence, the functions $\mathcal{H}_\lambda(z)$ are analytical for $|z| < 1$ except for the two simple poles for $z_{J/\psi(1S)} \simeq 0.18$ and $z_{\psi(2S)} \simeq -0.44$, corresponding to the resonances J/ψ and $\psi(2S)$ respectively. Therefore, after dividing out these singularities

$$\mathcal{H}_\lambda(z) = \frac{1 - z z^*_{J/\psi(1S)}}{z - z_{J/\psi(1S)}} \cdot \frac{1 - z z^*_{\psi(2S)}}{z - z_{\psi(2S)}} \cdot \hat{\mathcal{H}}_\lambda(z), \tag{B.2}$$

the remaining functions $\hat{\mathcal{H}}_\lambda(z)$ are analytical in $|z| < 1$ and can be Taylor-expanded around $z = 0$. In addition, in order to improve the convergence of the expansion, these non-local functions are assumed to inherit all the singularities of the form factors, resulting in

$$\hat{\mathcal{H}}_\lambda(z) = \Big[\sum_{k=0}^{K} \alpha_k^{(\lambda)} z^k\Big] \mathcal{F}_\lambda(z)\,, \tag{B.3}$$

where $\alpha_k^{(\lambda)}$ are complex coefficients, and the analytical expansion is truncated after the term z^k. This parametrization is expected to converge reasonably well in the region of interest, $|z| < 0.52$,[1] however, the truncation of the series unavoidably introduces some model dependence.

Finally, using the property that the longitudinal correlator must vanish at zero momentum transfer (see Ref. [23]), *i.e.* $\hat{\mathcal{H}}_0(q^2 = 0) = 0$, we can reduce by one the number of complex parameters through the replacement

$$\frac{\hat{\mathcal{H}}_0(z)}{\mathcal{F}_0(z)} = (z - z_0) \cdot \sum_{k=0}^{K-1} \alpha_k^0 z^k \,, \tag{B.4}$$

where $z_0 = z(q^2 = 0)$.

---

[1] The domain $|z| < 0.52$ includes q^2 values between -7 GeV^2 and $M^2_{\psi(2S)}$.

Appendix C

Amplitude fits to $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ data candidates including partially reconstructed backgrounds

In the following, the extension of the mass fits obtained in Sec. 8.3.2 with the inclusion of the variables $\cos\theta_K$, $\cos\theta_\ell$, ϕ and $m(K\pi)$ is presented. The signal and background components have been parametrized using the same datasets discussed in Sec. 8.3.2 and keeping the same mass parametrization already described. The list below reports the parametrization of the components in the remaining variables:

- $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ signal decays:
  The signal component is parametrized as reported in Eq. 9.6, with an effective acceptance ϵ_{eff} parametrized accounting for the specific set of selections applied. The floating parameters for the signal component correspond to the real and complex part of the amplitudes $\tilde{A}_{\perp,\parallel,S0}$, width and centre of Breit-Wigner in $m_{K\pi}$ and the shift and scale of the invariant mass peak. In electron fits the right-hand tail of the mass distribution is also allowed to float via the scaling s_α.

- *Combinatorial*:
  The combinatorial background is modelled using Chebyshev polynomials up to third (second) order for $\cos\theta_K$, up to second order for $\cos\theta_\ell$, up to second (first) order for ϕ and up to first (second) order for $m_{K\pi}$ for muons (electrons), requiring factorization between the different variables. The shape is allowed to float freely in data, together with the slope of its exponential mass distribution and the corresponding yield.

- Partially reconstructed $B \to X J/\psi(\to \ell^+\ell^-)$ decays:
  This contribution is modelled from simulation separately for different trigger categories but joining the different years of simulation available, regardless of the different integrated luminosities as already discussed. The invariant mass $m(K\pi\ell\ell)$ is modelled with a binned KDE while Chebyshev polynomials up to third order for $\cos\theta_K$, up to second order for $\cos\theta_\ell$, up to second order for ϕ are used for muons and electrons. The $m(K\pi)$ mass, due to the contribution from real $\overline{K}^{*0}$, presents a resonant behaviour, parametrized using a combination of a linear function and a Gaussian. Only for electrons, an additional Gaussian contribution is added

to the Chebychev parametrization of $\cos\theta_K$, to improve the description of the peaking behaviour around 0.6. This contribution, visible also in muon modes but ignored there due to its smaller size, corresponds to $B^+ \to K^+X(\to Y\pi J/\psi)$ decays, where the asymmetric behaviour in $\cos\theta_K$ is caused by the different average momentum of the kaon and pion, similarly to what happens in double-semileptonic decays. The shapes obtained are fixed from simulation, with the only parameter allowed to float corresponding to the yield of the partially reconstructed component. Figure C.1 shows the parametrization obtained in the trigger category L0L! for muons and electrons, respectively shown in the left and right column.

The projections of the fits for muons and electrons in the trigger categories L0L! for Run 2 are shown in Figs. C.2 and C.3. Similar results have been obtained for the L0I trigger categories and are not reported here for brevity.

Overall the fit projections show a good agreement with the data with four main exceptions: a large deviation in the region of $\cos\theta_K$ between -1 and 0.5 that is consistent between electrons and muons, a large deviation in the region of $\cos\theta_K$ between 0.85 and 1 in the muons, a consistent deviation in the $m_{K\pi}$ mass in muons and a sizable deviation in the variant mass distribution for both lepton flavours. The first three effects are also present in the simplified fit configuration and are thus explained in Sec. 9.2.3. The distortion observed in the invariant mass is instead different to what is observed in the simplified fit. Since the mass only fit described in Sec. 8.3.2 shows a much better agreement in data, it is expected that the deviations observed are caused by the introduction of the additional distributions, most likely caused by a non-perfect correspondence between the distributions obtained from simulation and the background content present in data. The effect of the background distortion is however not considered such to invalidate the proposed approach.

The comparison between the observables extracted and the reference values listed in Table 9.5 are summarized in Fig. C.4. The values of the observables for electron and muon channels in L0L! (L0I) can also be found explicitly in Table C.1, together with their compatibility with previous measurements and between themselves. To appropriately compare with the available results, we computed the expected values of $|A_S|^2$ and δ_S in four bins of $m_{K\pi}$: [826, 861] MeV/c^2, [861, 896] MeV/c^2, [896, 931] MeV/c^2 and [931, 966] MeV/c^2, referred for brevity as Δ_i with i from 1 to 4. The procedure used follows the definition in Eq. 9.11. The uncertainties reported are statistical only.

In general, the compatibility between P-wave observables for electrons and muons in both trigger categories suggest that the current approach does not introduce any difference between the decays that involve the two lepton generations. A sizable difference is observed instead when considering the S-wave observable δ_S, that is significantly lower in electrons with respect to muons. The origin of this difference is not fully understood, but could be related to a possible mismodelling of the partially reconstructed contribution from simulation. This difference disappears when, in the simplified configuration fit of Sec. 9.2.3, additional selections are applied to reduce as much as possible this contribution.

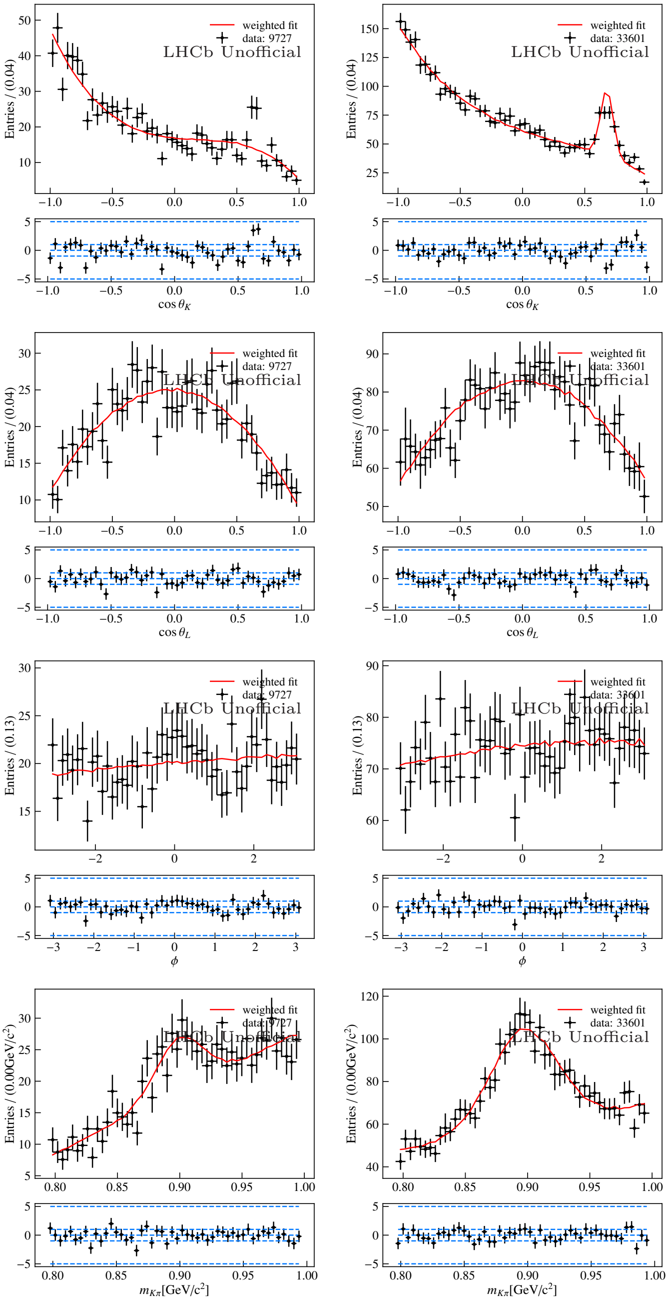


Figure C.1: Parametrization of the partially reconstructed components as obtained from a fit to simulated events for the trigger category L0L! in the muon and electron channel, shown in the left and right column respectively.

When comparing with previous measurements, the P-wave observables show a good agreement with respect to the results obtained with 1 fb^{-1} of data taken in 2011, corresponding with the first columns in Table 9.5. Some tensions can be noticed for S_4 and S_8 when comparing to the check performed in the angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$ up to the year 2016. A stronger disagreement can be seen when comparing S-wave and S-P interference terms. It is however important

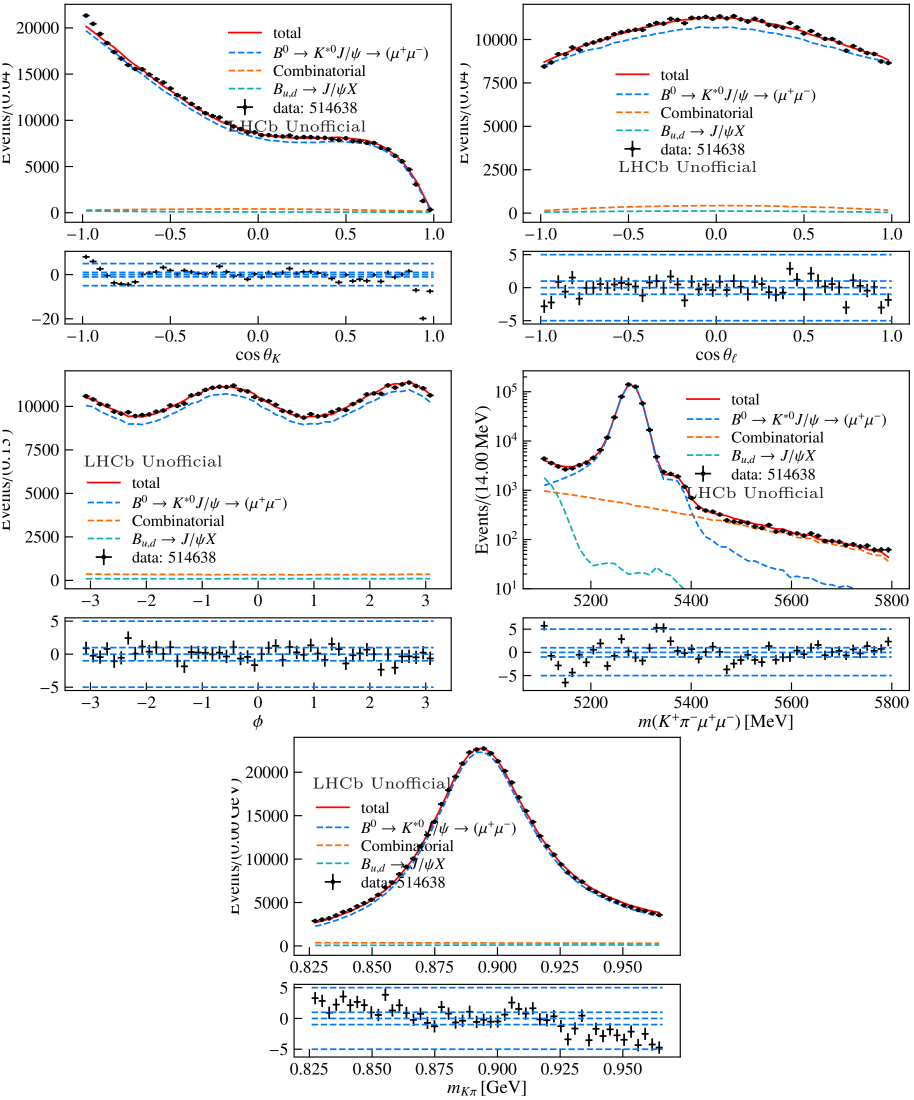


Figure C.2: Projection of the fit result for the $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ angular fit in the L0L! trigger category for Run 2.

to stress that the uncertainties reported here are purely statistical in nature. The compatibility between the different measurements must be re-evaluated once all the systematics are taken into account.

The impact of ignoring the additional components has been briefly studied. The data fits have been repeated including their contributions, with the angular and $m(K\pi)$ distributions fixed from simulation and constraining their expected

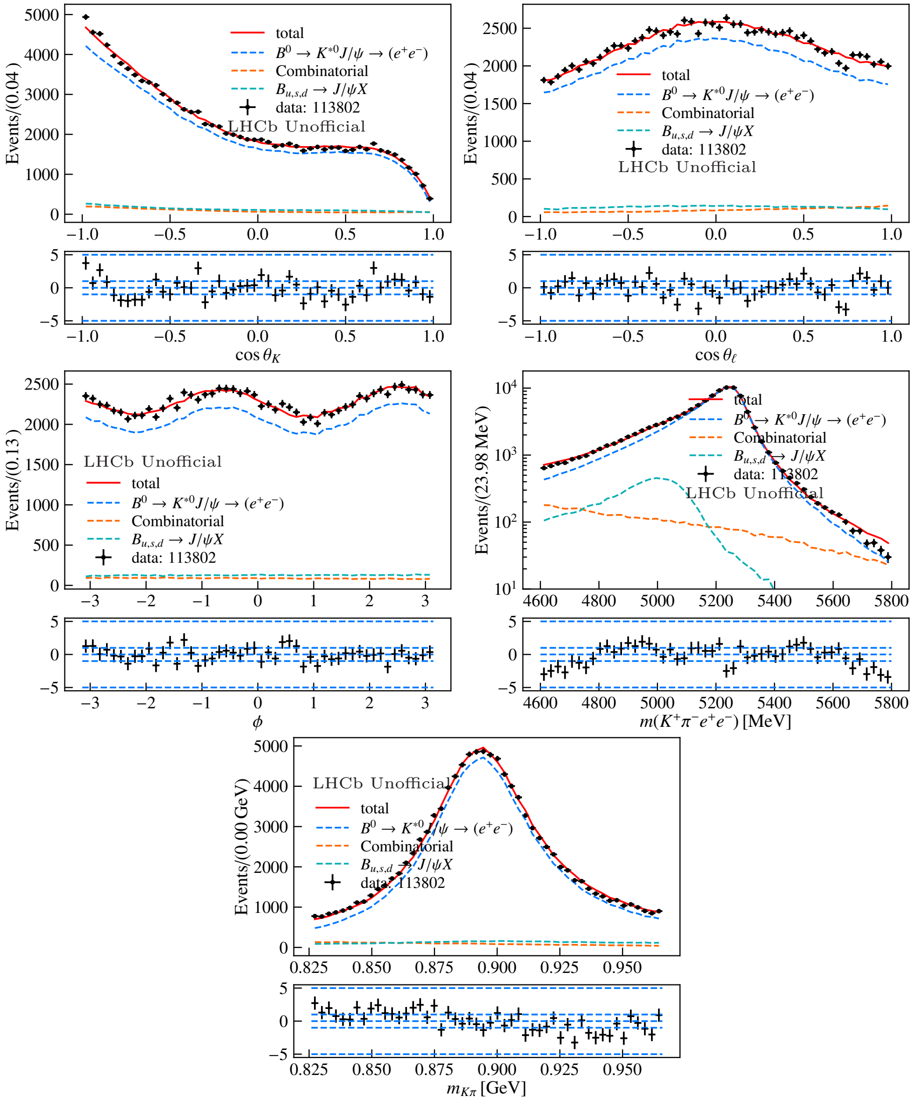


Figure C.3: Projection of the fit result for the $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ angular fit in the L0L! trigger category for Run 2.

