Facultat de Física

Departament de Física Atòmica, Molecular i Nuclear

Institut de Física Corpuscular (UV-CSIC)

Measurement of the mixing-induced and CP-violating parameters of $B_s^0 \to \phi\gamma$ decays at LHCb

Carlos Sánchez Mayordomo

Directed by:

María Aránzazu Oyanguren Campos

Doctorado en Física

Mayo, 2018

Dra. MARÍA ARÁNZAZU OYANGUREN CAMPOS
Investigadora Ramón y Cajal

CERTIFICA:
Que la presente memoria, ≪**Measurement of the mixing-induced and CP-violating parameters of $B_s^0 \to \phi\gamma$ decays at LHCb**≫, ha sido realizada bajo mi dirección en el *Departament de Física Atòmica, Molecular i Nuclear* de la *Universitat de València*, y constituye el trabajo de tesis doctoral de Carlos Sánchez Mayordomo para optar al grado de Doctor en Física.

Y para que así conste, en cumplimiento con la legislación vigente, firmo el presente certificado.

Firmado
Dra. María Aránzazu Oyanguren Campos

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university.

Carlos Sánchez Mayordomo

Preface

This thesis presents the first measurement of the mixing-induced and CP-violating parameters of the $B_s^0 \to \phi\gamma$ decay ($\mathcal{A}^{\Delta}_{\phi\gamma}$, $\mathcal{S}_{\phi\gamma}$ and $\mathcal{C}_{\phi\gamma}$) appearing in the time-dependent decay rate, using the data collected by the LHCb experiment during the 2011 and 2012 data taking periods, accounting for a total of 3 fb^{-1} integrated luminosity.

The Standard Model (SM) of particle physics describes the properties of all known elementary particles and their electromagnetic, strong and weak interactions. However, an extension of this model is needed to include currently unexplained phenomena like dark matter, neutrino oscillations, or the matter-antimatter asymmetry of the universe. The theories that explain these phenomena introduce new particles that can produce other New Physics (NP) phenomena. For example, they could change the helicity of the photon in $b \to s\gamma$ transitions, leading to observable effects in the time-dependent decay rate of $B_s^0 \to \phi\gamma$ decays. The observables $\mathcal{A}^{\Delta}_{\phi\gamma}$ and $\mathcal{S}_{\phi\gamma}$ are sensitive to the ratio of helicity amplitudes (photon polarization). An overview of the theoretical framework is given in Chapter 1.

The Large Hadron Collider (LHC) is a circular proton collider located at CERN, designed to study particle physics at an unprecedent centre-of-mass energy of $\sqrt{s} = 14$ TeV. The LHCb experiment, designed to perform flavour physics measurements, is one of the main experiments of the LHC. An introduction to the LHCb experiment and the detector description is found in Chapter 2.

The reconstruction of neutral particles at LHCb and the selection of $B_s^0 \to \phi\gamma$ candidates is detailed in Chapter 3. Firstly, the work performed to improve the energy profile of photons is described, which is useful to distinguish between photons γ and pions π^0 within the LHCb framework. Secondly, the event selection criteria for $B_s^0 \to \phi(\to K^+K^-)\gamma$ and $B^0 \to K^{*0}(\to K^\pm\pi^\mp)\gamma$ decays is explained, including the procedure followed to optimize the selection. Lastly, the background subtraction procedure is presented, as well as the reconstruction of the decay time, the main variable in this analysis.

The measurement of $\mathcal{A}^{\Delta}_{\phi\gamma}$ is detailed in Chapter 4, where the time-dependent decay rate is fitted without distinguishing the flavour of the initial meson (untagged analysis). The $B^0 \to K^{*0}\gamma$ decay is used to control the decay time acceptance. The measurement of $\mathcal{S}_{\phi\gamma}$ and $\mathcal{C}_{\phi\gamma}$ is detailed in Chapter 5, where the flavour of the initial meson is identified (tagged analysis). A special emphasis is made in the description of the decay time resolution, which can affect the measurement. The implications of the result are also discussed. Finally, the conclusions are reported in Chapter 6.

Contents

1 Theoretical overview

This chapter gives an overview of the theoretical framework used in this work. The fundamental aspects of the Standard Model of particle physics are established in Section 1.1. The backbones of the weak interaction phenomenology are described in Sections 1.2, 1.3, and 1.4, detailing the CKM mechanism, the mixing of neutral mesons and the CP symmetry violation, respectively. Finally, Section 1.5 includes an overview of the photon polarization theory and the experimental measurements in the quest for new physics.

1.1 The Standard Model of particle physics

The Standard Model (SM) is a quantum field theory that describes properties of the elementary particles and the electromagnetic, weak and strong interactions. It has been experimentally validated on numerous occasions, making very accurate predictions. However, it is an incomplete theory since it does not include the gravity interaction, and does not explain the dark matter, the neutrino masses nor the matter-antimatter asymmetry of the universe.

The SM is based on the following symmetry group:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \tag{1.1}$$

which is the direct product of $SU(3)_C$ of color (C), representing the strong interaction characterized by the Quantum Chromodynamics (QCD) theory, and $SU(2)_L \otimes U(1)_Y$, the electroweak interaction of weak isospin and hypercharge, a unification of the weak and electromagnetic forces.

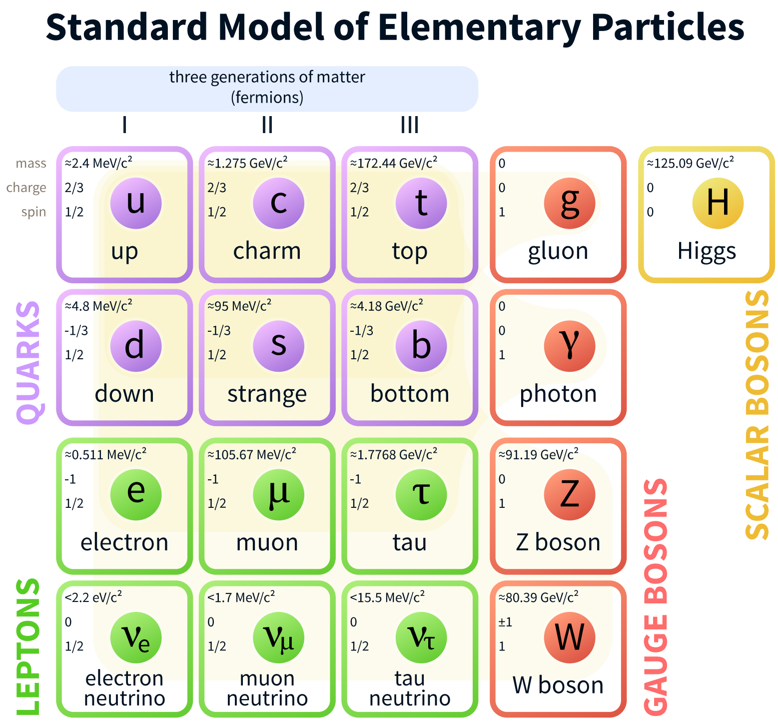


Figure 1.1: Illustration of the fundamental particles in the Standard Model and their properties [1].

Within the model, the whole space is filled with different types of fields, and the excitations of those fields are what we interpret as particles. A schematic view of the elementary particles and their properties is shown in Figure 1.1. None of them are known to have any underlying substructure. Two groups of particles can be distinguished: fermions, half-integer spin particles that form the building blocks of matter, and bosons, integer spin particles responsible of the interactions. The exchange of a boson between two particles is perceived as an interaction: the photon is the mediator of the electromagnetic force, the $W^{\pm}$ and Z^0 bosons are the mediators of the weak force, and eight gluons are responsible of the strong interaction. Along with them, the masses of the particles are generated by a scalar Higgs boson, discovered by the ATLAS and CMS collaborations in 2012 [2, 3].

Elementary fermions can be classified in six quarks[1] (u, d, c, s, t and b) and six leptons[1] (e, ν_e, μ, ν_μ, τ and ν_τ), where only the former feel the strong interaction. Quarks can be divided into up- and down-type quarks, while the leptons can be separated between electrically charged and neutral

[1] And their corresponding antiparticles.

particles, the latter known as neutrinos. They come up in three families or generations, sharing similar properties but differing in the masses. Most of the matter observed in the universe is formed from the first and lightest generation: the electron and electron neutrino, and the *up* and *down* quarks. Particles from the second and third generations are heavier and tend to rapidly decay into first-generation particles, hence they can only be observed when produced in a highly energetic environment, like in a particle collider.