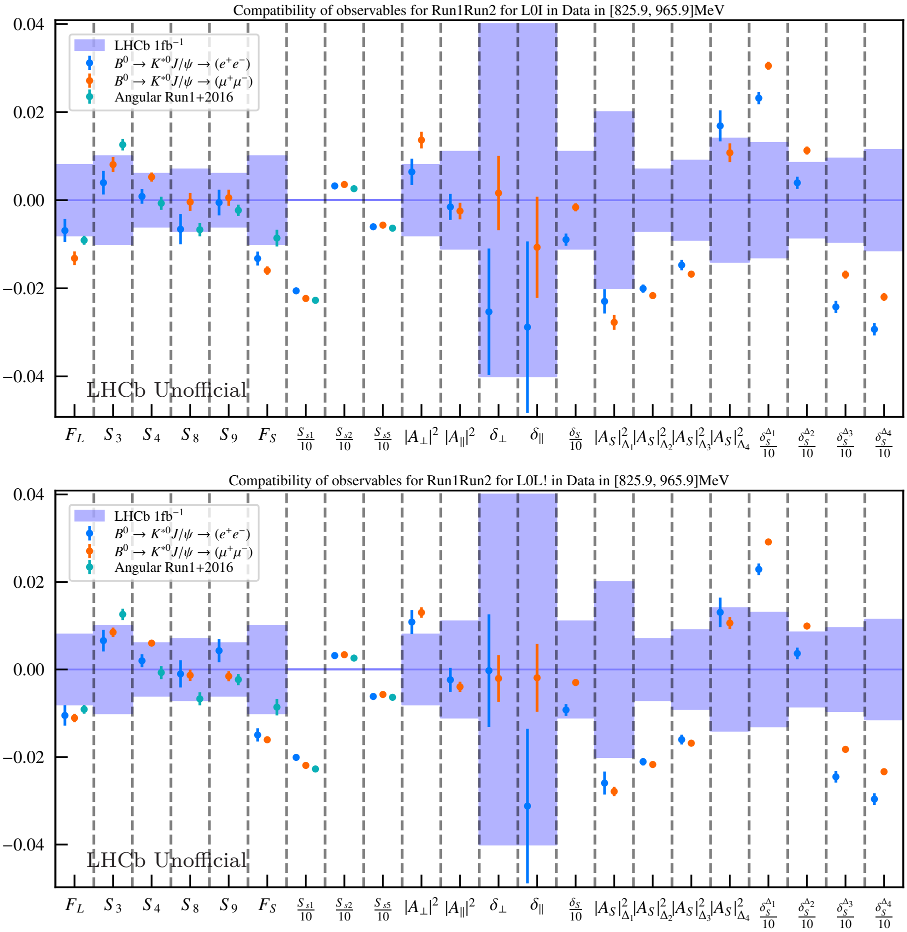


Figure C.4: Summary of amplitude parameters and the derived angular observables for the simultaneous fit in Run 1 and Run 2 of electrons and muons. The fits are performed separately in the trigger categories L0I (top) and L0L! (bottom).

yields analogously to Sec. 8.3.2. Due to their small size, their inclusion has next to no effect on the observables of the $B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$ decays and is not reported here for brevity.

Type	$(\to e^+e^-)$	$\Delta(\sigma)^{e^+e^-}_{\text{LHCb 1fb}^{-1}}$	$\Delta(\sigma)^{e^+e^-}$ LHCb RUN 1 +2016	$(\to \mu^+\mu^-)$	$\Delta(\sigma)^{\mu^+\mu^-}$ LHCb 1fb^{-1}	$\Delta(\sigma)^{\mu^+\mu^-}$ LHCb RUN 1 +2016	LFU
F_L	0.5651 ± 0.0026	-0.82	0.78	0.5588 ± 0.0016	-1.62	-2.20	-2.06
F_S	0.0508 ± 0.0016	-1.31	-1.87	0.0480 ± 0.0010	-1.59	-3.47	-1.48
S_3	-0.0090 ± 0.0027	0.38	-2.90	-0.0049 ± 0.0017	0.80	-2.11	1.30
S_4	-0.2491 ± 0.0017	0.14	0.70	-0.2447 ± 0.0010	0.86	3.27	2.26
S_8	-0.0546 ± 0.0034	-0.85	0.03	-0.0484 ± 0.0020	-0.06	2.49	1.55
S_9	-0.0845 ± 0.0029	-0.08	0.56	-0.0834 ± 0.0018	0.09	1.28	0.32
S_{s1}	-0.2057 ± 0.0058	–	3.37	-0.2231 ± 0.0035	–	0.95	-2.54
S_{s2}	0.0324 ± 0.0019	–	2.47	0.0356 ± 0.0011	–	4.67	1.43
S_{s5}	-0.0603 ± 0.0013	–	1.49	-0.0566 ± 0.0009	–	3.62	2.35
δ_S	2.1206 ± 0.0139	-0.81	–	2.1939 ± 0.0091	-0.15	–	4.42
$\delta_S^{\Delta_1}$	-2.9615 ± 0.0139	-46.29	–	-2.8881 ± 0.0091	-45.87	–	4.42
$\delta_S^{\Delta_2}$	2.6995 ± 0.0139	0.46	–	2.7729 ± 0.0091	1.32	–	4.42
$\delta_S^{\Delta_3}$	1.6978 ± 0.0139	-2.53	–	1.7712 ± 0.0091	-1.77	–	4.42
$\delta_S^{\Delta_4}$	1.2368 ± 0.0139	-2.55	–	1.3102 ± 0.0091	-1.92	–	4.42
$\delta_\perp$	2.9147 ± 0.0144	-0.60	–	2.9416 ± 0.0084	0.04	–	1.62
$\delta_\parallel$	-2.9688 ± 0.0195	-0.65	–	-2.9507 ± 0.0115	-0.26	–	0.80
$\lvert A_S\rvert^2_{\Delta_1}$	0.0920 ± 0.0028	-1.14	–	0.0873 ± 0.0017	-1.38	–	-1.48
$\lvert A_S\rvert^2_{\Delta_2}$	0.0289 ± 0.0009	-2.84	–	0.0273 ± 0.0006	-3.08	–	-1.48
$\lvert A_S\rvert^2_{\Delta_3}$	0.0373 ± 0.0012	-1.62	–	0.0352 ± 0.0007	-1.86	–	-1.48
$\lvert A_S\rvert^2_{\Delta_4}$	0.1219 ± 0.0035	1.17	–	0.1158 ± 0.0021	0.76	–	-1.48
$\lvert A_\perp\rvert^2$	0.2084 ± 0.0030	0.75	–	0.2156 ± 0.0019	1.66	–	2.04
$\lvert A_\parallel\rvert^2$	0.2265 ± 0.0030	-0.13	–	0.2255 ± 0.0019	-0.22	–	-0.27

Type	$(\to e^+e^-)$	$\Delta(\sigma)^{e^+e^-}_{\text{LHCb 1fb}^{-1}}$	$\Delta(\sigma)^{e^+e^-}$ LHCb RUN 1 +2016	$(\to \mu^+\mu^-)$	$\Delta(\sigma)^{\mu^+\mu^-}$ LHCb 1fb^{-1}	$\Delta(\sigma)^{\mu^+\mu^-}$ LHCb RUN 1 +2016	LFU
F_L	0.5615 ± 0.0023	-1.26	-0.55	0.5609 ± 0.0010	-1.37	-1.43	-0.23
F_S	0.0490 ± 0.0015	-1.48	-2.62	0.0479 ± 0.0006	-1.60	-3.73	-0.68
S_3	-0.0064 ± 0.0025	0.64	-2.14	-0.0045 ± 0.0011	0.85	-2.43	0.71
S_4	-0.2480 ± 0.0015	0.32	1.28	-0.2440 ± 0.0007	1.00	4.10	2.48
S_8	-0.0490 ± 0.0031	-0.13	1.65	-0.0493 ± 0.0013	-0.18	2.72	-0.09
S_9	-0.0797 ± 0.0026	0.66	2.24	-0.0855 ± 0.0012	-0.25	0.43	-2.03
S_{s1}	-0.2008 ± 0.0056	–	4.32	-0.2191 ± 0.0023	–	2.35	-3.04
S_{s2}	0.0318 ± 0.0018	–	2.30	0.0338 ± 0.0008	–	4.12	0.99
S_{s5}	-0.0616 ± 0.0013	–	0.91	-0.0570 ± 0.0005	–	3.62	3.32
δ_S	2.1177 ± 0.0134	-0.83	–	2.1803 ± 0.0058	-0.27	–	4.29
$\delta_S^{\Delta_1}$	-2.9644 ± 0.0134	-46.33	–	-2.9017 ± 0.0058	-46.05	–	4.29
$\delta_S^{\Delta_2}$	2.6966 ± 0.0134	0.43	–	2.7593 ± 0.0058	1.17	–	4.29
$\delta_S^{\Delta_3}$	1.6949 ± 0.0134	-2.56	–	1.7576 ± 0.0058	-1.92	–	4.29
$\delta_S^{\Delta_4}$	1.2339 ± 0.0134	-2.58	–	1.2966 ± 0.0058	-2.05	–	4.29
$\delta_\perp$	2.9397 ± 0.0128	-0.01	–	2.9380 ± 0.0053	-0.05	–	-0.13
$\delta_\parallel$	-2.9712 ± 0.0177	-0.71	–	-2.9419 ± 0.0078	-0.05	–	1.52
$\lvert A_S\rvert^2_{\Delta_1}$	0.0891 ± 0.0026	-1.29	–	0.0871 ± 0.0011	-1.39	–	-0.68
$\lvert A_S\rvert^2_{\Delta_2}$	0.0279 ± 0.0009	-2.99	–	0.0273 ± 0.0004	-3.10	–	-0.68
$\lvert A_S\rvert^2_{\Delta_3}$	0.0360 ± 0.0011	-1.77	–	0.0352 ± 0.0005	-1.87	–	-0.68
$\lvert A_S\rvert^2_{\Delta_4}$	0.1181 ± 0.0034	0.91	–	0.1156 ± 0.0014	0.75	–	-0.67
$\lvert A_\perp\rvert^2$	0.2129 ± 0.0027	1.28	–	0.2150 ± 0.0012	1.61	–	0.73
$\lvert A_\parallel\rvert^2$	0.2257 ± 0.0027	-0.21	–	0.2241 ± 0.0012	-0.36	–	-0.54

Table C.1: Overview of the amplitude parameters and the derived angular observables for the simultaneous fits in L0L! (top) and L0I (bottom) RUN 1 and RUN 2 for $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ and $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$, indicated here as $(\to e^+e^-)$ and $(\to \mu^+\mu^-)$ for brevity. Additionally the compatibility, expressed as a number of standard deviations, of the obtained values with respect to the reference reported in the first column of Table 9.5. The last columns shows instead the compatibility between electron and muon channel.

Appendix D

S-wave form factors

The scalar form factors f_+, f_T and f_0 introduced in Sec. 10.1.1 can be described by the following matrix elements,

$$\langle K_0^*(k)|\bar{s}\gamma_\mu\gamma_5 b|B(p)\rangle = \left[(p+k)_\mu - \frac{M_B^2 - M_{K_0^*}^2}{q^2} q_\mu\right] f_+(q^2) + \frac{M_B^2 - M_{K_0^*}^2}{q^2} q_\mu \, f_0(q^2) \tag{D.1}$$

and

$$\langle K_0^*(k)|\bar{s}\sigma_{\mu\nu}\gamma_5 q^\nu b|B(p)\rangle = i\Big[(M_B^2 - M_{K_0^*}^2)q_\mu - q^2(p+k)_\mu\Big]\frac{M_B}{q^2} f_T(q^2)\,. \tag{D.2}$$

They correspond to the form factors reported in Eq.10 of Ref. [166], except for the transformation

$$f_T \mapsto \frac{q^2}{M_B(M_B + M_{K_0^*})} f'_T\,, \tag{D.3}$$

where f'_T is defined in Ref. [166].

The form factors definition holds for a generic scalar-to-scalar transition, but the exact evaluation of the f_+, f_T and f_0 functions depends on the mesons involved in the decay. Since the current theoretical knowledge on the $B^0 \to [K\pi]_{J=0}$ form factors is very limited, the S-wave form factors are fixed in the amplitude fit and a systematic uncertainty will be assigned for this choice in the model.

The nominal model for the S-wave form factors is based on Ref. [188], where a calculation in the perturbative QCD approach and the dipole parametrization has been adopted to access the momentum transfer dependence (see Eq.A3 of [188])

$$F(q^2) = \frac{F(0)}{1 + a_F\, q^2/m_B^2 + b_F\,(q^2/m_B^2)^2}\,. \tag{D.4}$$

The resulting coefficients are summarized in Table D.1.

Table D.1: $B^0 \to K_0^*$ form factors as from Table II of [188].

	$F(0)$	a_F	b_F
f_+	0.27	-2.1	1.2
f_0	0.27	-1.2	0.08
f_T	0.30	-2.2	1.2

Appendix E

Improvements in the inclusion of the $K\pi$ invariant mass lineshape

The expression $\tilde{f}_J$, introduced in Eq. 10.6, has been used in binned angular analysis to describe the shape in q^2 and $m_{K\pi}$ of the P and S-wave amplitudes. As already mentioned, only the term f_{BW}^{892} and f_{LASS} contains information on the dynamics of the decay, while the remaining factors are obtained from conservation of linear and angular momentum, with the Blatt-Weisskopf factor functions that ensure a non-divergent behaviour of the amplitude. However, when considering the expression obtained in Eq. 2.16 and 10.3 for the transversity amplitudes of $B^0 \to K^{*0}\ell^+\ell^-$, it is possible to notice a correspondence between the functions $\mathcal{A}_\lambda^{L,R}$ and $\tilde{f}_{P,S}$, once it is realized that $p = \frac{\sqrt{\lambda}}{2M_B}$. In particular, considering only the p dependence, we have that:

$$\tilde{f}_P \quad \propto \quad \sqrt{p} \propto \mathcal{A}_{\perp,\parallel,0,t}^{L,R}\,, \tag{E.1}$$

$$\tilde{f}_S \quad \propto \quad \sqrt{p^3} \propto \mathcal{A}_{S0,St}^{L,R}\,. \tag{E.2}$$

This exact correspondence suggests that the factors $\sqrt{p}(p/p_0)$ introduced by Eq. 10.6 are actually already present in Eq. 2.16 and 10.3, and correspond therefore to a double counting. On a similar line of thought, it can be noticed that the phenomenological form factor $B'(p, p_0, d)$ simply corresponds to an approximation of the actual $B^0 \to K^{*0}$ form factors and should be therefore also removed.

A different discussion holds for the terms that describe the decay of the $K^*(892)^0$ or $K^*(1430)^0$, together with the corresponding non-resonant S-wave. The expressions obtained for the transversity amplitudes $\mathcal{A}_\lambda^{L,R}$ assume indeed only the transition to a resonant state and no amplitude for its subsequent decay is considered.

Therefore, the parametrization described in Eq. 10.6, could actually be improved with the modification of the $\tilde{f}_P$ and $\tilde{f}_S$ into:

$$\tilde{f}_P(m_{K\pi}) \quad \to \quad \left(\sqrt{k}\cdot\frac{k}{m_{K\pi}}\right)\frac{B'_{L=1}(k, k_{892}, d)}{m_{K\pi}^2 - m_{892}^2 - i m_{892}\Gamma_{892}(m_{K\pi})}\,, \tag{E.3}$$

$$\tilde{f}_S(m_{K\pi}) \quad \to \quad \sqrt{k}\left(\frac{1}{\cot\delta_B - i} + e^{2i\delta_B}\frac{1}{\cot\delta_R - i}\right)\,. \tag{E.4}$$

With this modification the modification of the transversity amplitudes to include the $m_{K\pi}$ contribution would correspond to

$$\begin{aligned}\mathcal{A}_{0,\perp,\parallel,t}^{L,R}(q^2, m_{K\pi}) &\to \mathcal{A}_{0,\perp,\parallel,t}^{L,R}(q^2)\cdot A_P(m_{K\pi})\,,\\ \mathcal{A}_{S0,St}^{L,R}(q^2, m_{K\pi}) &\to \mathcal{A}_{S0,St}^{L,R}(q^2)\cdot A_S(m_{K\pi})\,,\end{aligned} \tag{E.5}$$

where

$$\begin{aligned} A_P(m_{K\pi}) &= \mathcal{N}_P \cdot \tilde{f}_P(m_{K\pi})\,, \\ A_S(m_{K\pi}) &= \mathcal{N}_S \cdot |g_S| e^{i\delta_S} \cdot \tilde{f}_S(m_{K\pi})\,. \end{aligned} \tag{E.6}$$

The constants $\mathcal{N}_P$ and $\mathcal{N}_S$ are here defined requiring that the modulus squared of A_P and A_S are normalized to unity in the entire physical region

$$\int_0^\infty \left|\mathcal{N}_i \cdot f_i(m_{K\pi})\right|^2 \mathrm{d}m_{K\pi} = 1, \tag{E.7}$$

and no dependence on q^2 in the normalization is now present. In this work these improvements have not been considered due to the small impact that such modifications are expected to have, together with the additional needed time that reproducing all fits and toys would have meant.

It is however important to understand that, even with these improvements, the expressions discussed here are still only an effective approximation of the differential decay rate $\frac{\mathrm{d}^5\Gamma}{\mathrm{d}q^2\,\mathrm{d}m_{K\pi}\,\mathrm{d}\vec{\Omega}}$. An attempt to improve on such a parametrization, including the finite width of the $K^*(892)^0$ and its decay, including the full $m_{K\pi}$ spectrum as been proposed in Ref. [189]. However, while this is a clear step in the direction of having a coherent parametrization of $B^0 \to K^{*0}\ell^+\ell^-$ decays that includes its $m_{K\pi}$ shape, to be fully implemented in this context it would ideally require to have a corresponding parametrization in the case of $K^*(1430)^0$ and its non-resonant contributions. Such parametrization is currently not available in the literature.

Appendix F

Backgrounds modelling for $B^0 \to K^{*0}\ell^+\ell^-$

F.1 Double-semileptonic background

F.1.1 A rough comparison of DSL in $K\pi e\mu$ and $K\pi ee$ data

Even though the simulation for $B^0 \to D^-(\to K^{*0}e^-\bar{\nu}_e)e^+\nu_e$ and $B^0 \to D^-(\to K^{*0}\mu^-[e^-]\bar{\nu})e^+[\mu^+]\nu$ points to the fact that the angular shapes are minimally effected by the presence of an electron or a muon in the final state, it is important to check the portability of the parametrization obtained from $K\pi e\mu$ data to $K\pi ee$ data.

It is possible to do so by looking in both datasets at the mass window $m(K\pi ee) \in [3800, 4300]\,\mathrm{MeV}/c^2$ after a tight MVA cut (MVA > 0.99). Here we expect:

- only double-semileptonic cascade decays and combinatorial backgrounds for $K\pi e\mu$ data,
- double-semileptonic cascade decays, combinatorial, $B^0 \to K^{*0}J/\psi$ leakage and partially reconstructed for $K\pi ee$ data.

A comparison between the $\cos\theta_\ell$ distributions is shown in Fig. F.1(a) and allows for three comments:

- the qualitative behaviour of the two distributions is similar as expected, since the dominant contribution should be given by DSL,
- both samples present a second peak for values of $\cos\theta_\ell$ close to -1 that is not reproduced in simulation,
- the relative abundance of this peak is not the same between the two samples, however additional sources of background in the $K\pi ee$ could be responsible for such differences (e.g. J/ψ-leakage or partially reconstructed decays).

The presence of a second peak in data but not in simulation seems to point to a reconstruction effect. One possible explanation for this discrepancy is a contribution from combinatorial-like components in data, such as the combination of a random K^* with a true DSL event, which could show up as a symmetric structure with peaks at both -1 and 1. This would imply two things: first such combinatorial-like contribution could be probably merged in the standard parametrization of the combinatorial, secondly, due to the fact that also the

$B^0 \to K^{*0}J/\psi$ leakage and partially reconstructed are symmetric in $\cos\theta_\ell$, we could manage to compare the asymmetric part of the DSL distribution (that is the dominant) in data between the $K\pi e\mu$ and the $K\pi ee$ dataset, by subtracting the left-hand side of the $\cos\theta_\ell$ distribution from the right-hand side. This attempt is shown in Fig. F.1(b), where the agreement shown between the two is very good.

In conclusion, it seems that, where it is possible to have a meaningful comparison, the agreement between the DSL shapes obtained from $K\pi e\mu$ data to $K\pi ee$ is such that no differences are expected when using a parametrization of the DSL background from $K\pi e\mu$ as a proxy of the DSL in $K\pi ee$ data.

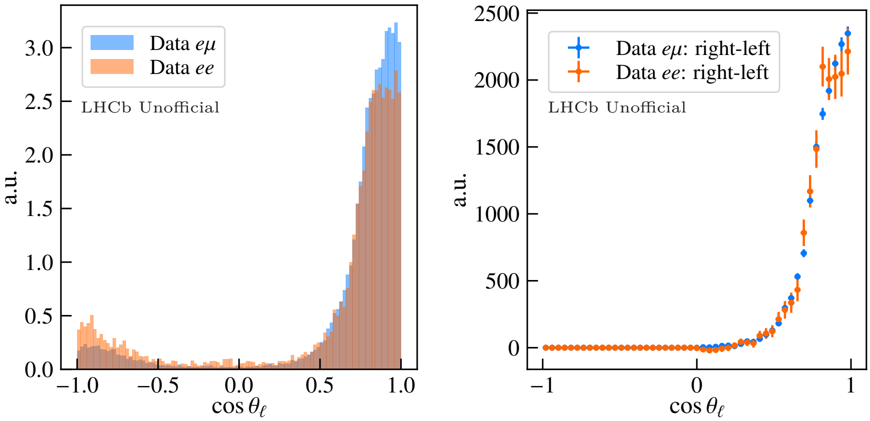


Figure F.1: On the left, a comparision between the $\cos\theta_\ell$ distribution in $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}e^+\pm^\mp$ data candidates selected in the DSL-enriched mass region of $m(K\pi ee) \in [3800, 4300]\,\text{MeV}/c^2$ with tight cut on the combinatorial MVA. On the right, the same plot where a subtraction of the symmetric component of the two distributions has been performed.

F.1.2 Including additional components to DSL simulation

To explain some of the differences observed between the LFV sample and the single component simulation, additional contributions have been studied and included in a mixture of different channels. Figure F.2 reports a screenshot of the different components included in the decay, while Figure F.3 shows a comparison between the simulation of $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\bar{\nu})e^+\nu$ and the mixture chosen. As can be noticed no clear difference between the two simulated samples can be noticed in the projections of the variables of interest.

```
Decay B0sig
  #HQET2 parameter taken from Spring 2019 HFLAV averages:
  #https://hflav-eos.web.cern.ch/hflav-eos/semi/spring19/html/ExclusiveVcb/exclBtoD.html
  #https://hflav-eos.web.cern.ch/hflav-eos/semi/spring19/html/ExclusiveVcb/exclBtoDstar.html
  0.021700   MyD-        e+ nu_e          PHOTOS  HQET2 1.131 1.081; #rho^2 as of HFLAV 2019 Spring, v1 unchanged
  0.016182   MyD*-   e+  nu_e             PHOTOS  HQET2 1.122 0.908 1.270 0.852; #rho^2 (ha1 unchanged) R1 and R2 as of HFLAV 2019
  0.0014355  MyD_0*-      e+  nu_e        PHOTOS  ISGW2;
  0.0006156  MyD'_1-      e+  nu_e        PHOTOS  ISGW2;
  0.0018477  MyD_1-       e+  nu_e        PHOTOS  ISGW2;
  0.0016516  MyD_2*-      e+  nu_e        PHOTOS  ISGW2;
  0.0001980  MyD-  pi0  e+  nu_e          PHOTOS  GOITY_ROBERTS;
  0.0002940  MyD-  pi0 pi0   e+  nu_e     PHOTOS  PHSP;
  0.0011970  MyD-  pi+ pi-   e+  nu_e     PHOTOS  PHSP;
  0.0001492  MyD*-  pi0  e+  nu_e         PHOTOS  GOITY_ROBERTS;
  0.0002374  MyD*-  pi0 pi0   e+  nu_e    PHOTOS  PHSP;
  0.0009021  MyD*-  pi+ pi-   e+  nu_e    PHOTOS  PHSP;
  0.0019096  MyD-    Mytau+ nu_tau        PHOTOS  ISGW2;
  0.0008411  MyD*-    Mytau+ nu_tau       PHOTOS  ISGW2;
  0.0000867  MyD_1-   Mytau+ nu_tau       PHOTOS  ISGW2;
  0.0000689  MyD_0*-   Mytau+ nu_tau      PHOTOS  ISGW2;
  0.0000534  MyD'_1-   Mytau+ nu_tau      PHOTOS  ISGW2;
  0.0001103  MyD_2*-   Mytau+ nu_tau      PHOTOS  ISGW2;
Enddecay
CDecay anti-B0sig
```

Figure F.2: Screenshot reporting the different contribution included in the mixture of different double-semileptonic decays used in the check.

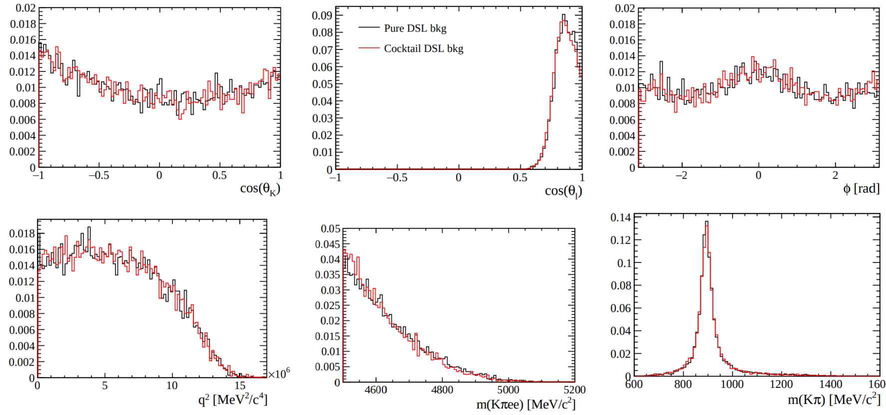


Figure F.3: Comparison between the simulation of $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\bar{\nu})e^+\nu$ and the mixture reported in Figure F.2.

F.1.3 DSL modelling from simulation

As an alternative parametrization, the double-semileptonic contribution can be obtained from simulated decays of $B^0 \to D^-(\to K^{*0}e^-\bar{\nu})e^+\nu$. The sample is reconstructed as $B^0 \to K^{*0}e^+e^-$ and undergoes the full selection chain of the rare mode. Additionally, the events are required to have a MVA > 0.97, q^2_c in central-q^2_c and to be in the mass range $m(K\pi e^+e^-) \in [4800, 5700]\,\text{MeV}/c^2$, slightly larger than the mass range of the fit to the rare mode, to increase the statistics available. The fit results obtained from a fit to these simulated samples, for different years and trigger categories, is reported in Table F.1. The same parametrization of Eq. 10.28 is used in these fits to parametrize the double-semileptonic component. Figure F.4 shows a projection of the fit result obtained for the Run 2 dataset in the L0I trigger category.

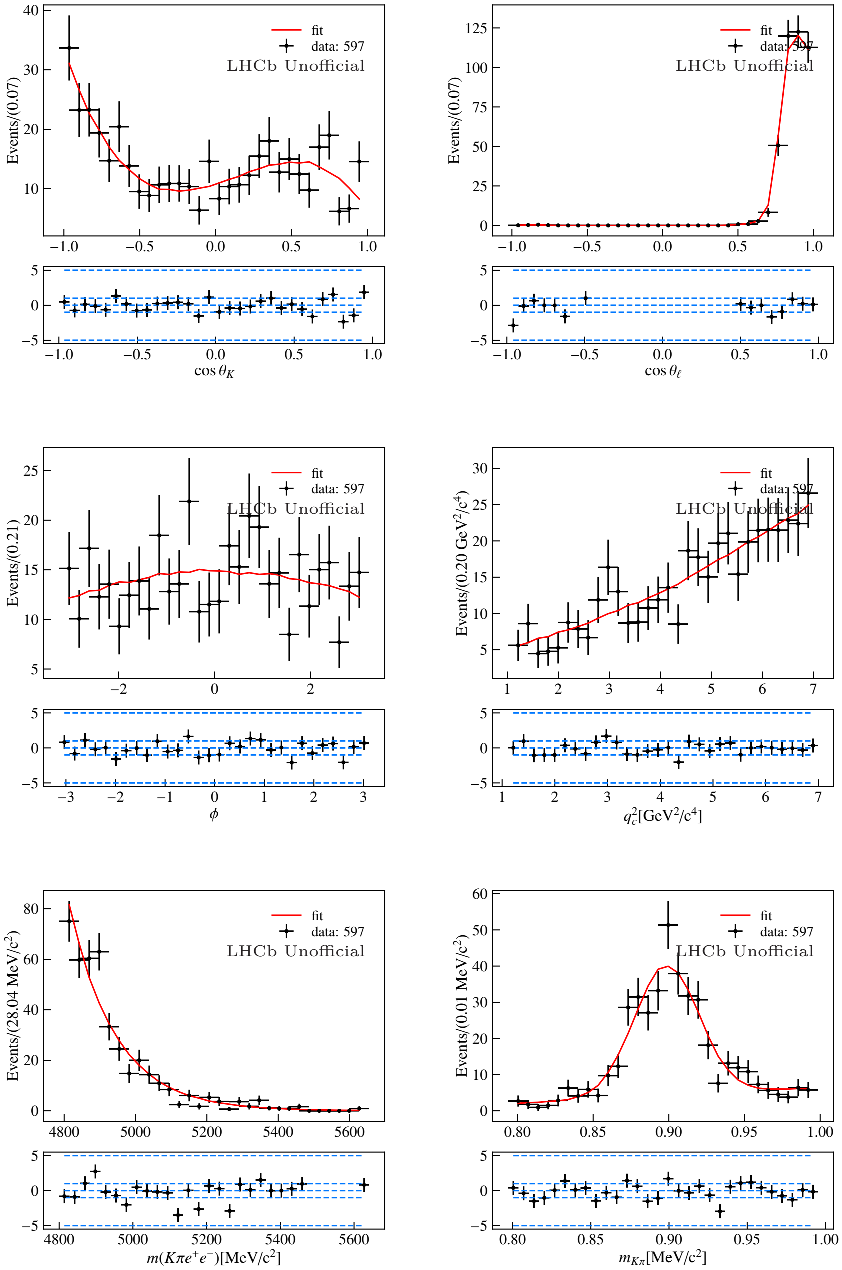


Figure F.4: Parametrization of the $B^0 \to D^-(\to K^{*0}e^-\bar{\nu})e^+\nu$ simulated sample corresponding to RUN 2 in the L0I trigger category.

Table F.1: Parameter values obtained from fits to the simulated double-semileptonic samples for all runs and trigger categories.

Year	RUN 1		RUN 2	
Type	L0E!	L0I	L0E!	L0I
$c^{K}_{1;DSL}$	-0.825 ± 0.168	-0.573 ± 0.143	-0.540 ± 0.120	-0.525 ± 0.090
$c^{\phi}_{1;DSL}$	0.011 ± 0.053	0.060 ± 0.045	-0.056 ± 0.041	0.001 ± 0.025
$c^{q^2}_{1;DSL}$	0.661 ± 0.164	0.589 ± 0.116	0.531 ± 0.099	0.704 ± 0.072
$c^{K}_{2;DSL}$	0.382 ± 0.139	0.334 ± 0.101	0.369 ± 0.083	0.297 ± 0.068
$c^{\phi}_{2;DSL}$	-0.015 ± 0.015	0.009 ± 0.014	0.023 ± 0.014	-0.010 ± 0.007
$c^{q^2}_{2;DSL}$	0.133 ± 0.176	-0.048 ± 0.121	0.068 ± 0.098	0.079 ± 0.072
$c^{K}_{3;DSL}$	-0.542 ± 0.139	-0.470 ± 0.113	-0.295 ± 0.092	-0.353 ± 0.070
$f^{K\pi}_{\mathcal{G};DSL}$	0.810 ± 0.086	0.720 ± 0.072	0.812 ± 0.070	0.699 ± 0.046
a_{DSL}	2.585 ± 0.946	0.344 ± 0.164	0.523 ± 0.553	-1.647 ± 10.086
$\mu^{K^{*0}}$	0.892 ± 0.004	0.898 ± 0.003	0.899 ± 0.003	0.899 ± 0.002
b_{DSL}	-1.986 ± 9.934	-0.237 ± 0.113	-0.393 ± 0.415	1.164 ± 7.582
$\sigma^{K^{*0}}$	0.024 ± 0.003	0.024 ± 0.003	0.027 ± 0.003	0.021 ± 0.002

F.2 Partially reconstructed $B^+ \to K_1^+ \to (K^+\pi^+\pi^-)e^+e^-$ decays

In the following are reported the fit results obtained from the parametrization of $B^+ \to K_1^+ \to (K^+\pi^+\pi^-)e^+e^-$ simulated decays corresponding to RUN 1 and RUN 2 conditions, separately for the trigger categories L0I and L0E!. The *pdf* used to parametrize the simulated samples corresponds exactly to the one described in Sec. 10.2.3. Figure F.5 shows the result of a fit to the weighted RUN 2 $B^+ \to K_1^+ \to (K^+\pi^+\pi^-)e^+e^-$ simulated decays in the L0I trigger category. Table F.2 reports the corresponding best parameters for all runs and trigger categories.

Table F.2: Parameter values obtained from fits to $B^+ \to K_1^+ e^+ e^-$ in Run 1 and Run 2 for the trigger categories L0I and L0E!.