From all the fundamental forces, the strong interaction is unique in the sense that the mediators of the force, the gluons, carry the same charge that they mediate, the color. Two colored particles feel a stronger interaction the further they are separated, until the energy becomes large enough to produce a new particle-antiparticle quark pair. As a consequence, quarks cannot be observed in isolation, but instead they are *glued* to other quarks to form colorless composite particles called hadrons. Depending on the number of quarks they are made of, hadrons are classified as mesons ($q\bar{q}$ with a quark and an antiquark with the corresponding anticolor charge) or baryons (qqq with three quarks combining all the colors). Exotic hadrons composed of four and five quarks [4, 5] have recently been observed [6–8]. When a quark is produced in a particle collider, it is quickly surrounded by other quarks to form hadrons in a process known as *hadronization*, being able to form particles like the beauty mesons B^0 ($d\bar{b}$) and B_s^0 ($s\bar{b}$) or the beauty baryon Λ_b^0 (udb) if the collision energy is enough to produce them.

1.1.1 Electroweak interaction

The electromagnetic and weak forces that we experience at low energies are unified into a single electroweak interaction at very high energies. The theory describing it was developed independently by Weinberg, Salam and Glashow [9–12], who used the following gauge symmetry group:

$$SU(2)_L \otimes U(1)_Y, \tag{1.2}$$

where the $SU(2)_L$ group gives rise to three massless bosons W_1^+, W_2^- and W_3^0 that only interact with non-zero weak isospin T_3 and left-handed chirality particles (or right-handed antiparticles). The L subindex of the group notation makes reference to this fact. The $U(1)_Y$ group generates another massless boson B^0 that acts on particles with weak hypercharge $Y = 2\,(Q - T_3)$, where Q represents the electric charge.

At energies below the electroweak scale, the symmetry of the theory is spontaneously broken by the Higgs mechanism, decoupling the electromagnetic and weak interactions. The complex Higgs doublet contributes with four additional degrees of freedom to the theory: three of them are absorbed by the electroweak bosons that become massive, while the fourth can be interpreted as the scalar Higgs particle [13, 14] whose field permeates all the space and gives mass to the particles that interact with it. The charged bosons W^+ and W^- of the weak interaction are a mixture of the W_1^+ and W_2^- electroweak bosons. Correspondingly, the neutral Z^0 and the photon γ are a mixture of W_3^0 and B^0, driven by the weak mixing angle θ_W:

$$\begin{aligned} W^{\pm} &= \frac{1}{\sqrt{2}}(W_1 \mp iW_2), \\ \begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} &= \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B^0 \\ W^0 \end{pmatrix}. \end{aligned} \tag{1.3}$$

The photon is the only boson that remains massless, thereby the electromagnetic interaction has an infinite range of action over electrically-charged particles. On the contrary, the fact that the weak interaction is mediated by three massive bosons ($W^{\pm}$ and Z^0) translates into a short-ranged force.

1.2 The CKM matrix

Experimental results indicate that the quark states participating in the weak interaction are different than the flavour eigenstates. The Cabibbo-Kobayashi-Maskawa (CKM) matrix is a 3×3 complex matrix that connects the flavour eigenstates of the quarks (d, s, b) with the weak eigenstates (d', s', b') [15, 16]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{\text{weak}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{flavour}}, \tag{1.4}$$

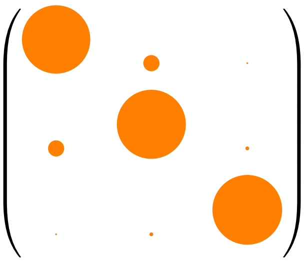
after choosing the flavour and weak eigenstates of the up-type quarks to be equal. The CKM mixing matrix contains all the flavour-changing and CP-violating couplings of the SM, where each matrix element V_{ij} represent the strength of the $j \to i$ process in the weak interaction, between a down- and an up-type quark. However, the SM does not predict the values of the elements; they have to be obtained from experiments. In the lepton sector, an equivalent matrix known as the Pontecorvo-Maki-Nagakawa-Sakata (PMNS) describes the mixing of neutrinos [17, 18].

The CKM matrix has to be unitary if weak universality holds. This requirement simplifies the matrix description down to four parameters: three angles and a complex phase, the latter being responsible of the CP violation in the SM. There are several ways to parametrize the CKM matrix, but one of the most popular choices is the Wolfenstein parametrization [19] in which the elements are expanded in a power series in the parameter λ:

$$\boldsymbol{V}_{\text{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{1.5}$$

where $\lambda \simeq 0.22$, $A \simeq 0.81$, $\rho \simeq 0.14$ and $\eta \simeq 0.35$ are determined experimentally [20]. This parametrization highlights its hierarchical structure:

$$|V_{\text{CKM}}|_{ij} = \begin{pmatrix} 0.97417 \pm 0.00021 & 0.2248 \pm 0.0006 & (4.09 \pm 0.39) \times 10^{-3} \\ 0.220 \pm 0.005 & 0.995 \pm 0.016 & (40.5 \pm 1.5) \times 10^{-3} \\ (8.2 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.009 \pm 0.031 \end{pmatrix}$$

$\sim$  ,

(1.6)

where the quarks clearly prefer to couple to the same family. Off-diagonal terms are order λ between the first and second generations, order λ^2 between second and third generations, and order λ^3 between first and third generations. Since the off-diagonal elements are significantly smaller than the diagonal ones, processes involving these terms are called Cabibbo-suppressed. The parameter η, related to the size of the CP violating effects, appear in the smallest elements.

The unitary of the matrix $V^\dagger_{\text{CKM}} V_{\text{CKM}} = 1$ imposes six conditions that can be represented as triangles in the complex plane, in which the sum of three products of specific off-diagonal terms must be zero. One of these conditions is referred as the unitarity triangle:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{1.7}$$

in which all terms are order λ^3. The invariance of the equation under rotations in the complex plane is exploited to define two of the vertices $(0,0)$ and $(0,1)$ in a normalized triangle, leaving its characterization to the third vertex $(\bar{\rho},\bar{\eta})$, where $\bar{\rho} \equiv \rho\left(1-\frac{\lambda^2}{2}\right)$ and $\bar{\eta} \equiv \eta\left(1-\frac{\lambda^2}{2}\right)$. The triangle angles are usually named α, β and γ [21]:

$$\begin{aligned}
\alpha &= \arg\left(-\frac{V_{tb}^* V_{td}}{V_{ub}^* V_{ud}}\right) = 92.0 \pm 1.2°, \\
\beta &= \arg\left(-\frac{V_{cb}^* V_{cd}}{V_{tb}^* V_{td}}\right) = 22.6 \pm 0.4°, \\
\gamma &= \arg\left(-\frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}}\right) = 65.4 \pm 1.1°.
\end{aligned} \tag{1.8}$$

Having only one phase responsible of the CP violation means that the area of all the possible triangles has to be equal (unitarity condition) and, if the weak universality holds, the sum of the three angles $\alpha+\beta+\gamma$ must be equal to 180°. The current status of the CKM triangle constraints is shown in Figure 1.2 [21].

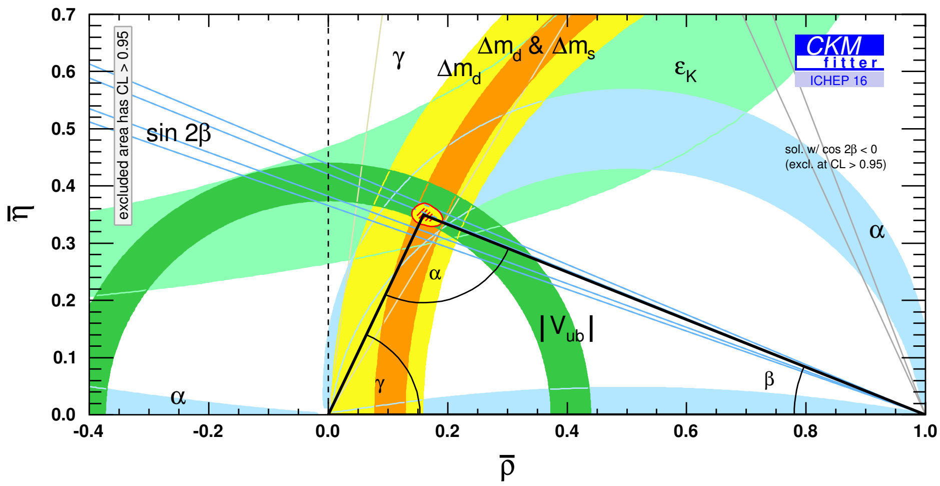


Figure 1.2: Global fit to the unitarity triangle, as provided by the CKMfitter group [21]. Colored bands represent the 1σ constraints of the different measurements.

1.3 Mixing of neutral mesons

Neutral mesons can oscillate over time between its particle and antiparticle states, in a process known as mixing. Figure 1.3 shows the transitions involved in the mixing process of the B_s^0 system mediated by the weak force, where a change of flavour $b \to s$ takes place.

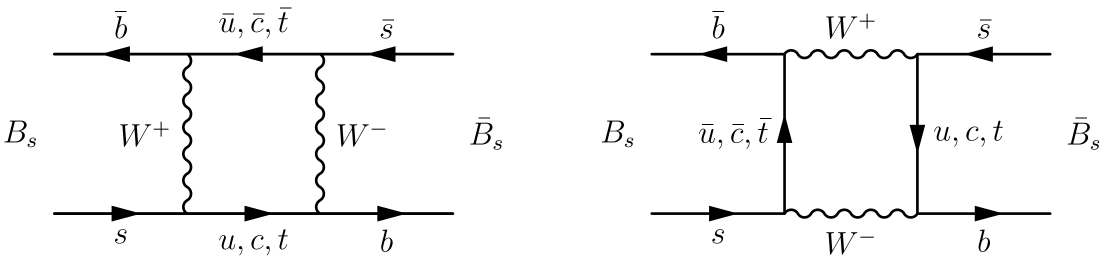


Figure 1.3: Feynman box diagrams contributing to the B_s^0-$\bar{B}_s^0$ mixing, mediated by the weak interaction.

This phenomenon occurs because the physical states[2] $|M_L\rangle$ and $|M_R\rangle$ are not the same as the states of definite flavour $|M^0\rangle$ and $|\bar{M}^0\rangle$. Instead, the former can be expressed as a linear combination of the latter:

$$\begin{aligned} |M_H\rangle &= p\,|M^0\rangle - q\,|\bar{M}^0\rangle\,, \\ |M_L\rangle &= p\,|M^0\rangle + q\,|\bar{M}^0\rangle\,, \end{aligned} \tag{1.9}$$

where p and q are complex numbers that satisfy $|p|^2 + |q|^2 = 1$ and are different for each meson system. The notation L and H makes reference to the *light* and *heavy* states, since the mass difference is significant in meson systems with beauty quark content, in contraposition to the neutral kaon system that is usually named *short* and *long* to indicate the difference in lifetime.

The meson mixing has been observed in all the four systems K^0, D^0, B^0 and B_s^0. K^0 oscillations were the first to be established [22, 23]. Two decades passed until the discovery of the B^0 oscillations by the ARGUS and UA1 collaborations in 1987 [24, 25]. Another two decades later, the B_s^0 oscillations were observed by the CDF and D0 experiments at Fermilab in

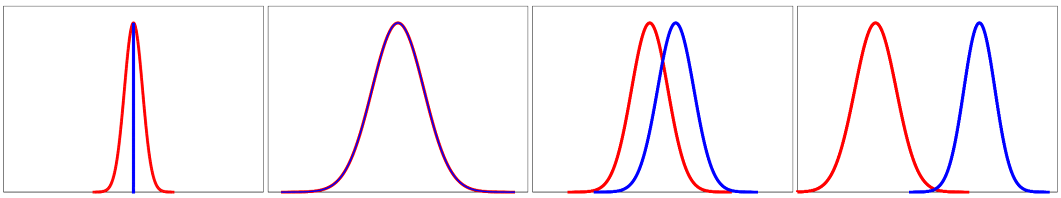
[2]Physical states are those of definite energy (and consequently, of definite mass and lifetime) under processes mediated by the weak interaction. Flavour states are those of definite energy under the strong and electromagnetic interactions. As a consequence, the weak force is the only interaction able to change the flavour.

2006 [26]. The oscillation in the D^0 system was evidenced by BaBar and Belle [27, 28] in 2007, and observed by LHCb from a single measurement [29].

Table 1.1 lists the mass and decay width differences between the light and heavy components of the four neutral meson systems. In Figure 1.3, a visual representation is shown. The K_S^0 and K_L^0 mesons have very similar masses despite the latter having a lifetime two orders of magnitude larger, thus they are centered at the same point with the long kaon represented as a narrower peak. The similar properties of the charmed D^0 components make the curves indistinguishable, a situation that is not true in the case of the B^0 system, in which the different masses can be appreciated but not the difference in lifetime. Finally, the B_s^0 case is unique in the sense that the difference can be appreaciated in both the masses and the lifetimes.

Table 1.1: Mixing parameters (mass and relative decay width differences) of the four oscillating neutral meson systems, averaged by the PDG [20].

System	Δm	$\Delta\Gamma/\bar{\Gamma}$
K^0	5.29 ± 0.01 ns^{-1}	2
D^0	9.5 ± 4.3 ns^{-1}	0.0129 ± 0.0016
B^0	0.5065 ± 0.0019 ps^{-1}	-0.002 ± 0.010
B_s^0	17.757 ± 0.021 ps^{-1}	0.130 ± 0.009



(a) K^0 (b) D^0 (c) B^0 (d) B_s^0

Figure 1.4: Representation of the light and heavy components of the different neutral meson systems. The x-axis represent the mass in arbitrary units, while the lifetime is related to the width of the curve, $\Gamma = 1/\tau$. The quantities have been modified for visualization purposes and are not scaled.

1.3.1 Time evolution of neutral mesons

Hamiltonian description

The time evolution of a quantum system can be described by the Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix}|P_1\rangle \\ |P_2\rangle\end{pmatrix} = \boldsymbol{H}\begin{pmatrix}|P_1\rangle \\ |P_2\rangle\end{pmatrix}, \tag{1.10}$$

where $\boldsymbol{H} = \boldsymbol{M} - \frac{i}{2}\boldsymbol{\Gamma}$ is the Hamiltonian of the system[3], where the mass matrix $\boldsymbol{M}$ represents the static component of the energy, and the decay matrix $\boldsymbol{\Gamma}$ is related to the particle lifetime, needed to account for final states outside of the system. The formalism has been simplified using the Weisskopf-Wigner approximation by only considering linear combinations of the two initial states, as well as timescales much higher than the strong interaction [30].

The mass eigenstates $|M_H\rangle$ and $|M_L\rangle$ are trivially described by a diagonal Hamiltonian since the states do not mix:

$$\boldsymbol{H}_{\text{mass}} = \begin{pmatrix}\omega_H & 0 \\ 0 & \omega_L\end{pmatrix} = \begin{pmatrix}m_H & 0 \\ 0 & m_L\end{pmatrix} - \frac{i}{2}\begin{pmatrix}\Gamma_H & 0 \\ 0 & \Gamma_L\end{pmatrix}. \tag{1.11}$$

The description of the flavour eigenstates $|M\rangle$ and $\left|\bar{M}\right\rangle$ is more complicated due to the mixing phenomena, requiring a general non-diagonal Hamiltonian:

$$\boldsymbol{H}_{\text{flavour}} = \begin{pmatrix}H_{11} & H_{12} \\ H_{21} & H_{22}\end{pmatrix} = \begin{pmatrix}m_{11} & m_{12} \\ m_{12}^* & m_{22}\end{pmatrix} - \frac{i}{2}\begin{pmatrix}\Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{22}\end{pmatrix}, \tag{1.12}$$

where the off-diagonal terms give rise to the mixing: H_{12} and H_{21} represent the $M \to \bar{M}$ and $\bar{M} \to M$ transitions, respectively. The Hamiltonian can be simplified from the general case if the CPT symmetry holds[4], where the condition

$$\langle M|\,\boldsymbol{H}\,|M\rangle = \left\langle \bar{M}\right|\boldsymbol{H}\left|\bar{M}\right\rangle \tag{1.13}$$

leads to the diagonal terms relations $m \equiv m_{11} = m_{22}$ and $\Gamma \equiv \Gamma_{11} = \Gamma_{22}$:

$$\boldsymbol{H}_{\text{flavour}} = \begin{pmatrix}m & m_{12} \\ m_{12}^* & m\end{pmatrix} - \frac{i}{2}\begin{pmatrix}\Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma\end{pmatrix}. \tag{1.14}$$