Year	Run 1		Run 2	
Type	L0I	L0E!	L0I	L0E!
$c^{K}_{1;PR}$	-0.715 ± 0.105	-0.666 ± 0.114	-0.887 ± 0.027	-0.855 ± 0.033
$c^{\ell}_{1;PR}$	-0.067 ± 0.109	-0.152 ± 0.104	-0.017 ± 0.025	0.002 ± 0.029
$c^{\phi}_{1;PR}$	-0.050 ± 0.038	-0.001 ± 0.037	0.006 ± 0.009	-0.009 ± 0.011
$c^{q^2}_{1;PR}$	0.183 ± 0.111	-0.092 ± 0.119	0.035 ± 0.030	0.068 ± 0.035
$c^{K}_{2;PR}$	0.024 ± 0.105	0.025 ± 0.112	0.261 ± 0.025	0.256 ± 0.029
$c^{\ell}_{2;PR}$	-0.508 ± 0.122	-0.763 ± 0.112	-0.801 ± 0.029	-0.838 ± 0.033
$c^{\phi}_{2;PR}$	0.006 ± 0.012	-0.003 ± 0.011	-0.005 ± 0.003	-0.005 ± 0.003
$c^{q^2}_{2;PR}$	-0.014 ± 0.108	0.155 ± 0.116	0.141 ± 0.027	0.114 ± 0.032
$c^{K}_{3;PR}$	-0.140 ± 0.099	-0.129 ± 0.101	-0.211 ± 0.024	-0.163 ± 0.029
$f^{K\pi}_{\mathcal{G};PR}$	0.314 ± 0.063	0.400 ± 0.147	0.248 ± 0.024	0.209 ± 0.024
a_{PR}	0.001 ± 0.001	0.001 ± 0.002	0.002 ± 0.006	0.001 ± 0.008
$\mu^{K^{*0}}_{PR}$	0.887 ± 0.004	0.887 ± 0.007	0.896 ± 0.002	0.898 ± 0.002
b_{PR}	-0.001 ± 0.001	-0.001 ± 0.001	0.001 ± 0.004	0.000 ± 0.001
$\sigma^{K^{*0}}_{PR}$	0.018 ± 0.004	0.028 ± 0.011	0.025 ± 0.002	0.020 ± 0.002

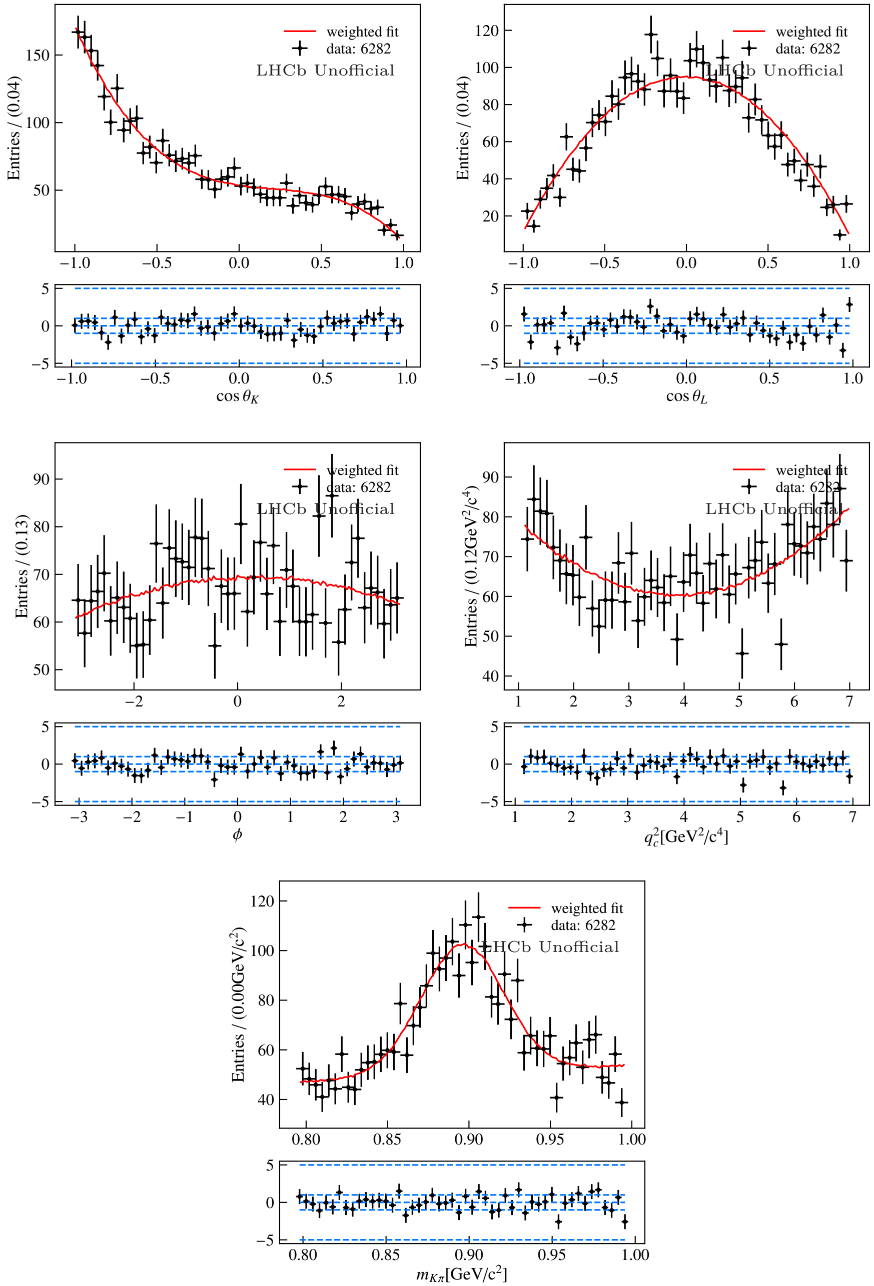


Figure F.5: Result of the fit to the $B^+ \to K_1^+e^+e^-$ simulation sample in RUN 2 for the trigger category L0I.

F.3 On the possibility of joining the parametrization of the combinatorial component

The possibility of joining the combinatorial components of different runs while keeping them split for different trigger categories in $B^0 \to K^{*0}e^+e^-$ data fits is studied using the same-sign sample. The selections applied are aligned as much as possible to the ones applied to the opposite-sign mode $B^0 \to K^{*0}e^+e^-$. but additionally the following changes/selections are applied:

- a selection in q^2_c corresponding to the central-q^2_c region;
- a looser mass range: $m(K\pi ee) \in [4900, 6800]\,\mathrm{MeV}/c^2$;
- a looser MVA cut: MVA> 0.3;
- the only events taken are the ones corresponding to leptons of the same sign. Samples including hadrons of the same sign are here ignored.

As can be seen, some of the selections have been loosened. This has been done to increase the statistics available for this check. No major distortion is expected due to these choices. To simplify the comparison the selected data has been fitted using a standard combinatorial shape, as introduced in Sec. 10.2.1, where the effect of the $B^+ \to K^+\ell^+\ell^-$ veto has not been considered. Additionally, the order of the Chebyshev polynomials used are up to order 3 for $\cos\theta_K$, up to order 5 for $\cos\theta_\ell$, up to order 2 for ϕ, q^2 and $m_{K\pi}$. Figures F.6, F.7, F.8 and F.7 show the result of the fits performed for Run 1 and Run 2 for the different trigger categories. Three main features can be noticed: the L0I sample is three times bigger than the L0E! and there is an overall good agreement between the fit and the data, except for the $\cos\theta_\ell$ distribution in Run 2 L0I, where a dip around zero can be observed. It is now possible compare the distributions obtained from the fits, by including also the uncertainty as obtained from the fit result. This comparison is shown in Figures F.10 and F.11. As can be seen there is an overall good agreement between the distributions, suggesting that using a single shapes for different runs should introduce a negligible bias.

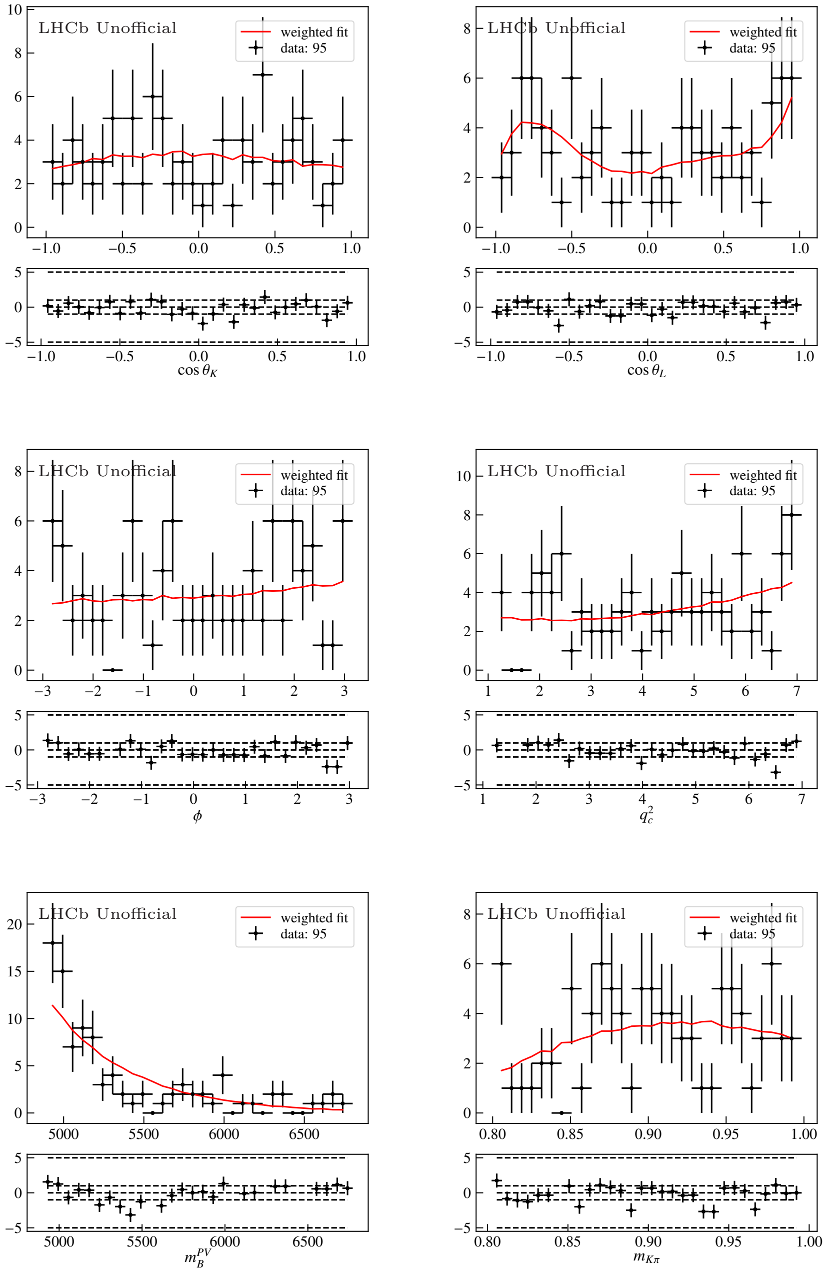


Figure F.6: Fit to the same-sign sample for Run 1 in the L0E! trigger category. The black points represent the data while the red line the best fit function.

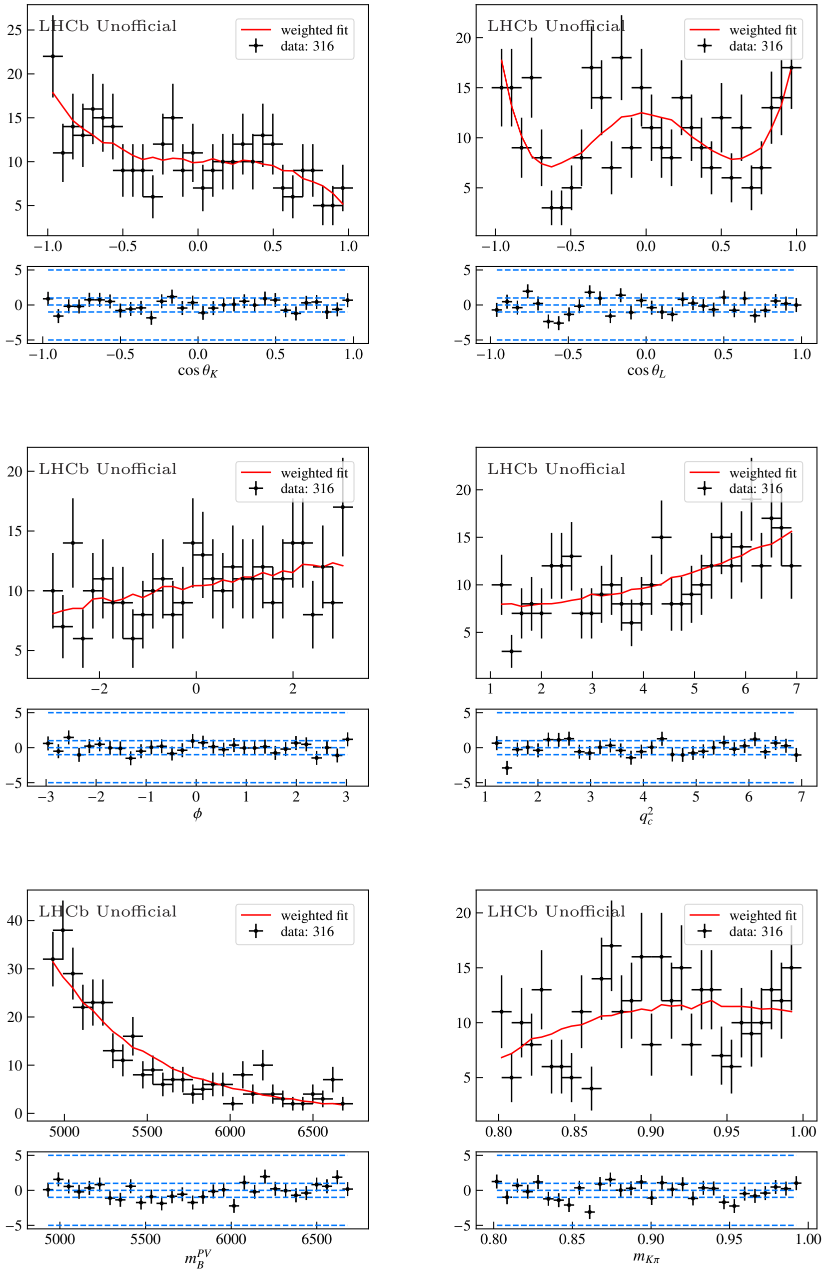


Figure F.7: Fit to the same-sign sample for Run 2 in the L0E! trigger category. The black points represent the data while the red line the best fit function.

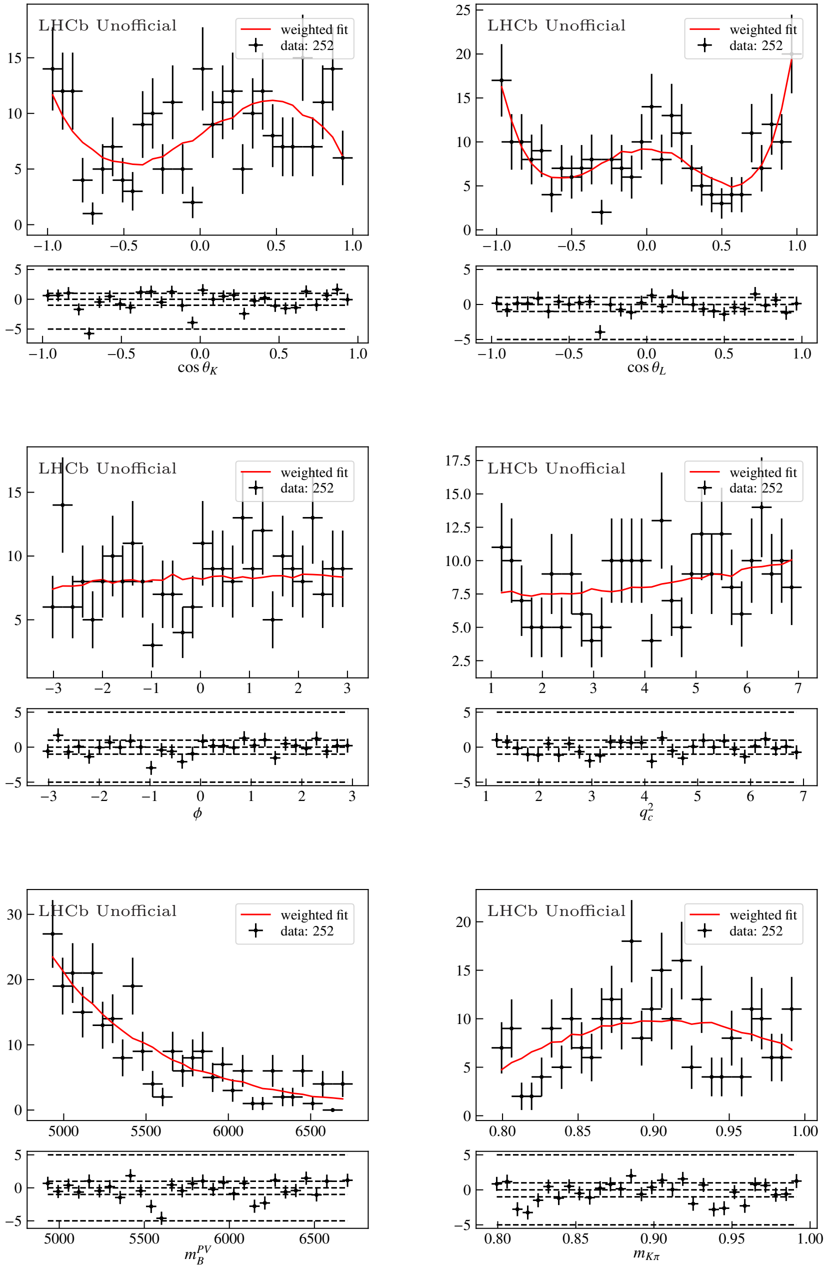


Figure F.8: Fit to the same-sign sample for Run 1 in the L0I trigger category. The black points represent the data while the red line the best fit function.

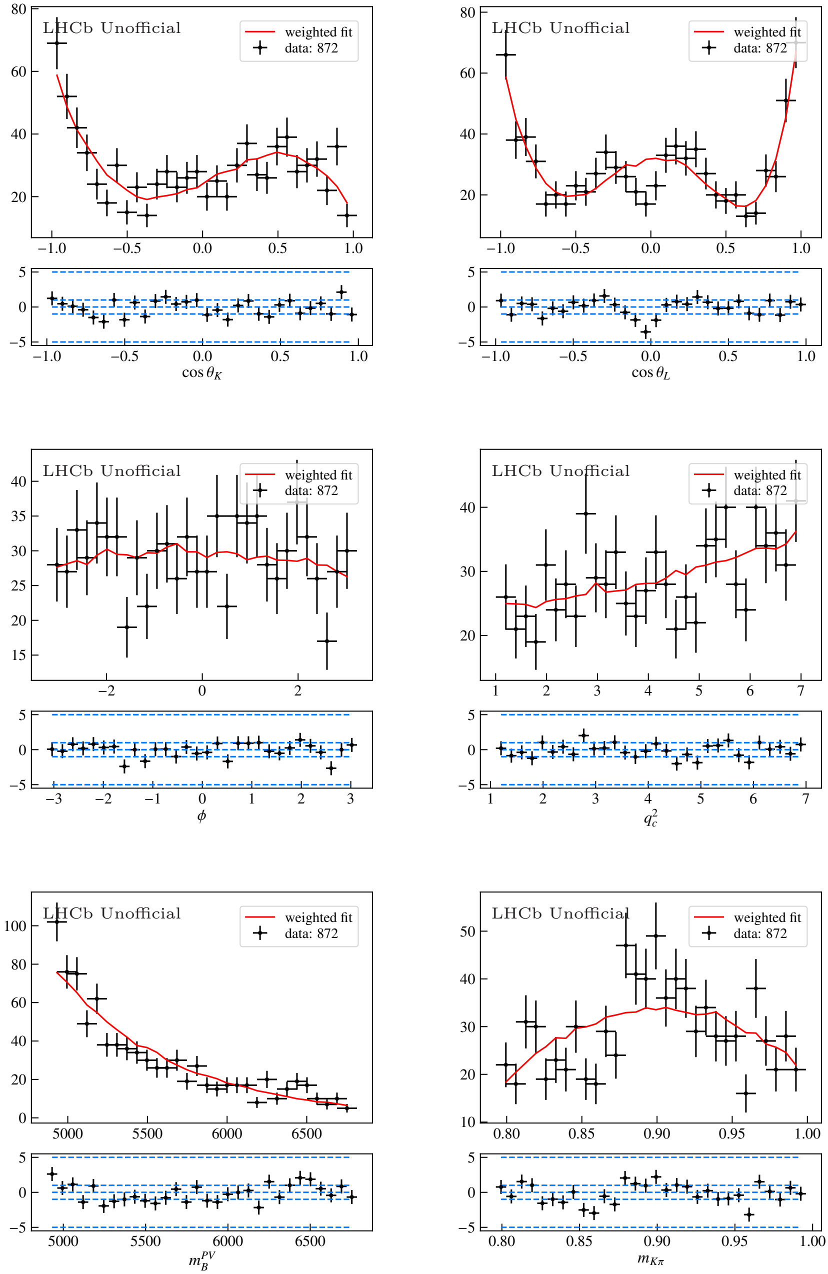


Figure F.9: Fit to the same-sign sample for Run 2 in the L0I trigger category. The black points represent the data while the red line the best fit function.