[3] A complex matrix can be separated into an hermitian and an antihermitian part, $\boldsymbol{H} = \boldsymbol{H_H} + \boldsymbol{H_{aH}}$, where $\boldsymbol{H_H} = \frac{1}{2}(\boldsymbol{H} + \boldsymbol{H}^\dagger)$ and $\boldsymbol{H_{aH}} = \frac{1}{2}(\boldsymbol{H} - \boldsymbol{H}^\dagger)$. This definition imply that the off-diagonal terms of any 2×2 complex matrix should obey $H^{21}_{(a)H} = H^{12\,*}_{(a)H}$.

[4] The formalism has also been developed in the case of CPT violation [31].

The parameters of the $\boldsymbol{H}_{\text{mass}}$ and $\boldsymbol{H}_{\text{flavour}}$ Hamiltonians can be related by inserting on Equation 1.13 the matrix $\boldsymbol{A} = \left(\begin{smallmatrix} p & -q \\ p & q \end{smallmatrix}\right)$ that links the states in the two bases, as shown in Equation 1.9. The resulting expression $\boldsymbol{H}_{\text{flavour}} = \boldsymbol{A}^{-1}\boldsymbol{H}_{\text{mass}}\boldsymbol{A}$ leads to:

$$\boldsymbol{H}_{\text{flavour}} = \frac{1}{2}\begin{pmatrix} \omega_H + \omega_L & -\frac{p}{q}(\omega_H - \omega_L), \\ -\frac{q}{p}(\omega_H - \omega_L) & \omega_H + \omega_L. \end{pmatrix}, \tag{1.15}$$

relating the physical observables with the fundamental parameters of the mixing theory of Equation 1.14:

$$\begin{aligned} \omega &\equiv \frac{\omega_H + \omega_L}{2} = m - \frac{i}{2}\Gamma, \\ \frac{p}{q} &= \sqrt{\frac{m_{12} - \frac{i}{2}\Gamma_{12}}{m_{12}^* - \frac{i}{2}\Gamma_{12}^*}}. \end{aligned} \tag{1.16}$$

Time evolution of flavour states

The time evolution equation (1.10) can be solved in the case of stationary states, like the mass eigenstates, which evolve following an exponential decay:

$$\begin{aligned} |M_H(t)\rangle &= e^{-(im_H + \Gamma_H/2)t}\,|M_H\rangle\,, \\ |M_L(t)\rangle &= e^{-(im_L + \Gamma_L/2)t}\,|M_L\rangle\,. \end{aligned} \tag{1.17}$$

However, a flavour eigenstate will evolve as a superposition of two states, as a consequence of the mixing. The change of basis is achieved by using Equation 1.9 that relates the flavour and the mass eigenstates $\boldsymbol{A} = \left(\begin{smallmatrix} p & -q \\ p & q \end{smallmatrix}\right)$:

$$\begin{pmatrix} |M(t)\rangle \\ |\bar{M}(t)\rangle \end{pmatrix} = \boldsymbol{A^{-1}\,U A} \begin{pmatrix} |M\rangle \\ |\bar{M}\rangle \end{pmatrix}, \tag{1.18}$$

where:

$$\begin{aligned} \boldsymbol{A^{-1}\,U A} &= \begin{pmatrix} \frac{1}{2p} & \frac{1}{2p} \\ -\frac{1}{2q} & \frac{1}{2q} \end{pmatrix} \begin{pmatrix} e^{-i\omega_H t} & 0 \\ 0 & e^{-i\omega_L t} \end{pmatrix} \begin{pmatrix} p & -q \\ p & q \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} e^{-i\omega_H t} + e^{-i\omega_L t} & \frac{q}{p}\left(-e^{-i\omega_H t} + e^{-i\omega_L t}\right) \\ \frac{p}{q}\left(-e^{-i\omega_H t} + e^{-i\omega_L t}\right) & e^{-i\omega_H t} + e^{-i\omega_L t} \end{pmatrix} \\ &= \begin{pmatrix} g_+(t) & \frac{q}{p} g_-(t) \\ \frac{p}{q} g_-(t) & g_+(t) \end{pmatrix}. \end{aligned} \tag{1.19}$$

Explicitly, the flavour states will evolve following the expression:

$$\begin{aligned} |M(t)\rangle &= g_+(t)\,|M\rangle + \frac{q}{p} g_-(t)\,|\bar{M}\rangle\,, \\ |\bar{M}(t)\rangle &= \frac{p}{q} g_-(t)\,|M\rangle + g_+(t)\,|\bar{M}\rangle\,, \end{aligned} \tag{1.20}$$

where $g_+(t)$ and $g_-(t)$ are the amplitude probabilities of no-oscillation and oscillation of the system, respectively:

$$g_\pm(t) = \frac{1}{2}\left(\pm e^{-i\omega_H t} + e^{-i\omega_L t}\right). \tag{1.21}$$

The time-dependent decay rate of $B_s^0 \to \phi\gamma$ decays is calculated using Equation 1.20, as it will be detailed in Section 1.5.1.

1.4 CP violation

1.4.1 Symmetries in particle physics

A physical system exhibits a symmetry if it remains invariant under a certain transformation. In classical physics the symmetries are continuous. Each continuous symmetry leads to the conservation law of a given quantity, associated with the generators of the symmetry group, as the Noether's theorem states [32]. For example, the energy is the conserved quantity under time translations, the momentum is the conserved quantity under space translations, and angular momentum is conserved under space rotations. In quantum physics, the symmetries are discrete, and they are satisfied if an operator commutes with the Hamiltonian. If a discrete symmetry holds, it leads to a conserved quantum number and a selection rule for the transitions between different states [30].

There are three basic discrete symmetries in particle physics: P, C and T, depicted in Figure 1.5. Parity transformation P inverts the spatial coordinates of a system, changing the chirality[5] of the particles. Charge conjugation C changes the sign of the quantum numbers, transforming particles

[5]The chirality is a fundamental quantum property of a particle, related to its spin. A particle has left- or right-handed chirality if it transforms under a left- or right-handed representation of the Poincaré group. For massless particles, chirality and helicity are equivalent; a particle is said to have right-handed helicity if the direction of the spin and momentum are the same. For massive particles, the helicity of a particle depends on the reference frame, but its chirality remains invariant.

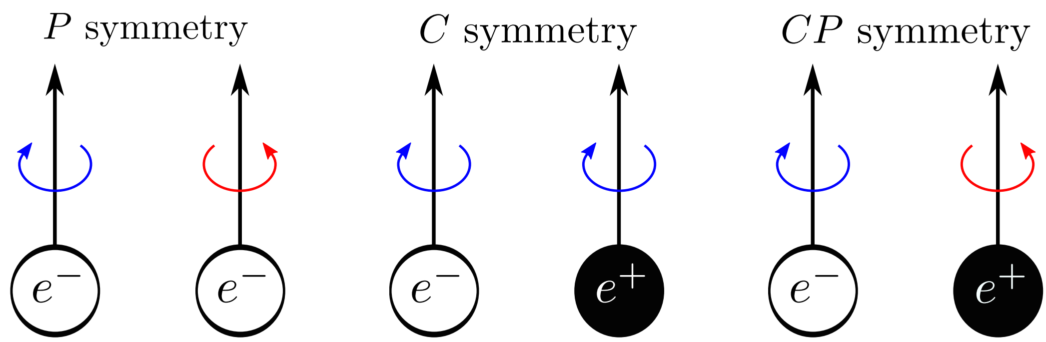


Figure 1.5: Symmetries in particle physics. The vertical arrow represents the momentum direction, while the curved arrows represent the spin. Particles and antiparticles are represented by white and black balls, respectively.

into antiparticles, but do not modify their chirality. Time transformation T reverse the time direction of a process. These discrete symmetries can be combined in any order to form the CPT symmetry, related to Lorentz invariance and therefore expected to hold in nature [33–35]. One of the consequences of the CPT invariance is that particles and antiparticles must have the same masses and lifetimes, a statement validated by experimental results with a high level of precision [20]. The CP symmetry is equivalent to the T symmetry if the CPT symmetry holds. It is the only fundamental discrete symmetry that is broken by a small amount, becoming one of the main targets of study of flavour physics experiments.