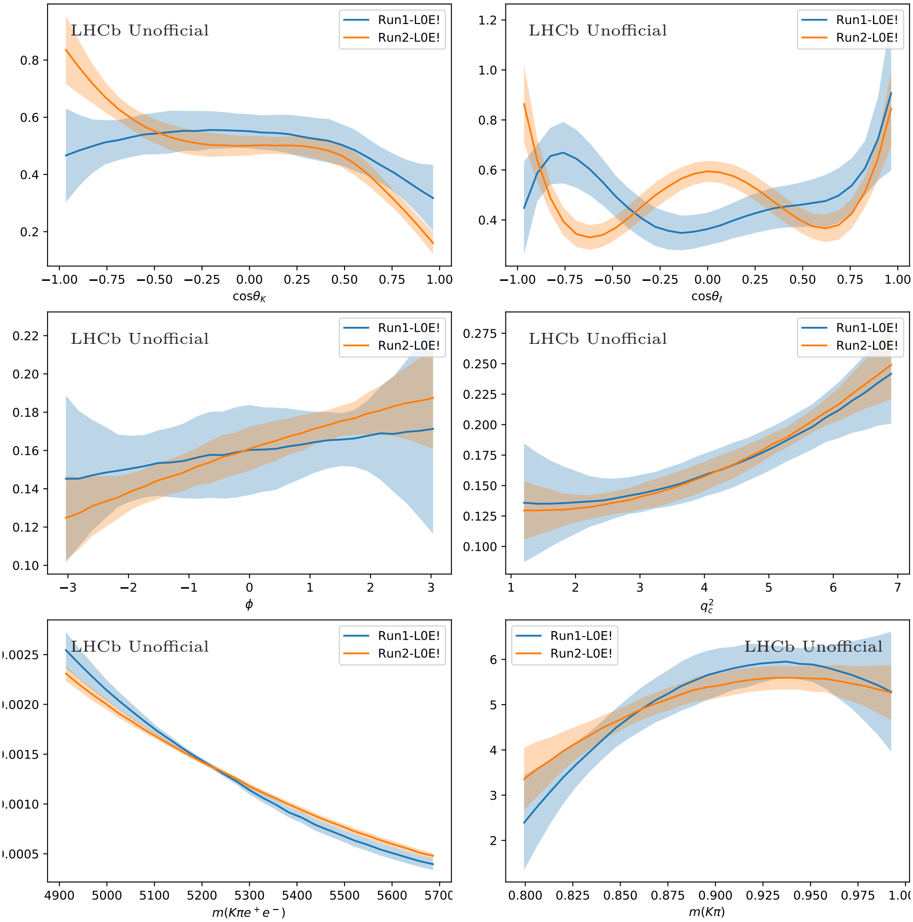


Figure F.10: Comparison between the shapes obtained from fits to same-sign data for the trigger category L0E! in Run 1 and Run 2.

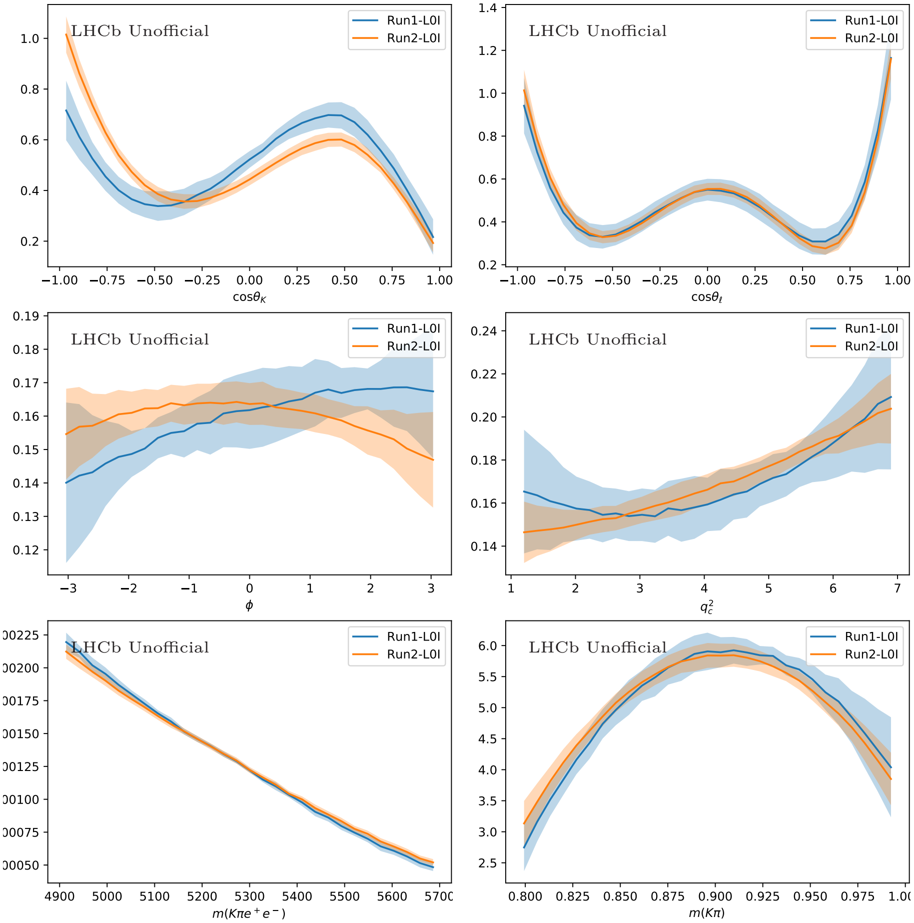


Figure F.11: Comparison between the shapes obtained from fits to same-sign data for the trigger category L0I in Run 1 and Run 2.

Appendix G

Blinded fit results to $B^0 \to K^{*0}\ell^+\ell^-$ decay candidates

In the following the fit results obtained for $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ candidates, separately for different years and trigger categories are shown in Figures from G.1 to G.6.

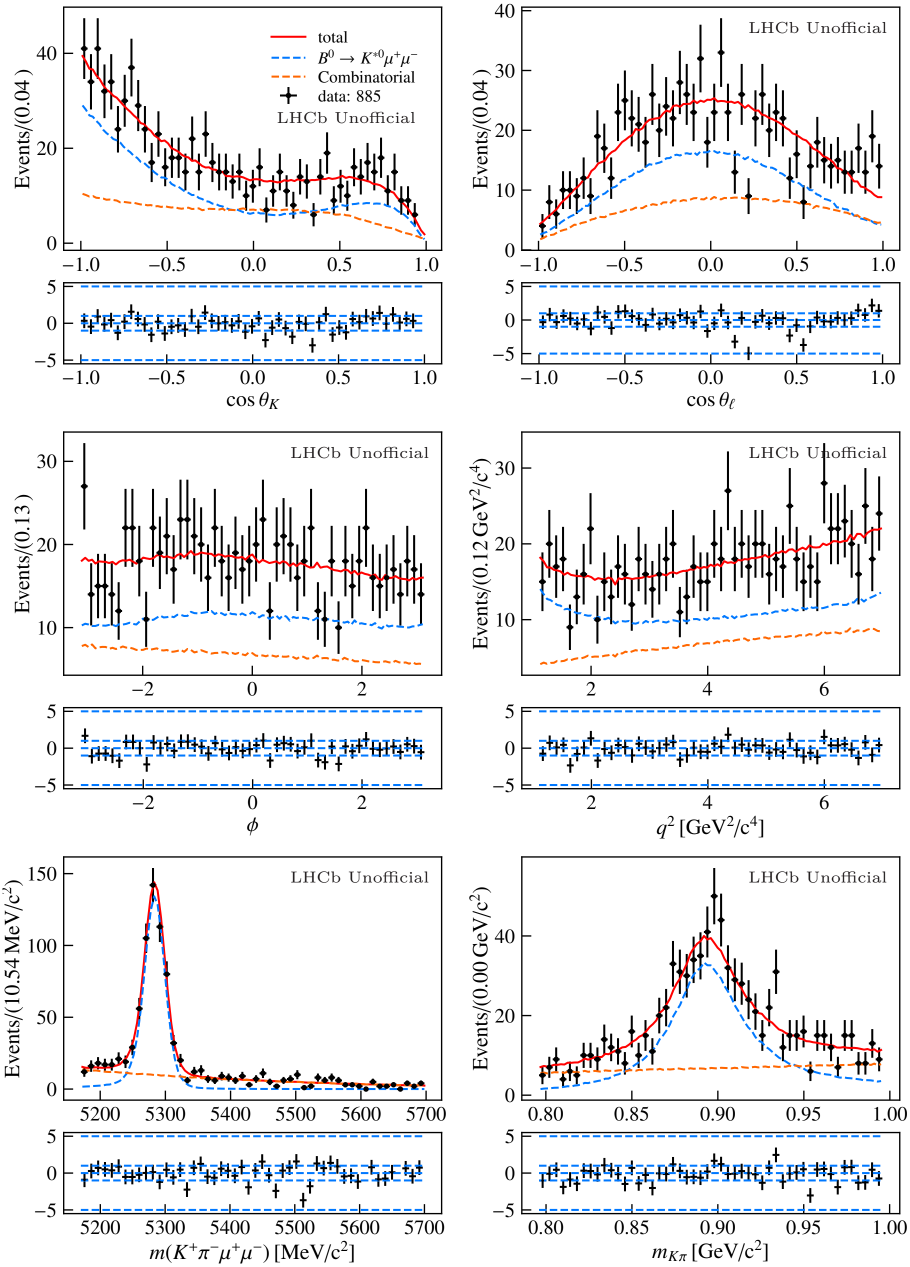


Figure G.1: Projection of the fit result for the $B^0 \to K^{*0}\mu^+\mu^-$ amplitude fit for Run 1 in central-q^2.

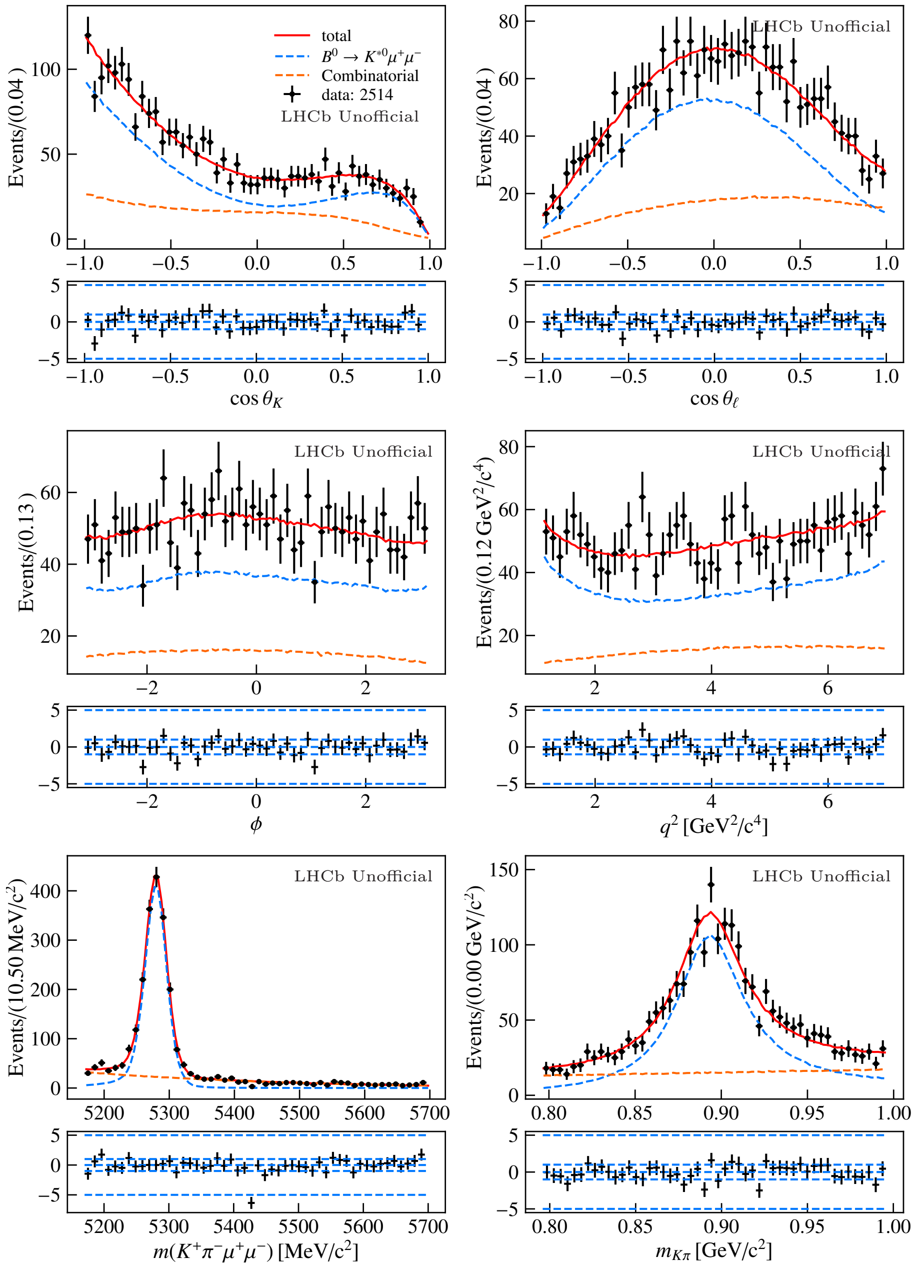


Figure G.2: Projection of the fit result for the $B^0 \to K^{*0}\mu^+\mu^-$ amplitude fit for Run 2 in central-q^2.

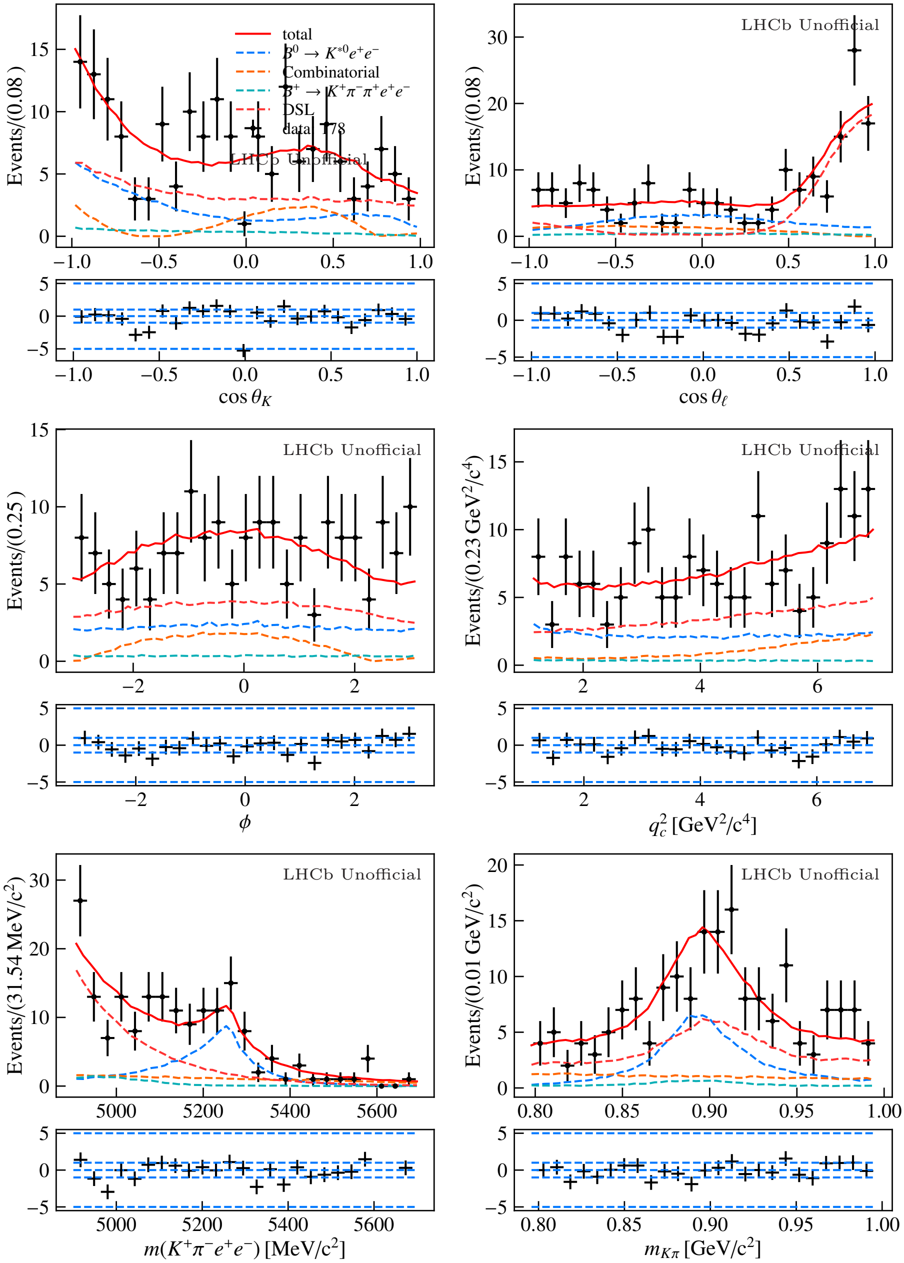


Figure G.3: Projection of the fit result for the $B^0 \to K^{*0}e^+e^-$ amplitude fit for L0I Run 1 in central-q^2_c.

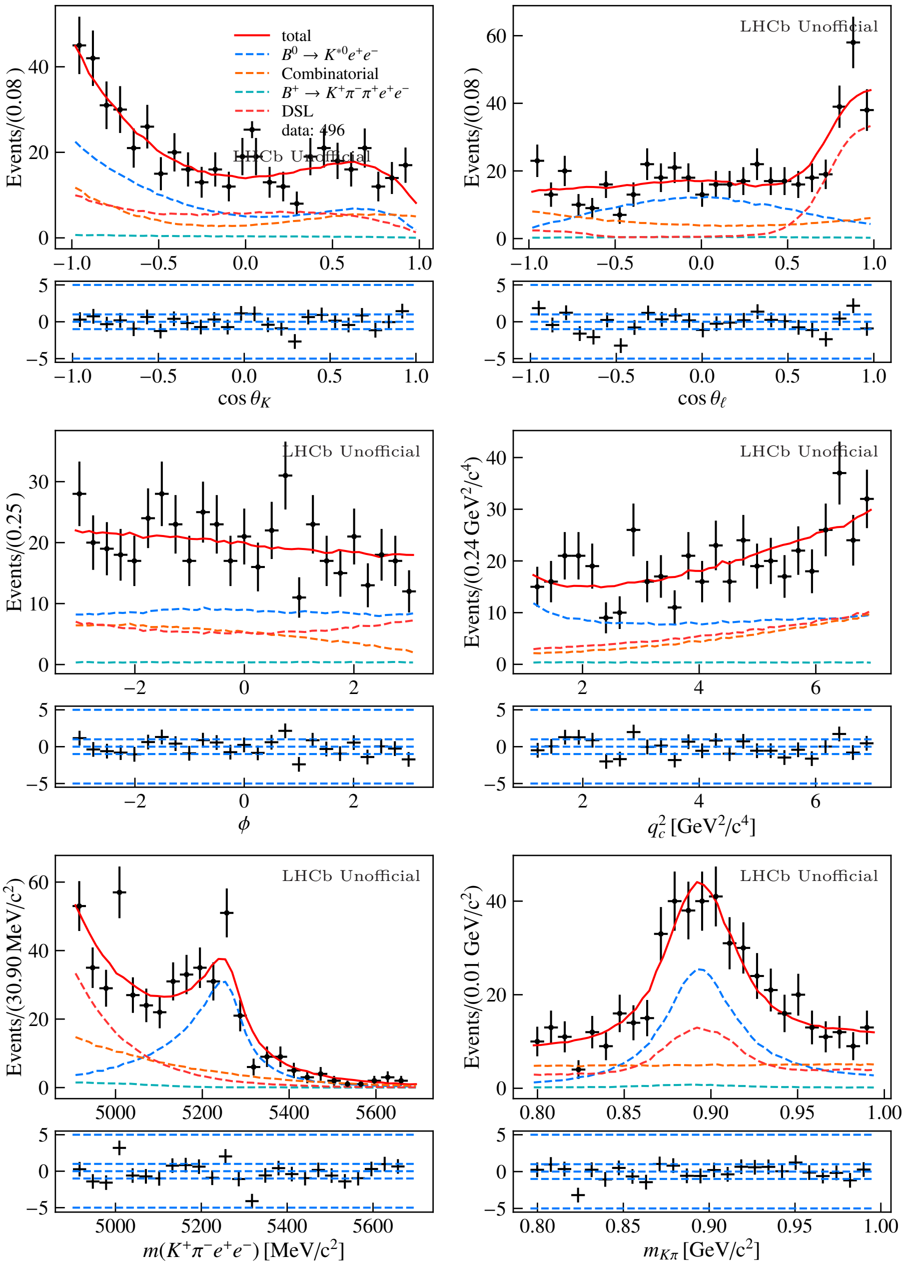


Figure G.4: Projection of the fit result for the $B^0 \to K^{*0}\mu^+\mu^-$ amplitude fit for L0I Run 2 in central-q_c^2.

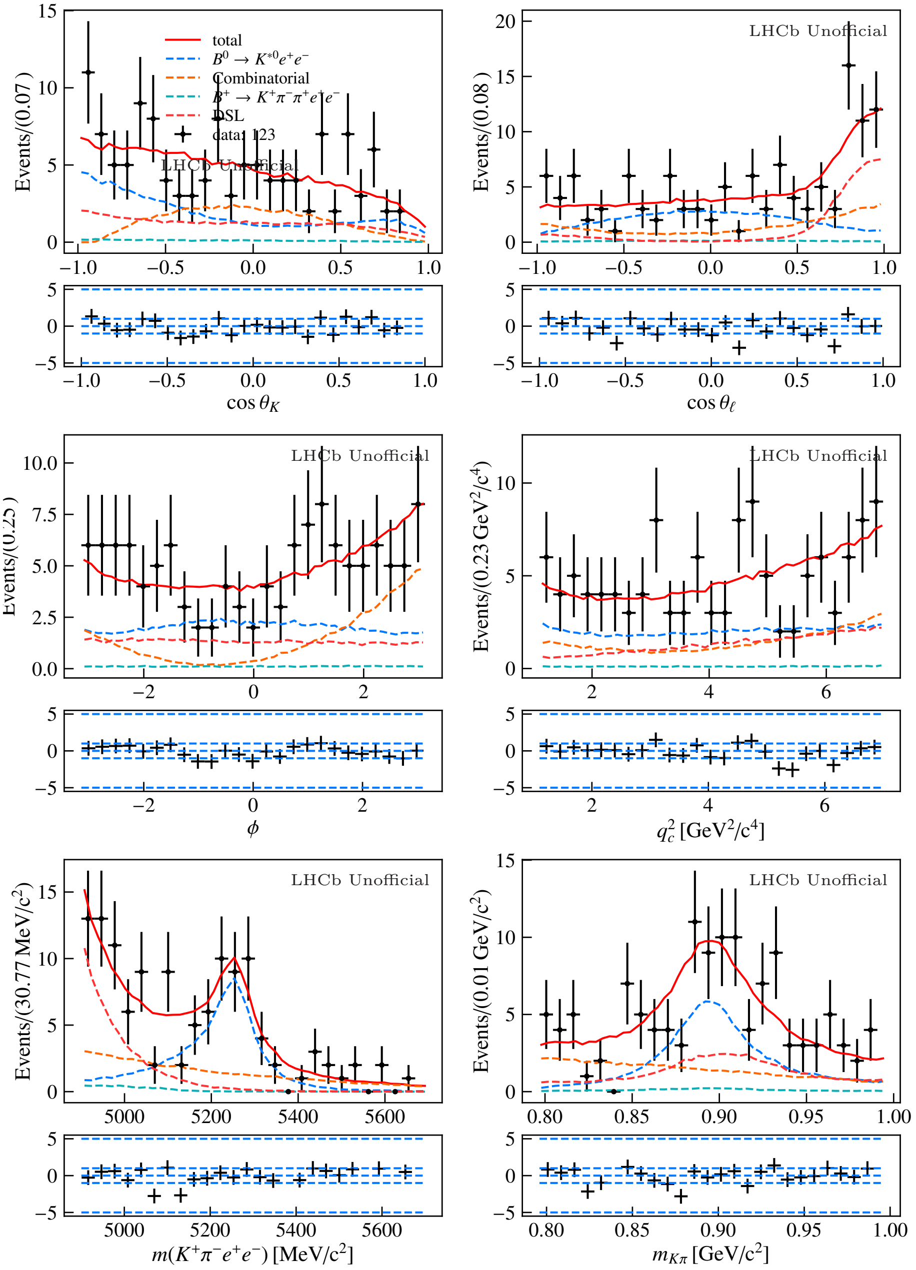


Figure G.5: Projection of the fit result for the $B^0 \to K^{*0}\mu^+\mu^-$ amplitude fit for L0E! Run 1 in central-q^2_c.

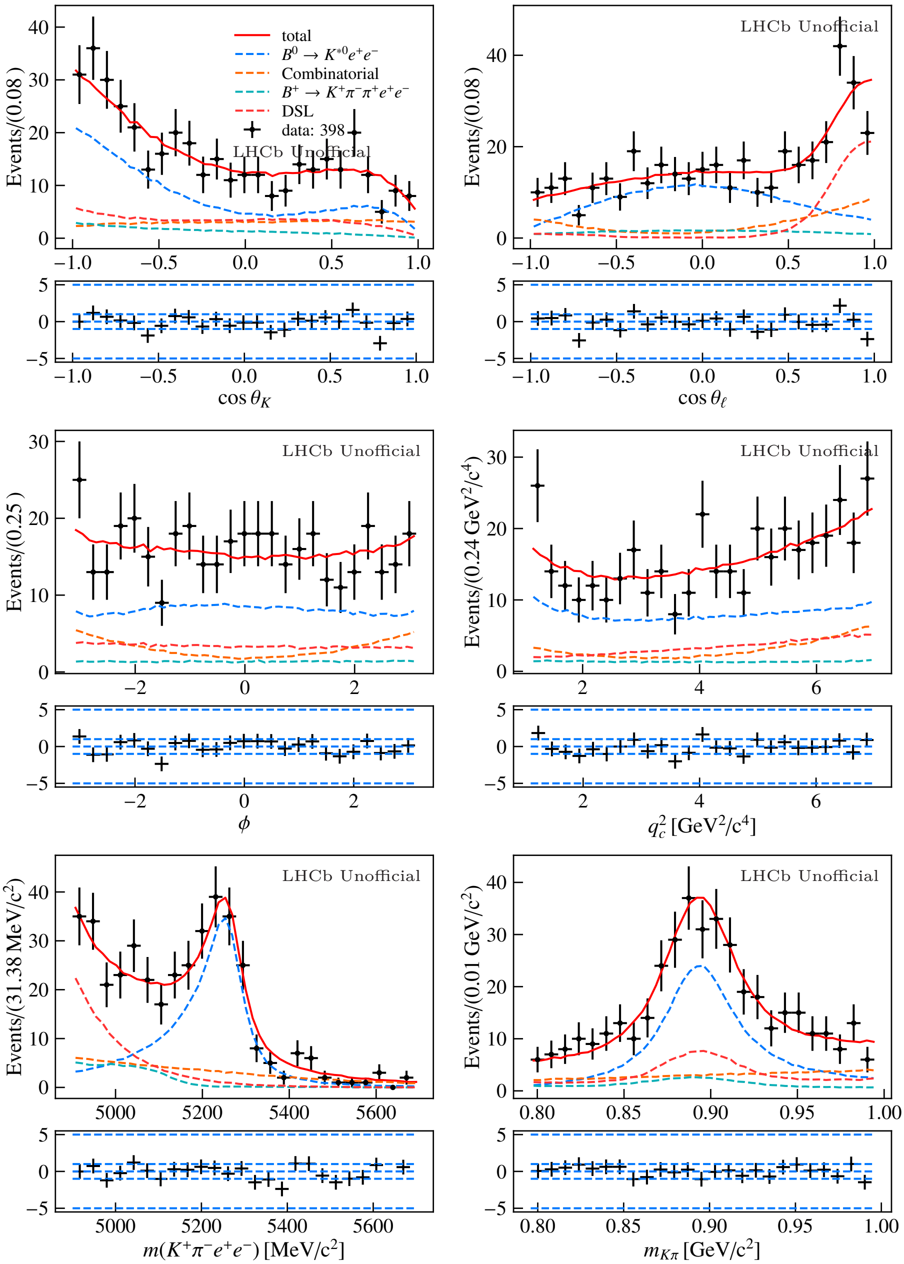


Figure G.6: Projection of the fit result for the $B^0 \to K^{*0}\mu^+\mu^-$ amplitude fit for L0E! Run 2 in central-q_c^2.

Bibliography

[1] Aad, G. et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* vol. 716 (2012), pp. 1–29. arXiv: `1207.7214 [hep-ex]`.

[2] Chatrchyan, S. et al. “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”. In: *Phys. Lett. B* vol. 716 (2012), pp. 30–61. arXiv: `1207.7235 [hep-ex]`.

[3] Standard Model. *Standard Model — Wikipedia, The Free Encyclopedia.* [Online; accessed 29-June-2022].

[4] Abel, C. et al. “Measurement of the Permanent Electric Dipole Moment of the Neutron”. In: *Phys. Rev. Lett.* vol. 124, no. 8 (2020), p. 081803. arXiv: `2001.11966 [hep-ex]`.

[5] Chau, L.-L. and Keung, W.-Y. “Comments on the Parametrization of the Kobayashi-Maskawa Matrix”. In: *Phys. Rev. Lett.* vol. 53 (1984), p. 1802.

[6] Wolfenstein, L. “Parametrization of the Kobayashi-Maskawa Matrix”. In: *Phys. Rev. Lett.* vol. 51 (1983), p. 1945.

[7] *J. Charles et al.* http://ckmfitter.in2p3.fr. Accessed: 2022-29-11.

[8] *M.Bona et al.* http://www.utfit.org/. Accessed: 2022-29-11.

[9] Lenz, A. “Constraints on a fourth generation of fermions from Higgs Boson searches”. In: *Adv. High Energy Phys.* vol. 2013 (2013), p. 910275.

[10] Schael, S. et al. “Precision electroweak measurements on the Z resonance”. In: *Phys. Rept.* vol. 427 (2006), pp. 257–454. arXiv: hep-ex/0509008.

[11] Fukuda, Y. et al. “Measurement of the flux and zenith angle distribution of upward through going muons by Super-Kamiokande”. In: *Phys. Rev. Lett.* vol. 82 (1999), pp. 2644–2648. arXiv: hep-ex/9812014.

[12] Ahmad, Q. R. et al. “Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory”. In: *Phys. Rev. Lett.* vol. 89 (2002), p. 011301. arXiv: nucl-ex/0204008.

[13] Gouvêa, A. de. “Neutrino Mass Models”. In: *Ann. Rev. Nucl. Part. Sci.* vol. 66 (2016), pp. 197–217.

[14] Polchinski, J. “Effective field theory and the Fermi surface”. In: *Theoretical Advanced Study Institute (TASI 92): From Black Holes and Strings to Particles.* June 1992, pp. 0235–276. arXiv: hep-th/9210046.

[15] Wilson, K. G. and Zimmermann, W. “Operator product expansions and composite field operators in the general framework of quantum field theory”. In: *Commun. Math. Phys.* vol. 24 (1972), pp. 87–106.

[16] Mannel, T. "Effective Field Theories in Flavor Physics". In: *Springer Tracts Mod. Phys.* vol. 203 (2004), pp. 1–175.

[17] Wilson, K. G. "The Renormalization Group: Critical Phenomena and the Kondo Problem". In: *Rev. Mod. Phys.* vol. 47 (1975), p. 773.

[18] Callan Jr., C. G. "Broken scale invariance in scalar field theory". In: *Phys. Rev. D* vol. 2 (1970), pp. 1541–1547.

[19] Symanzik, K. "Small distance behavior in field theory and power counting". In: *Commun. Math. Phys.* vol. 18 (1970), pp. 227–246.

[20] Buras, A. *Gauge Theory of Weak Decays.* Cambridge University Press, June 2020.

[21] Chetyrkin, K. G., Misiak, M., and Munz, M. "Weak radiative B meson decay beyond leading logarithms". In: *Phys. Lett. B* vol. 400 (1997). [Erratum: Phys.Lett.B 425, 414 (1998)], pp. 206–219. arXiv: hep-ph/9612313.

[22] Blake, T., Gershon, T., and Hiller, G. "Rare b hadron decays at the LHC". In: *Ann. Rev. Nucl. Part. Sci.* vol. 65 (2015), pp. 113–143. arXiv: `1501.03309 [hep-ex]`.

[23] Bobeth, C. et al. "Long-distance effects in $B \to K^* \ell\ell$ from analyticit y". In: *Eur. Phys. J. C* vol. 78, no. 6 (2018), p. 451. arXiv: `1707.07305 [hep-ph]`.

[24] Aaij, R. et al. "Improved measurement of $B^0_{(s,d)} \to \mu^+\mu^-$ decays". In: (2021). in preparation.

[25] Isidori, G. et al. "On the significance of new physics in $b \to s\ell^+\ell^-$ decays". In: *Phys. Lett. B* vol. 822 (2021), p. 136644. arXiv: `2104.05631 [hep-ph]`.

[26] Altmannshofer, W. et al. "Symmetries and Asymmetries of $B \to K^*\mu^+\mu^-$ Decays in the Standard Model and Beyond". In: *JHEP* vol. 01 (2009), p. 019. arXiv: `0811.1214 [hep-ph]`.

[27] Bobeth, C., Hiller, G., and Piranishvili, G. "CP Asymmetries in $\bar{B} \to \bar{K}^*(\to \bar{K}\pi)\bar{\ell}\ell$ and Untagged $\bar{B}_s$, $B_s \to \phi(\to K^+K^-)\bar{\ell}\ell$ Decays at NLO". In: *JHEP* vol. 07 (2008), p. 106. arXiv: `0805.2525 [hep-ph]`.