Symmetry violations

Processes mediated by the electromagnetic or strong interactions conserve each of the three basic symmetries, while the weak interaction violates all of them. Actually, the weak force violates maximally the P and C symmetries since it only couples to particles with left-handed chirality. The P violation was discovered with β decays in the Wu experiment in 1957 [36], while the CP violation was discovered by Cronin and Fitch with neutral kaons in 1964 [37]. In this experiment, neutral kaons were propagated up to a distance where only long kaons K^0_L could survive, but instead a small amount of $K^0_S \to \pi^0\pi^0$ decays were measured. Since $CP(K^0_S) \simeq +1$ and $CP(K^0_L) \simeq -1$, this experiment discovered the neutral meson mixing and the violation of the CP symmetry [38].

1.4.2 Types of CP violation

There are three types of CP violation in the Standard Model:

1. CP **violation in the decay.** This type of CP violation occurs when the probability of a particle P to decay to a final state f is not the same as the probability of the antiparticle $\bar{P}$ to decay to the CP-conjugated final state $\bar{f}$:
$$\mathcal{P}(P \to f) \neq \mathcal{P}(\bar{P} \to \bar{f}). \tag{1.22}$$

2. CP **violation in the mixing.** It occurs when the probability of the mixing $P^0 \to \bar{P}^0$ is not the same as the reverse process:
$$\mathcal{P}(P^0 \to \bar{P}^0) \neq \mathcal{P}(\bar{P}^0 \to P^0), \tag{1.23}$$
which will happen if
$$\left|\frac{q}{p}\right| \neq 1. \tag{1.24}$$
The beauty meson systems B^0 and B_s^0 do not exhibit CP violation in the mixing up to the per-mil level, with world averages $|q/p|_{B^0} = 1.0010 \pm 0.0008$ and $|q/p|_{B_s^0} = 1.0003 \pm 0.0014$ [39].

3. CP **violation in the interference between mixing and decay.** Also called mixing-induced CP violation, it occurs when both neutral mesons P^0 and $\bar{P}^0$ can decay into the same final state f, which can be reached in two ways: $P^0 \to f$ and $P^0 \to \bar{P}^0 \to f$. There is CP violation if the following statement holds:
$$\mathcal{P}(P^0(\to \bar{P}^0) \to f) \neq \mathcal{P}(\bar{P}^0(\to P^0) \to f). \tag{1.25}$$
The work developed in this thesis for the $B_s^0 \to \phi\gamma$ decay is based on this type of CP violation, as it is explained in Section 1.5.1.

The type (1) is also known as *direct* CP violation and can occur in both charged and neutral particles, while the types (2) and (3) are *indirect* in the sense that they involve the mixing of neutral mesons.

1.5 Photon polarization in $b \to s\gamma$ transitions

The radiative $b \to s\gamma$ transition is a Flavour Changing Neutral Current (FCNC), since it changes the flavour of the initial quark with no effective alteration of the electric charge. This transition occurs in the SM through one-loop $W^{\pm}$ exchanges, as illustrated in the Figure 1.6 (left) diagram. The exchange of new virtual particles in the loop, represented by Figure 1.6 (right), could modify the decay rates[6] of the processes and the helicity structure of the vertex.

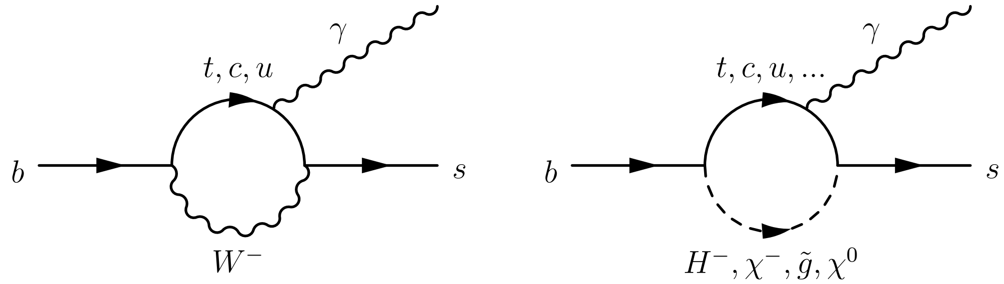


Figure 1.6: The $b \to s\gamma$ penguin diagram, mediated by SM particles (left) and new physics particles (right).

In the SM, the $b \to s\gamma$ photon is emitted predominantly with left-handed helicity (right-handed in $\bar{b} \to \bar{s}\gamma$), since the W boson couples exclusively to left-handed quarks. A significant right-handed contribution would be a clear signature of New Physics, as it is predicted in several extensions of the SM [41]. For a more detailed theoretical discussion in the context of the effective weak Hamiltonian, see Section 5.8.1.

1.5.1 Mixing-induced and CPV parameters in $B \to X\gamma$

In this section, the time-dependent decay rates of $B_s^0 \to \phi\gamma$ decays and flavour-specific $B^0 \to K^{*0}\gamma$ decays are derived. The obtained CPV parameters of the decay rate are then related with the photon polarization observable.

[6] Measurements of the inclusive branching ratios $\mathcal{B}(B \to X_s\gamma) = (3.4 \pm 0.2) \times 10^{-4}$ [20] and exclusive branching ratios $\mathcal{B}(B_s^0 \to \phi\gamma) = (3.5 \pm 0.4) \times 10^{-5}$ [40] show no significant deviation respect to the SM prediction.

$B_s^0 \to \phi\gamma$ decays

In $B_s^0 \to \phi\gamma$ decays, the final states $\phi\gamma_L$ and $\phi\gamma_R$ with left- and right-handed helicity photons are accessible from both B_s^0 and $\bar{B}_s^0$ initial states, allowing for CP violation in the interference between decays with and without mixing:

$$\begin{tikzcd}
& \phi\gamma_L & \\
B_s^0 \arrow[ur, "A_L"] \arrow[dr, "A_R"'] \arrow[rr, dashed, leftrightarrow, "\text{mixing}"] & & \bar{B}_s^0 \arrow[ul, "\bar{A}_L"'] \arrow[dl, "\bar{A}_R"] \\
& \phi\gamma_R &
\end{tikzcd}$$

The quantities next to the arrows represent the amplitude probabilities of the different processes:

$$\begin{aligned} A_{L,R} &= A(B_s^0 \to \phi\gamma_{L,R}), \\ \bar{A}_{L,R} &= A(\bar{B}_s^0 \to \phi\gamma_{L,R}). \end{aligned} \tag{1.26}$$

The decay rate evolution of an initial flavour state is constructed by adding incoherently the probabilities of the final states with left- and right-handed photons:

$$\begin{aligned} \Gamma_{B_s^0\to\phi\gamma}(t) &= |\langle\phi\gamma_L|B_s^0(t)\rangle|^2 + |\langle\phi\gamma_R|B_s^0(t)\rangle|^2, \\ \Gamma_{\bar{B}_s^0\to\phi\gamma}(t) &= |\langle\phi\gamma_L|\bar{B}_s^0(t)\rangle|^2 + |\langle\phi\gamma_R|\bar{B}_s^0(t)\rangle|^2. \end{aligned} \tag{1.27}$$