[28] Aaij, R. et al. "Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,\text{fb}^{-1}$ of integrated luminosity". In: *JHEP* vol. 02 (2016), p. 104. arXiv: `1512.04442 [hep-ex]`.

[29] Aaij, R. et al. "Measurement of *CP*-averaged observables in the $B^0 \to K^{*0}\mu^+\mu^-$ decay". In: *Phys. Rev. Lett.* vol. 125 (2020), p. 011802. arXiv: `2003.04831 [hep-ex]`.

[30] Kruger, F. and Matias, J. "Probing new physics via the transverse amplitudes of $B^0 \to K^{*0}(\to K^-\pi^+)l^+l^-$ at large recoil". In: *Phys. Rev. D* vol. 71 (2005), p. 094009. arXiv: hep-ph/0502060.

[31] Gratrex, J., Hopfer, M., and Zwicky, R. “Generalised helicity formalism, higher moments and the $B \to K_{J_K} (\to K\pi)\bar{\ell}_1\ell_2$ angular distributions”. In: *Phys. Rev. D* vol. 93, no. 5 (2016), p. 054008. arXiv: 1506.03970 [hep-ph].

[32] Bobeth, C., Hiller, G., and Dyk, D. van. “General analysis of $\bar{B} \to \bar{K}^{(*)}\ell^+\ell^-$ decays at low recoil”. In: *Phys. Rev.* vol. D87, no. 3 (2013). [Phys. Rev.D87,034016(2013)], p. 034016. arXiv: 1212.2321 [hep-ph].

[33] Aaij, R. et al. “Measurements of the S-wave fraction in $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \to K^*(892)^0\mu^+\mu^-$ differential branching fraction”. In: *JHEP* vol. 11 (2016), p. 047. arXiv: 1606.04731 [hep-ex].

[34] Bharucha, A., Straub, D. M., and Zwicky, R. “$B \to V\ell^+\ell^-$ in the Standard Model from light-cone sum rules”. In: *JHEP* vol. 08 (2016), p. 098. arXiv: 1503.05534 [hep-ph].

[35] Horgan, R. R. et al. “Lattice QCD calculation of form factors describing the rare decays $B \to K^*\ell^+\ell^-$ and $B_s \to \phi\ell^+\ell^-$”. In: *Phys. Rev.* vol. D89, no. 9 (2014), p. 094501. arXiv: 1310.3722 [hep-lat].

[36] Altmannshofer, W. and Straub, D. M. “New physics in $b \to s$ transitions after LHC run 1”. In: *Eur. Phys. J. C* vol. 75, no. 8 (2015), p. 382. arXiv: 1411.3161 [hep-ph].

[37] Aaij, R. et al. “Angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay”. In: (2020). to appear in Phys. Rev. Lett. arXiv: 2012.13241 [hep-ex].

[38] Wehle, S. et al. “Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$”. In: (2016). arXiv: 1612.05014 [hep-ex].

[39] Aaboud, M. et al. “Angular analysis of $B_d^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *JHEP* vol. 10 (2018), p. 047. arXiv: 1805.04000 [hep-ex].

[40] “Measurement of the P_1 and P_5' angular parameters of the decay $\mathrm{B}^0 \to \mathrm{K}^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV”. In: (2017).

[41] Descotes-Genon, S. et al. “On the impact of power corrections in the prediction of $B \to K^*\mu^+\mu^-$ observables”. In: *JHEP* vol. 12 (2014), p. 125. arXiv: 1407.8526 [hep-ph].

[42] Khodjamirian, A. et al. “Charm-loop effect in $B \to K^{(*)}\ell^+\ell^-$ and $B \to K^*gamma$”. In: *JHEP* vol. 09 (2010), p. 089. arXiv: 1006.4945 [hep-ph].

[43] Matias, J. et al. “Complete Anatomy of $\bar{B}_d \to \bar{K}^{*0}(\to K\pi)l^+l^-$ and its angular distribution”. In: *JHEP* vol. 04 (2012), p. 104. arXiv: 1202.4266 [hep-ph].

[44] Descotes-Genon, S. et al. “Implications from clean observables for the binned analysis of $B \to K * \mu^+\mu^-$ at large recoil”. In: *JHEP* vol. 01 (2013), p. 048. arXiv: 1207.2753 [hep-ph].

[45] Khodjamirian, A., Mannel, T., and Wang, Y. M. "$B \to K\ell^+\ell^-$ decay at large hadronic recoil". In: *JHEP* vol. 02 (2013), p. 010. arXiv: 1211.0234 [hep-ph].

[46] Horgan, R. R. et al. "Calculation of $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ observables using form factors from lattice QCD". In: *Phys. Rev. Lett.* vol. 112 (2014), p. 212003. arXiv: 1310.3887 [hep-ph].

[47] Beaujean, F., Bobeth, C., and Dyk, D. van. "Comprehensive Bayesian analysis of rare (semi)leptonic and radiative B decays". In: *Eur. Phys. J. C* vol. 74 (2014). [Erratum: Eur.Phys.J.C 74, 3179 (2014)], p. 2897. arXiv: 1310.2478 [hep-ph].

[48] Hambrock, C. et al. "$B \to K^\star$ form factors from flavor data to QCD and back". In: *Phys. Rev. D* vol. 89, no. 7 (2014), p. 074014. arXiv: 1308.4379 [hep-ph].

[49] Lyon, J. and Zwicky, R. "Resonances gone topsy turvy - the charm of QCD or new physics in $b \to s\ell^+\ell^-$?" In: (June 2014). arXiv: 1406.0566 [hep-ph].

[50] Jäger, S. and Martin Camalich, J. "Reassessing the discovery potential of the $B \to K^*\ell^+\ell^-$ decays in the large-recoil region: SM challenges and BSM opportunities". In: *Phys. Rev.* vol. D93, no. 1 (2016), p. 014028. arXiv: 1412.3183 [hep-ph].

[51] Jäger, S. and Martin Camalich, J. "On $B \to V\ell\ell$ at small dilepton invariant mass, power corrections, and new physics". In: *JHEP* vol. 05 (2013), p. 043. arXiv: 1212.2263 [hep-ph].

[52] Aaij, R. et al. "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays". In: *JHEP* vol. 08 (2017), p. 055. arXiv: 1705.05802 [hep-ex].

[53] Lees, J. P. et al. "Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \to K^{(*)}l^+l^-$". In: *Phys. Rev. D* vol. 86 (2012), p. 032012. arXiv: 1204.3933 [hep-ex].

[54] Abdesselam, A. et al. "Test of Lepton-Flavor Universality in $B \to K^*\ell^+\ell^-$ Decays at Belle". In: *Phys. Rev. Lett.* vol. 126, no. 16 (2021), p. 161801. arXiv: 1904.02440 [hep-ex].

[55] Bordone, M., Isidori, G., and Pattori, A. "On the Standard Model predictions for R_K and R_{K^*}". In: *Eur. Phys. J.* vol. C76, no. 8 (2016), p. 440. arXiv: 1605.07633 [hep-ph].

[56] Capdevila, B. et al. "Assessing lepton-flavour non-universality from $B \to K^*\ell\ell$ angular analyses". In: *JHEP* vol. 10 (2016), p. 075. arXiv: 1605.03156 [hep-ph].

[57] Aaij, R. et al. "Differential branching fractions and isospin asymmetries of $\to K^{(*)}\mu^+\mu^-$ decays". In: *JHEP* vol. 06 (2014), p. 133. arXiv: 1403.8044 [hep-ex].

[58] Aaij, R. et al. "Differential branching fraction and angular analysis of the decay $B_s^0 \to \phi\mu^+\mu^-$". In: *JHEP* vol. 07 (2013), p. 084. arXiv: `1305.2168 [hep-ex]`.

[59] Aaij, R. et al. "Angular analysis of the rare decay $B_s^0 \to \phi\mu^+\mu^-$". In: *JHEP* vol. 11 (2021), p. 043. arXiv: `2107.13428 [hep-ex]`.

[60] Aaij, R. et al. "Test of lepton universality in beauty-quark decays". In: (2021). Submitted to Nature Physics. arXiv: `2103.11769 [hep-ex]`.

[61] Aaij, R. et al. "Tests of lepton universality using $B^0 \to K_S^0\ell^+\ell^-$ and $B^+ \to K^{*+}\ell^+\ell^-$ decays". In: *Phys. Rev. Lett.* vol. 128 (2022), p. 191802. arXiv: `2110.09501 [hep-ex]`.

[62] Aaij, R. et al. "Test of lepton universality using $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays". In: *JHEP* vol. 05 (2020), p. 040. arXiv: `1912.08139 [hep-ex]`.

[63] Cornella, C. et al. "Reading the footprints of the B-meson flavor anomalies". In: *JHEP* vol. 08 (2021), p. 050. arXiv: `2103.16558 [hep-ph]`.

[64] Bauer, M. and Neubert, M. "Minimal Leptoquark Explanation for the $R_{D^{(*)}}$, R_K , and $(g-2)_\mu$ Anomalies". In: *Phys. Rev. Lett.* vol. 116, no. 14 (2016), p. 141802. arXiv: `1511.01900 [hep-ph]`.

[65] Cai, Y. et al. "Reconsidering the One Leptoquark solution: flavor anomalies and neutrino mass". In: *JHEP* vol. 10 (2017), p. 047. arXiv: `1704.05849 [hep-ph]`.

[66] Buttazzo, D. et al. "B-physics anomalies: a guide to combined explanations". In: *JHEP* vol. 11 (2017), p. 044. arXiv: `1706.07808 [hep-ph]`.

[67] Hiller, G., Loose, D., and Schönwald, K. "Leptoquark Flavor Patterns & B Decay Anomalies". In: *JHEP* vol. 12 (2016), p. 027. arXiv: `1609.08895 [hep-ph]`.

[68] Graverini, E. "Flavour anomalies: a review". In: *J. Phys. Conf. Ser.* vol. 1137, no. 1 (2019). Ed. by Bar~ao, F. et al., p. 012025. arXiv: `1807.11373 [hep-ex]`.

[69] Ciuchini, M. et al. "Lessons from the $B^{0,+} \to K^{*0,+}\mu^+\mu^-$ angular analyses". In: *Phys. Rev. D* vol. 103, no. 1 (2021), p. 015030. arXiv: `2011.01212 [hep-ph]`.

[70] Altmannshofer, W. and Stangl, P. "New physics in rare B decays after Moriond 2021". In: *Eur. Phys. J. C* vol. 81, no. 10 (2021), p. 952. arXiv: `2103.13370 [hep-ph]`.

[71] Algueró, M. et al. "$b \to s\ell^+\ell^-$ global fits after R_{K_S} and $R_{K^{*+}}$". In: *Eur. Phys. J. C* vol. 82, no. 4 (2022), p. 326. arXiv: `2104.08921 [hep-ph]`.

[72] Hurth, T. et al. "More Indications for Lepton Nonuniversality in $b \to s\ell^+\ell^-$". In: *Phys. Lett. B* vol. 824 (2022), p. 136838. arXiv: `2104.10058 [hep-ph]`.

[73] "LHC Machine". In: *JINST* vol. 3 (2008). Ed. by Evans, L. and Bryant, P., S08001.

[74] Aamodt, K. et al. "The ALICE experiment at the CERN LHC". In: *JINST* vol. 3 (2008), S08002.

[75] Aad, G. et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* vol. 3 (2008), S08003.

[76] Chatrchyan, S. et al. "The CMS Experiment at the CERN LHC". In: *JINST* vol. 3 (2008), S08004.

[77] Alves Jr., A. A. et al. "The LHCb Detector at the LHC". In: *JINST* vol. 3 (2008), S08005.

[78] Aaij, R. et al. "Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions". In: *Phys. Rev. Lett.* vol. 118, no. 5 (2017). [Erratum: Phys.Rev.Lett. 119, 169901 (2017)], p. 052002. arXiv: 1612.05140 `[hep-ex]`.

[79] Aaij, R. et al. "Prompt charm production in pp collisions at sqrt(s)=7 TeV". In: *Nucl. Phys. B* vol. 871 (2013), pp. 1–20. arXiv: 1302.2864 `[hep-ex]`.

[80] Aaij, R. et al. "Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* vol. 03 (2016). [Erratum: JHEP 09, 013 (2016), Erratum: JHEP 05, 074 (2017)], p. 159. arXiv: 1510.01707 `[hep-ex]`.

[81] Mangano, M. L. "Two lectures on heavy quark production in hadronic collisions". In: *Proc. Int. Sch. Phys. Fermi* vol. 137 (1998). Ed. by Bigi, I. I. Y. and Moroni, L., pp. 95–137. arXiv: hep-ph/9711337.

[82] Elsässer, C. *$\bar{b}b$ production angle plots.* URL: %7B%5Chref%7Bhttps://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb%5C_ProductionAngles.html%7D%7Bhttps://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb%5C_ProductionAngles.html%7D%7D.

[83] Aubert, B. et al. "The BaBar detector". In: *Nucl. Instrum. Meth. A* vol. 479 (2002), pp. 1–116. arXiv: hep-ex/0105044.

[84] Abashian, A. et al. "The Belle Detector". In: *Nucl. Instrum. Meth. A* vol. 479 (2002), pp. 117–232.

[85] Alves Jr., A. A. et al. "The LHCb detector at the LHC". In: *JINST* vol. 3, no. LHCb-DP-2008-001 (2008), S08005.

[86] *LHCb magnet: Technical Design Report.* Geneva, 2000.

[87] Aaij, R. et al. "Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC". In: *JINST* vol. 14, no. 04 (2019), P04013. arXiv: 1812.10790 `[hep-ex]`.

[88] Abellan Beteta, C. et al. "Monitoring radiation damage in the LHCb Tracker Turicensis". In: *JINST* vol. 15, no. 08 (2020), P08016. arXiv: 1809.05063 `[physics.ins-det]`.

[89] *LHCb Silicon Tracker - Material for Publications.* https://www.physik.uzh.ch/groups/lhcb/public/material/. Accessed: 2022-06-10.

[90] Arink, R. et al. "Performance of the LHCb Outer Tracker". In: *JINST* vol. 9 (2014), P01002. arXiv: 1311.3893 [physics.ins-det].

[91] Aaij, R. et al. "LHCb detector performance". In: *Int. J. Mod. Phys.* vol. A30 (2015), p. 1530022. arXiv: 1412.6352 [hep-ex].

[92] *LHCb reoptimized detector design and performance: Technical Design Report.* Geneva, 2003.

[93] Rulten, C. B. "Performance studies for the Cherenkov Telescope Array (CTA) with prospects for detecting pulsed gamma-ray emission". In: 2012.

[94] *PID Plots for Conferences.* https://twiki.cern.ch/twiki/bin/view/LHCb/PIDConferencePlots. Accessed: 2022-06-10.

[95] Fabjan, C. W. and Gianotti, F. "Calorimetry for particle physics". In: *Rev. Mod. Phys.* vol. 75 (2003), pp. 1243–1286.

[96] Virdee, T. S. "Experimental techniques". In: (1999).

[97] Leo, W. R. *Techniques for Nuclear and Particle Physics Experiments: A How to Approach.* 1987.

[98] Derkach, D., Hushchyn, M., and Kazeev, N. "Machine Learning based Global Particle Identification Algorithms at the LHCb Experiment". In: *EPJ Web Conf.* vol. 214 (2019). Ed. by Forti, A. et al., p. 06011.

[99] Sciascia, B. "LHCb Run2 trigger performance". In: *PoS* vol. BEAUTY2016 (2016), p. 029.

[100] Puig, A. *The LHCb trigger in 2011 and 2012.* Tech. rep. Geneva: CERN, Nov. 2014.

[101] Williams, M. et al. "The HLT2 Topological Lines". In: (Jan. 2011).

[102] Gauld, R., Goertz, F., and Haisch, U. "An explicit Z'-boson explanation of the $B \to K^* \mu^+ \mu^-$ anomaly". In: *JHEP* vol. 01 (2014), p. 069. arXiv: 1310.1082 [hep-ph].

[103] Buras, A. J. and Girrbach, J. "Left-handed Z' and Z FCNC quark couplings facing new $b \to s\mu^+\mu^-$ data". In: *JHEP* vol. 12 (2013), p. 009. arXiv: 1309.2466 [hep-ph].

[104] Altmannshofer, W. and Straub, D. M. "New Physics in $B \to K^* \mu\mu$?" In: *Eur. Phys. J.* vol. C73 (2013), p. 2646. arXiv: 1308.1501 [hep-ph].

[105] Altmannshofer, W. et al. "Quark flavor transitions in $L_\mu - L_\tau$ models". In: *Phys. Rev. D* vol. 89 (2014), p. 095033. arXiv: 1403.1269 [hep-ph].

[106] Crivellin, A. et al. "Lepton-flavour violating B decays in generic Z' models". In: *Phys. Rev.* vol. D92, no. 5 (2015), p. 054013. arXiv: 1504.07928 [hep-ph].

[107] Altmannshofer, W., Davighi, J., and Nardecchia, M. “Gauging the accidental symmetries of the standard model, and implications for the flavor anomalies”. In: *Phys. Rev. D* vol. 101, no. 1 (2020), p. 015004. arXiv: `1909.02021 [hep-ph]`.

[108] Hiller, G. and Schmaltz, M. “R_K and future $b \to s\ell\ell$ physics beyond the standard model opportunities”. In: *Phys. Rev.* vol. D90 (2014), p. 054014. arXiv: `1408.1627 [hep-ph]`.

[109] Biswas, S. et al. “Explaining the lepton non-universality at the LHCb and CMS within a unified framework”. In: *JHEP* vol. 02 (2015), p. 142. arXiv: `1409.0882 [hep-ph]`.

[110] Gripaios, B., Nardecchia, M., and Renner, S. A. “Composite leptoquarks and anomalies in B-meson decays”. In: *JHEP* vol. 05 (2015), p. 006. arXiv: `1412.1791 [hep-ph]`.