Using Equations 1.20 and 1.21, we obtain the decay rate for an initial B_s^0 or $\bar{B}_s^0$:

$$\begin{aligned} \Gamma_{B_s^0\to\phi\gamma}(t) \propto e^{-\Gamma_s t}[&\cosh(\Delta\Gamma_s t/2) - \mathcal{A}^\Delta \sinh(\Delta\Gamma_s t/2) \\ &+ C\cos(\Delta m_s t) - S\sin(\Delta m_s t)], \\ \Gamma_{\bar{B}_s^0\to\phi\gamma}(t) \propto e^{-\Gamma_s t}[&\cosh(\Delta\Gamma_s t/2) - \mathcal{A}^\Delta \sinh(\Delta\Gamma_s t/2) \\ &- C\cos(\Delta m_s t) + S\sin(\Delta m_s t)], \end{aligned} \tag{1.28}$$

where Γ_s is the decay width of B_s^0 mesons, $\Delta\Gamma_s$ the decay width difference, Δm_s the mass difference, and the coefficients $\mathcal{A}^\Delta$, C and S depend on the left and right complex amplitudes:

$$\begin{aligned} \mathcal{A}^\Delta &= \frac{2\,\mathrm{Re}[\frac{q}{p}\left(\bar{A}_L A_L^* + \bar{A}_R A_R^*\right)]}{|A_L|^2 + |A_R|^2 + |\bar{A}_L|^2 + |\bar{A}_R|^2}, \\ S &= \frac{2\,\mathrm{Im}[\frac{q}{p}\left(\bar{A}_L A_L^* + \bar{A}_R A_R^*\right)]}{|A_L|^2 + |A_R|^2 + |\bar{A}_L|^2 + |\bar{A}_R|^2}, \\ C &= \frac{|A_L|^2 + |A_R|^2 - |\bar{A}_L|^2 - |\bar{A}_R|^2}{|A_L|^2 + |A_R|^2 + |\bar{A}_L|^2 + |\bar{A}_R|^2}. \end{aligned} \tag{1.29}$$

The coefficients $\mathcal{A}^{\Delta}$ and S depend on the real and imaginary parts of the helicity amplitudes, while the coefficient C represent the direct CP violation. In the most general description, the amplitudes receive the contribution of the three up-type quarks $q = \{u, c, t\}$ with different strong δ^q and weak ϕ^q phases, leaving the $|A^q|$ coefficients real[7]:

$$\begin{aligned} A_{L,R} &= \sum_q |A^q_{L,R}| e^{i\delta^q_{L,R}} e^{i\phi^q_{L,R}}, \\ \bar{A}_{L,R} = CP(A_{R,L}) &= \xi \sum_q |A^q_{R,L}| e^{i\delta^q_{R,L}} e^{-i\phi^q_{R,L}}, \end{aligned} \tag{1.30}$$

where $\xi = +1$ is the CP eigenvalue of the ϕ meson. However, the contribution from the top quark dominates, and under such assumption (and $|q/p| = 1$) the parameters of Equation 1.29 can be simplified to [42]:

$$\begin{aligned} \mathcal{A}^{\Delta} &\simeq \frac{2|A_L||A_R|}{|A_L|^2 + |A_R|^2} \cos(\phi_s - \phi_L - \phi_R) \cos(\delta_L - \delta_R), \\ S &\simeq \frac{2|A_L||A_R|}{|A_L|^2 + |A_R|^2} \sin(\phi_s - \phi_L - \phi_R) \cos(\delta_L - \delta_R), \\ C &\simeq 0, \end{aligned} \tag{1.31}$$

where the amplitude factor can be written in terms of the photon polarization fraction $r = \frac{|A_{R,L}|}{|A_{L,R}|}$ as:

$$\frac{2|A_L||A_R|}{|A_L|^2 + |A_R|^2} = \frac{2r}{1 + r^2} = \sin(2 \arctan r), \tag{1.32}$$

a factor that can take values from 0 (where $r = 0$ or $r = \infty$) to 1 (where $r = 1$). The Standard Model predictions for the three parameters are [42]:

$$\begin{aligned} \mathcal{A}^{\Delta}_{\text{SM}} &= 0.047\,^{+0.029}_{-0.025}, \\ S_{\text{SM}} &= 0 \pm 0.002, \\ C_{\text{SM}} &\simeq 0.005 \pm 0.005, \end{aligned} \tag{1.33}$$

where the uncertainties come from the different theoretical contributions. An experimental value different to these predictions would be a signature of new physics.

[7] The real coefficients depend on the corresponding CKM matrix elements. The phases of the electromagnetic final states also contribute to the overall rate and are absorbed in the coefficients.

Flavour-specific $B^0 \to K^{*0}\gamma$ decays

In flavour-specific decays, *i.e.* where the initial state is well determinated from the final products, there is no interference between mixing and decay. In this case, the time-dependent decay rate is not sensitive to the photon polarization. This is the case of $B^0 \to K^{*0}\gamma$ decays, where the particle and antiparticle vector mesons decay to different states $K^{*0} \to K^+\pi^-$ and $\bar{K}^{*0} \to K^-\pi^+$, respectively:

$$\begin{tikzcd}
& K^{*0}\gamma_L & & \bar{K}^{*0}\gamma_L & \\
B^0 \arrow[ur, "A_L"] \arrow[dr, "A_R"'] \arrow[rrrr, leftrightarrow, dashed, "\text{mixing}"] & & & & \bar{B}^0 \arrow[ul, "\bar{A}_L"'] \arrow[dl, "\bar{A}_R"] \\
& K^{*0}\gamma_R & & \bar{K}^{*0}\gamma_R &
\end{tikzcd}$$

As in the previous case, the left- and right-handed amplitude probabilities are added incoherently:

$$\begin{aligned}
\Gamma_{B^0\to K^{*0}\gamma}(t) &= |\langle K^*\gamma_L | B^0(t)\rangle|^2 + |\langle K^*\gamma_R | B^0(t)\rangle|^2, \\
\Gamma_{\bar{B}^0\to \bar{K}^{*0}\gamma}(t) &= |\langle \bar{K}^*\gamma_L | \bar{B}^0(t)\rangle|^2 + |\langle \bar{K}^*\gamma_R | \bar{B}^0(t)\rangle|^2,
\end{aligned} \tag{1.34}$$

obtaining the following expression for the time-dependent decay rate:

$$\begin{aligned}
\Gamma_{B^0\to K^{*0}\gamma}(t) &\propto e^{-\Gamma_d\, t}\left[\cosh\left(\Delta\Gamma_d\, t/2\right) + C\cos\left(\Delta m_d\, t\right)\right], \\
\Gamma_{\bar{B}^0\to \bar{K}^{*0}\gamma}(t) &\propto e^{-\Gamma_d\, t}\left[\cosh\left(\Delta\Gamma_d\, t/2\right) - C\cos\left(\Delta m_d\, t\right)\right].
\end{aligned} \tag{1.35}$$

Note that the coefficients $\mathcal{A}^{\Delta}$ nor S do not appear in these expressions, and C corresponds to the direct CP violation:

$$C = \frac{|A_L|^2 + |A_R|^2 - |\bar{A}_L|^2 - |\bar{A}_R|^2}{|A_L|^2 + |A_R|^2 + |\bar{A}_L|^2 + |\bar{A}_R|^2} = 0, \tag{1.36}$$

which vanishes in the B^0 system. In addition, the $\cosh\left(\Delta\Gamma_d\, t/2\right)$ term is negligible since $\Delta\Gamma_d$ is not sizeable. Therefore, the time-dependent decay rate can be approximated as a simple exponential:

$$\Gamma_{B^0\to K^{*0}\gamma}(t) \simeq e^{-\Gamma_d\, t}, \tag{1.37}$$

with no dependency on the photon polarization.

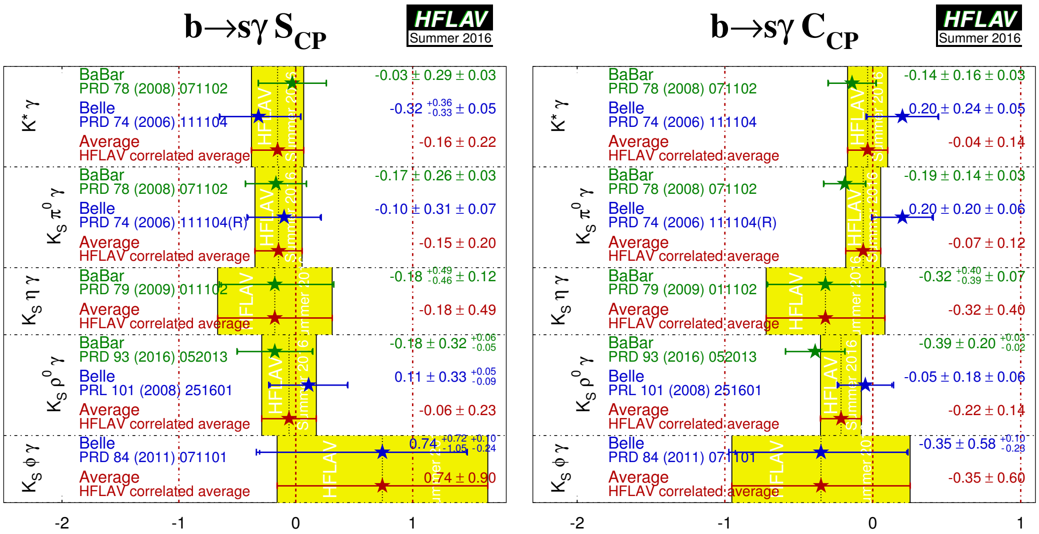
1.5.2 Experimental status

The measurement of the helicity amplitudes determine the photon polarization of the $b \to s\gamma$ transition. It can be measured indirectly through angular analyses and CP-violating parameters present in time-dependent decay rates. The LHCb collaboration observed for the first time that the photon in $b \to s\gamma$ is polarized, using $B^+ \to K^+\pi^-\pi^+\gamma$ decays with three hadrons in the final state [43]. The up-down asymmetry of the photon with respect to the decay plane defined by the three final hadrons allows to measure the photon polarization, which was determined to be different than zero. However, to obtain a single value for the photon polarization, the multiple resonances contributing to the $K^+\pi^-\pi^+$ system have to be studied [44, 45].

The best sensitivity to the photon polarization from a single measurement was obtained at LHCb from the angular analysis of $B \to K^{*0}e^-e^+$ decays at the very low dilepton invariant mass, q^2, in which the electron-positron pair is mediated by a virtual photon [46]. The amplitude of $B \to Ve^+e^-$ decays depends on two terms, $A_\parallel$ and $A_\perp$, that at $q^2 = 0$ correspond respectively to the parallel and transverse components of the virtual photon polarization respect to that of the vector meson [47]. Other methods to measure the photon polarization have been proposed, like angular analyses in $\Lambda_b \to \Lambda\gamma$ [48, 49] and $\Lambda_b \to pK\gamma$ [50] baryon decays.

On the other hand, the photon polarization can be probed in decays of the type $B \to f_{CP}\gamma$, where f_{CP} is a CP eigenstate accessible either from B or $\bar{B}$ mesons [41, 51]. The left- and right-handed helicity amplitudes can interfere between them to generate measurable CP-violating effects in the time-dependent decay rate.

The B factories Belle and BaBar measured the mixing-induced S_{CP} and C_{CP} observables of several exclusive $b \to s\gamma$ decays from B^0 initial states, extracted from the time-dependent CP asymmetry. The full list of measurements is shown in Figure 1.7. The ones with the less uncertainty come from the $B^0 \to K_s\pi^0\gamma$ decay [52, 53], where two different results are quoted: one in the K^{*0} invariant mass range (referred to as $K^{*0}\gamma$) and another in an inclusive region (referred to as $K_s\pi^0\gamma$). The results of the $K^{*0}\gamma$ resonance have a cleaner theoretical interpretation. In addition, the asymmetries were also measured in the $B^0 \to K_s\eta\gamma$ [54], $B^0 \to K_s\rho^0\gamma$ [55, 56] and



(a) S_{CP} (b) C_{CP}

Figure 1.7: The measured CP-violating parameters in $b \to s\gamma$ transitions from B^0 decays [39].

$B^0 \to K_s\phi\gamma$ [57] decays. Regarding $b \to d\gamma$ transitions, they were only measured in $B^0 \to \rho^0\gamma$ [58] decays. All the measurements are compatible with the null hypothesis, but the uncertainties still remain large.

In the B_s^0 system, the A^{Δ} observable is also measurable due to the sizeable mixing width $\Delta\Gamma_s = 0.090 \pm 0.005\ \mathrm{ps}^{-1}$ [42]. The work performed in this thesis has given rise to the first measurement of this observable in $B_s^0 \to \phi\gamma$ decays, $A^{\Delta}_{\phi\gamma} = -0.98\,^{+0.46}_{-0.52}\,^{+0.23}_{-0.20}$ [59, 60], from a fit to the untagged time-dependent decay rate (see Chapter 4). The $S_{\phi\gamma}$ and $C_{\phi\gamma}$ observables can also be measured from the time-dependent decay rate in an analysis with flavour tagging, which is also described in this thesis (see Chapter 5).

2 The LHCb experiment

This thesis is developed in the context of the LHCb experiment, whose description is presented in this chapter. First, the CERN accelerator complex is introduced in Section 2.1, followed by the generalities of the LHCb experiment and its most relevant results in Section 2.2. The LHCb detector and its subsystems are presented in detail in Section 2.3. Finally, the trigger system is explained in Section 2.4, and a brief description of the computing and software resources of the experiment is given in Section 2.5.

2.1 CERN and the LHC machine

The Large Hadron Collider (LHC) is the largest particle accelerator in the world, and the one operating at the highest energy [61]. It is a proton-proton circular collider lying in the former LEP tunnel, a 26.7 kilometers long ring situated 100 meters underground, hosted by the European Organization for Nuclear Research (CERN) laboratory, in the Franco-Swiss border near Geneva. Thousands of scientists from countries all around the world are collaborating to analyze the data produced by the collisions and push our knowledge about particle physics.

The LHC collides two beams of protons or ions in opposite directions, with velocities close to the speed of light. A sea of particles is produced in the collisions, which is recorded and later analyzed by the physicists. A beam carries 2808 nominal bunches, each containing 10^{11} protons. Bunches are separated by 25 ns so that the bunch frequency is 40 MHz. The CERN accelerator complex, depicted in Figure 2.1, produces and inserts the beams into the LHC tunnel, where they are accelerated by radiofrequency structures

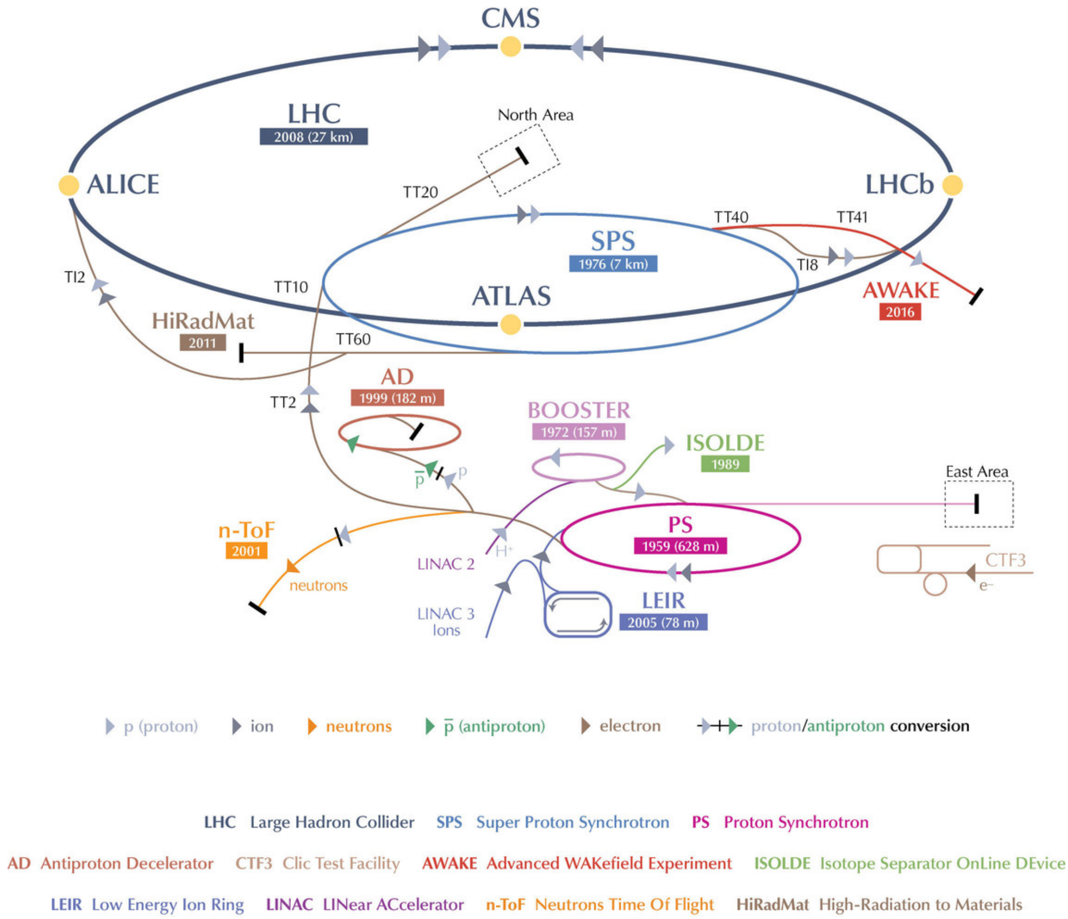


Figure 2.1: Scheme of the CERN accelerator complex [62].

and guided by several superconducting magnets. Figure 2.2 shows a view of the LHC tunnel.