[111] Freytsis, M., Ligeti, Z., and Ruderman, J. T. “Flavor models for $\bar{B} \to D^{(*)}\tau\bar{\nu}$”. In: *Phys. Rev. D* vol. 92, no. 5 (2015), p. 054018. arXiv: `1506.08896 [hep-ph]`.

[112] Fornal, B., Gadam, S. A., and Grinstein, B. “Left-Right SU(4) Vector Leptoquark Model for Flavor Anomalies”. In: *Phys. Rev. D* vol. 99, no. 5 (2019), p. 055025. arXiv: `1812.01603 [hep-ph]`.

[113] Bečirević, D. et al. “Scalar leptoquarks from grand unified theories to accommodate the B-physics anomalies”. In: *Phys. Rev. D* vol. 98, no. 5 (2018), p. 055003. arXiv: `1806.05689 [hep-ph]`.

[114] Blanke, M. and Crivellin, A. “B Meson Anomalies in a Pati-Salam Model within the Randall-Sundrum Background”. In: *Phys. Rev. Lett.* vol. 121, no. 1 (2018), p. 011801. arXiv: `1801.07256 [hep-ph]`.

[115] Gherardi, V., Marzocca, D., and Venturini, E. “Low-energy phenomenology of scalar leptoquarks at one-loop accuracy”. In: *JHEP* vol. 01 (2021), p. 138. arXiv: `2008.09548 [hep-ph]`.

[116] Di Luzio, L., Greljo, A., and Nardecchia, M. “Gauge leptoquark as the origin of B-physics anomalies”. In: *Phys. Rev. D* vol. 96, no. 11 (2017), p. 115011. arXiv: `1708.08450 [hep-ph]`.

[117] Bordone, M. et al. “Low-energy signatures of the PS^3 model: from B-physics anomalies to LFV”. In: *JHEP* vol. 10 (2018), p. 148. arXiv: `1805.09328 [hep-ph]`.

[118] Angelescu, A. et al. “Closing the window on single leptoquark solutions to the B-physics anomalies”. In: *JHEP* vol. 10 (2018), p. 183. arXiv: `1808.08179 [hep-ph]`.

[119] Di Luzio, L. et al. “Maximal Flavour Violation: a Cabibbo mechanism for leptoquarks”. In: *JHEP* vol. 11 (2018), p. 081. arXiv: `1808.00942 [hep-ph]`.

[120] Cornella, C., Fuentes-Martin, J., and Isidori, G. “Revisiting the vector leptoquark explanation of the B-physics anomalies”. In: *JHEP* vol. 07 (2019), p. 168. arXiv: `1903.11517 [hep-ph]`.

[121] Aaij, R. et al. “Measurement of the phase difference between short- and long-distance amplitudes in the $B^+ \to K^+\mu^+\mu^-$ decay”. In: (2016). arXiv: `1612.06764 [hep-ex]`.

[122] Blake, T. et al. “An empirical model to determine the hadronic resonance contributions to $\overline{B}^0 \to \overline{K}^{*0}\mu^+\mu^-$ transitions”. In: *Eur. Phys. J. C* vol. 78, no. 6 (2018), p. 453. arXiv: `1709.03921 [hep-ph]`.

[123] Hurth, T., Langenbruch, C., and Mahmoudi, F. “Direct determination of Wilson coefficients using $B^0 \to K^{*0}\mu^+\mu^-$ decays”. In: *JHEP* vol. 11 (2017), p. 176. arXiv: `1708.04474 [hep-ph]`.

[124] Chrzaszcz, M. et al. “Prospects for disentangling long- and short-distance effects in the decays $B \to K^*\mu^+\mu^-$”. In: *JHEP* vol. 10 (2019), p. 236. arXiv: `1805.06378 [hep-ph]`.

[125] Cornella, C. et al. “Hunting for $B^+ \to K^+\tau^+\tau^-$ imprints on the $B^+ \to K^+\mu^+\mu^-$ dimuon spectrum”. In: *Eur. Phys. J. C* vol. 80, no. 12 (2020), p. 1095. arXiv: `2001.04470 [hep-ph]`.

[126] Mauri, A., Serra, N., and Silva Coutinho, R. “Towards establishing lepton flavor universality violation in $\bar{B} \to \bar{K}^*\ell^+\ell^-$ decays”. In: *Phys. Rev. D* vol. 99, no. 1 (2019), p. 013007. arXiv: `1805.06401 [hep-ph]`.

[127] Paul, A. and Straub, D. M. “Constraints on new physics from radiative B decays”. In: *JHEP* vol. 04 (2017), p. 027. arXiv: `1608.02556 [hep-ph]`.

[128] Albrecht, J. et al. *Test of lepton flavour universality with $b \to s\ell^+\ell^-$ decays.* LHCb-ANA-2020-069. 2020.

[129] Patrignani, C. et al. “Review of particle physics”. In: *Chin. Phys.* vol. C40 (2016). and 2017 update, p. 100001.

[130] Sjostrand, T., Mrenna, S., and Skands, P. Z. “A Brief Introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* vol. 178 (2008), pp. 852–867. arXiv: `0710.3820 [hep-ph]`.

[131] Belyaev, I. et al. “Handling of the generation of primary events in Gauss, the LHCb simulation framework”. In: *J. Phys. Conf. Ser.* vol. 331 (2011). Ed. by Lin, S. C., p. 032047.

[132] Lange, D. J. “The EvtGen particle decay simulation package”. In: *Nucl. Instrum. Meth. A* vol. 462 (2001). Ed. by Erhan, S., Schlein, P., and Rozen, Y., pp. 152–155.

[133] Davidson, N., Przedzinski, T., and Was, Z. “PHOTOS interface in C++: Technical and Physics Documentation”. In: *Comput. Phys. Commun.* vol. 199 (2016), pp. 86–101. arXiv: `1011.0937 [hep-ph]`.

[134] Agostinelli, S. et al. “GEANT4–a simulation toolkit”. In: *Nucl. Instrum. Meth. A* vol. 506 (2003), pp. 250–303.

[135] Clemencic, M. et al. “The LHCb simulation application, Gauss: Design, evolution and experience”. In: *J. Phys. Conf. Ser.* vol. 331 (2011). Ed. by Lin, S. C., p. 032023.

[136] Ball, P. and Zwicky, R. “$B_{d,s} \to \rho, \omega, K^*, \phi$ decay form-factors from light-cone sum rules revisited”. In: *Phys. Rev.* vol. D71 (2005), p. 014029. arXiv: hep-ph/0412079 [hep-ph].

[137] *Patrick Robbe.* https://gitlab.cern.ch/lhcb-datapkg/Gen/DecFiles/-/blob/v30r80/dkfiles/Bd_JpsiKst,ee=DecProdCut.dec. Accessed: 2022-29-11.

[138] Aaij, R. et al. “Strong constraints on the $b \to s\gamma$ photon polarisation from $B^0 \to K^{*0}e^+e^-$ decays”. In: *JHEP* vol. 12 (2020), p. 081. arXiv: 2010.06011 [hep-ex].

[139] Hulsbergen, W. D. “Decay chain fitting with a Kalman filter”. In: *Nucl. Instrum. Meth. A* vol. 552 (2005), pp. 566–575. arXiv: physics/0503191.

[140] Atzeni, M. et al. “Angular analysis of $B^0 \to K^{*0}e^+e^-$ decays”. In: *In preparation* (2021).

[141] Friedman, J. H. “Greedy Function Approximation: A Gradient Boosting Machine”. In: *The Annals of Statistics* vol. 29, no. 5 (2001), pp. 1189–1232.

[142] Chen, T. and Guestrin, C. “XGBoost: A Scalable Tree Boosting System”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD ’16. San Francisco, California, USA: ACM, 2016, pp. 785–794.

[143] Silverman, B. *Density estimation for statistics and data analysis.* Chapman and Hall, 1986.

[144] Mauri, A. “Direct and Indirect Searches for New Physics in $b \to s\ell\ell$ Decays”. Presented 20 Dec 2018. 2018.

[145] Pivk, M. and Le Diberder, F. R. “SPlot: A Statistical tool to unfold data distributions”. In: *Nucl. Instrum. Meth. A* vol. 555 (2005), pp. 356–369. arXiv: physics/0402083.

[146] Aaij, R. et al. “Measurement of the electron reconstruction efficiency at LHCb”. In: *JINST* vol. 14 (2019), P11023. arXiv: 1909.02957 [hep-ex].

[147] Breiman, L. et al. “Classification and Regression Trees”. In: 1983.

[148] Schapire, R. E. “The Strength of Weak Learnability”. In: *Mach. Learn.* vol. 5, no. 2 (July 1990), pp. 197–227.

[149] Rogozhnikov, A. “Reweighting with Boosted Decision Trees”. In: *Journal of Physics: Conference Series* vol. 762 (Oct. 2016), p. 012036.

[150] Skwarnicki, T. “A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances”. PhD thesis. Cracow, INP, 1986.

[151] Isidori, G. et al. "QED in $\bar{B} \to \bar{K}\ell^+\ell^-$ LFU ratios: Theory versus Experiment, a Monte Carlo Study". In: (May 2022). arXiv: 2205.08635 [hep-ph].

[152] Asatrian, H. H. et al. "Two loop virtual corrections to $B \to X_s l^+ l^-$ in the standard model". In: *Phys. Lett. B* vol. 507 (2001), pp. 162–172. arXiv: hep-ph/0103087.

[153] Abadi, M. et al. "Tensorflow: Large-scale machine learning on heterogeneous distributed systems". In: *arXiv preprint arXiv:1603.04467* (2016).

[154] James, F. and Roos, M. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations". In: *Comput. Phys. Commun.* vol. 10 (1975), pp. 343–367.

[155] Verkerke, W. and Kirkby, D. P. "The RooFit toolkit for data modeling". In: *eConf* vol. C0303241 (2003). Ed. by Lyons, L. and Karagoz, M., MOLT007. arXiv: physics/0306116.

[156] Aaij, R. et al. "Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi pK^-$ decays". In: *Phys. Rev. Lett.* vol. 115 (2015), p. 072001. arXiv: 1507.03414 [hep-ex].

[157] Aaij, R. et al. "Measurement of relative branching fractions of decays to $\psi(2S)$ and J/ψ mesons". In: *Eur. Phys. J.* vol. C72 (2012), p. 2118. arXiv: 1205.0918 [hep-ex].

[158] Aaij, R. et al. "Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays". In: *Phys. Rev.* vol. D88 (2013), p. 052002. arXiv: 1307.2782 [hep-ex].

[159] Aubert, B. et al. "Measurement of decay amplitudes of $B \to J/\psi K^*, \psi(2S)K^*$, and $\chi_{c1}K^*$ with an angular analysis". In: *Phys. Rev. D* vol. 76 (2007), p. 031102. arXiv: 0704.0522 [hep-ex].

[160] Chilikin, K. et al. "Experimental constraints on the spin and parity of the $Z(4430)^+$". In: *Phys. Rev. D* vol. 88, no. 7 (2013), p. 074026. arXiv: 1306.4894 [hep-ex].

[161] Chilikin, K. et al. "Observation of a new charged charmoniumlike state in $\bar{B}^0 \to J/\psi K^-\pi^+$ decays". In: *Phys. Rev. D* vol. 90, no. 11 (2014), p. 112009. arXiv: 1408.6457 [hep-ex].

[162] Aaij, R. et al. "Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays". In: *Phys. Rev. D* vol. 88 (2013), p. 052002. arXiv: 1307.2782 [hep-ex].

[163] Aaij, R. et al. "Measurement of CP-Averaged Observables in the $B^0 \to K^{*0}\mu^+\mu^-$ Decay". In: *Phys. Rev. Lett.* vol. 125, no. 1 (2020), p. 011802. arXiv: 2003.04831 [hep-ex].

[164] Cunliffe, S. et al. "Measurement of the differential branching fraction of $B^0 \to K^*(892)^0\mu^+\mu^-$ and S-wave fraction of $B^0 \to K^+\pi^-\mu^+\mu^-$ decays using 3 fb^{-1} of integrated luminosity". In: *LHCb-ANA-2013-096* (2016).

[165] Aaij, R. et al. “Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays”. In: *Phys. Rev. D* vol. 88 (2013), p. 052002. arXiv: 1307.2782 [hep-ex].

[166] Becirevic, D. and Tayduganov, A. “Impact of $B \to K_0^* \ell^+ \ell^-$ on the New Physics search in $B \to K^* \ell^+ \ell^-$ decay”. In: *Nucl. Phys.* vol. B868 (2013), pp. 368–382. arXiv: 1207.4004 [hep-ph].

[167] Aston, D. et al. “A Study of K- pi+ Scattering in the Reaction K- p —> K- pi+ n at 11-GeV/c”. In: *Nucl. Phys.* vol. B296 (1988), pp. 493–526.

[168] Dunwoodie, B. *Fits to $K\pi$ $I = 1/2$ S-wave Amplitude and Phase Data.* http: //www.slac.stanford.edu/~wmd/kpi_swave/kpi_swave_fit.note. 2013.

[169] Bohm, G. and Zech, G. “Chebyshev”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* vol. 748 (2014), pp. 1–6.

[170] Aaij, R. et al. “Angular analysis of the $B^0 \to K^* e^+ e^-$ decay in the low-q^2 region”. In: *JHEP* vol. 04 (2015), p. 064. arXiv: 1501.03038 [hep-ex].

[171] Bona, M. et al. “The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of Δm_s and BR($B \to \tau\nu_\tau$)”. In: *JHEP* vol. 0610 (2006). Updated results from summer 2019 (post EPS 2019 conference), p. 081. arXiv: hep-ph/0606167 [hep-ph].

[172] Lista, L. *Statistical Methods for Data Analysis in Particle Physics.* Vol. 909. Springer, 2016.

[173] Wilks, S. S. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *Annals Math. Statist.* vol. 9, no. 1 (1938), pp. 60–62.

[174] Karbach, T. M. and Schlupp, M. *Constraints on Yield Parameters in Extended Maximum Likelihood Fits.* 2012. arXiv: 1210.7141 [physics.data-an].

[175] Straub, D. M. “flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond”. In: (Oct. 2018). arXiv: 1810.08132 [hep-ph].

[176] Aaij, R. et al. “Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era”. In: (Aug. 2018). arXiv: 1808.08865 [hep-ex].

[177] “Measurement of lepton universality parameters in $B^+ \to K^+ \ell^+ \ell^-$ and $B^0 \to K^{*0} \ell^+ \ell^-$ decays”. In: (Dec. 2022). arXiv: 2212.09153 [hep-ex].

[178] Ball, P. and Braun, V. M. “Exclusive semileptonic and rare B meson decays in QCD”. In: *Phys. Rev.* vol. D58 (1998), p. 094016. arXiv: hep-ph/9805422 [hep-ph].

[179] Khodjamirian, A., Mannel, T., and Offen, N. “Form-factors from light-cone sum rules with B-meson distribution amplitudes”. In: *Phys. Rev.* vol. D75 (2007), p. 054013. arXiv: hep-ph/0611193 [hep-ph].

[180] Becirevic, D., Lubicz, V., and Mescia, F. “An Estimate of the $B \to K^* \gamma$ form factor”. In: *Nucl. Phys.* vol. B769 (2007), pp. 31–43. arXiv: hep-ph/0611295 `[hep-ph]`.

[181] Bobeth, C., Hiller, G., and Dyk, D. van. “The Benefits of $\bar{B} \to \bar{K}^* l^+ l^-$ Decays at Low Recoil”. In: *JHEP* vol. 07 (2010), p. 098. arXiv: 1006.5013 `[hep-ph]`.

[182] Horgan, R. R. et al. “Rare B decays using lattice QCD form factors”. In: *PoS* vol. LATTICE2014 (2015), p. 372. arXiv: 1501.00367 `[hep-lat]`.

[183] Aaij, R. et al. “Measurement of the phase difference between the short- and long-distance amplitudes in the $B^+ \to K^+ \mu^+ \mu^-$ decay”. In: *Eur. Phys. J.* vol. C77 (2017), p. 161. arXiv: 1612.06764 `[hep-ex]`.

[184] Egede, U., Patel, M., and Petridis, K. A. “Method for an unbinned measurement of the q^2 dependent decay amplitudes of $\overline{B}^0 \to K^{*0} \mu^+ \mu^-$ decays”. In: *JHEP* vol. 06 (2015), p. 084. arXiv: 1504.00574 `[hep-ph]`.

[185] Gubernari, N., Dyk, D. van, and Virto, J. “Non-local matrix elements in $B_{(s)} \to \{K^{(*)}, \phi\} \ell^+ \ell^-$”. In: *JHEP* vol. 02 (2021), p. 088. arXiv: 2011.09813 `[hep-ph]`.

[186] Boyd, C. G., Grinstein, B., and Lebed, R. F. “Model independent extraction of |V(cb)| using dispersion relations”. In: *Phys. Lett.* vol. B353 (1995), pp. 306–312. arXiv: hep-ph/9504235 `[hep-ph]`.

[187] Bourrely, C., Caprini, I., and Lellouch, L. “Model-independent description of $B \to \pi \ell \nu$ decays and a determination of $|V(ub)|$”. In: *Phys. Rev.* vol. D79 (2009). [Erratum: Phys. Rev.D82,099902(2010)], p. 013008. arXiv: 0807.2722 `[hep-ph]`.

[188] Döring, M., Meißner, U.-G., and Wang, W. “Chiral Dynamics and S-wave Contributions in Semileptonic B decays”. In: *JHEP* vol. 10 (2013), p. 011. arXiv: 1307.0947 `[hep-ph]`.

[189] Descotes-Genon, S., Khodjamirian, A., and Virto, J. “Light-cone sum rules for $B \to K\pi$ form factors and applications to rare decays”. In: *JHEP* vol. 12 (2019), p. 083. arXiv: 1908.02267 `[hep-ph]`.