Along the tunnel, there are a total of seven detectors aiming for different purposes: four main experiments (ATLAS, CMS, ALICE and LHCb) and three specialized experiments (TOTEM, LHCf and MoEDAL). ATLAS and CMS are general purpose experiments with the objective of searching for new physics at the TeV scale. ALICE study the phenomenology of heavy-ion collisions and some QCD mechanisms related to quark-gluon plasma. LHCb is specialized in flavour physics, studying the decays of heavy particles containing b- or c-quarks. LHCf, located next to the ATLAS detector, is designed to measure the production of particles parallel to the proton beam (forward direction), which is useful to study the particle showers produced by the collision of highly-energetic cosmic rays in the Earth's atmosphere. TOTEM is situated near CMS, and studies the total cross section of pp interactions, the proton structure and scattering, and several diffractive pro-

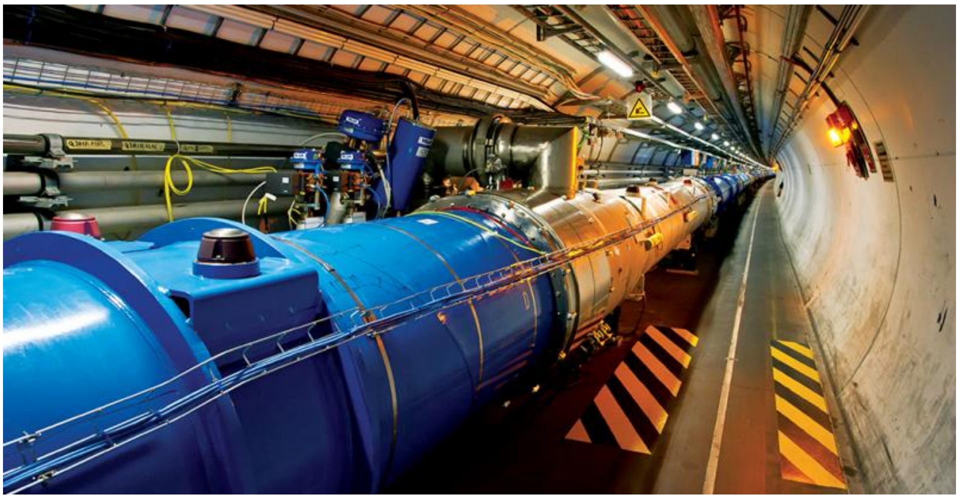


Figure 2.2: View inside the LHC tunnel [63].

cesses. MoEDAL is located close to the LHCb interaction point and started taking data in 2015. It pushes the discovery frontier by searching for new exotica phenomena, like magnetic monopoles or massive charged particles.

The first beams circulated in September 2008, but due to an electrical problem in the superconducting magnets a leakage of helium liquid damaged the accelerator and delayed the operation until 2010, a year devoted to commissioning. 2011 was the first data production year. After several production years, the LHC performs Long Shutdowns regularly to maintain or upgrade the machine; the data production period between these stops is called a *Run*. The 2011 and 2012 years correspond to the Run 1 period, and 2015, 2016, 2017 and 2018 to the Run 2 period. This thesis analyzes Run 1 data.

2.2 The LHCb experiment

The LHCb experiment is dedicated to the study of flavour physics, performing precision measurements of the CP symmetry violation and searching for rare decays of particles containing a beauty b or a charm c quark. With the current models we cannot explain the observed matter-antimatter asymmetry, given that the same quantity of matter and antimatter was produced at the beginning of the universe. New models beyond the SM predict new sources of CP violation and changes in the decay branching fractions, which would help to find an explanation of this anomaly [64].

The LHCb experiment is, at present, the most important source of heavy hadrons. A $b\bar{b}$ cross-section of around 500 μb, expected at the nominal center-of-mass energy of 14 TeV, allows for a massive production of not only B_d, B_s, and D mesons, but also charmed B_c and heavy baryons, which were not accessible in the previous e^+e^- factories Belle and BaBar.

Although the instantaneous luminosity of the LHC is $\mathcal{L}_{\text{LHC}} = 10^{34}\ \text{cm}^{-2}\,\text{s}^{-1}$, the LHCb detector is designed to work at a constant lower rate, $\mathcal{L}_{\text{LHCb}} = 2 \times 10^{32}\ \text{cm}^{-2}\,\text{s}^{-1}$, achieved by changing the angle of incidence of the beams at the interaction point as shown in Figure 2.3. This is a particularity of LHCb with respect to the other LHC experiments, and a useful feature for precision physics. Cleaner events are produced with few pp collisions per bunch crossing, facilitating the reconstruction and decreasing the radiation damage on the detector. LHCb recorded around 3 fb^{-1} during Run 1, and an additional 5 fb^{-1} is expected in Run 2. The integrated luminosity breakdown by year is shown at Figure 2.4.

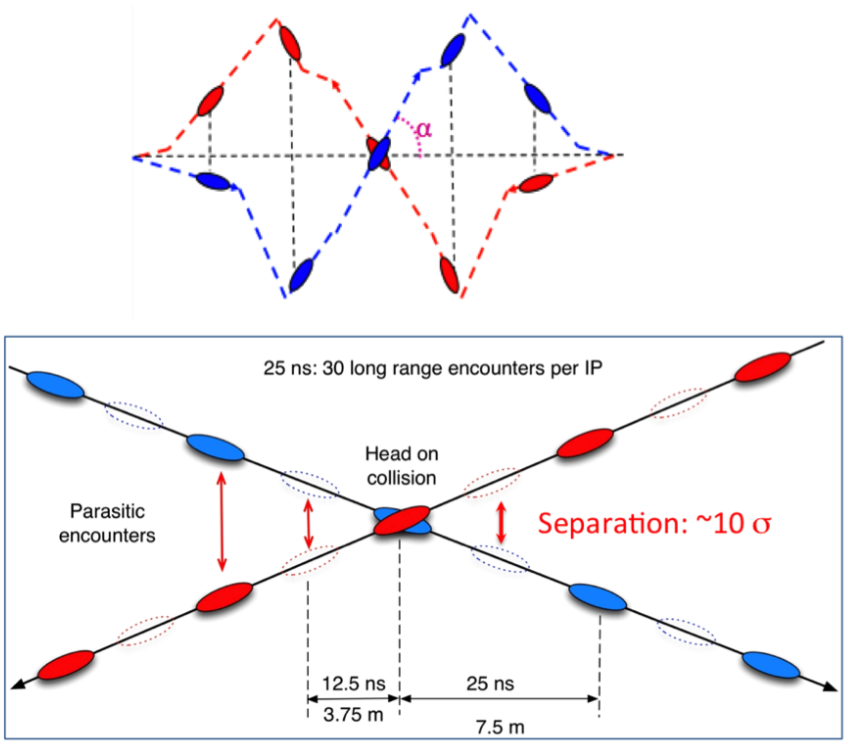


Figure 2.3: Crossing angle of the proton beams at the LHCb interaction point, to reduce the instantaneous luminosity [65].

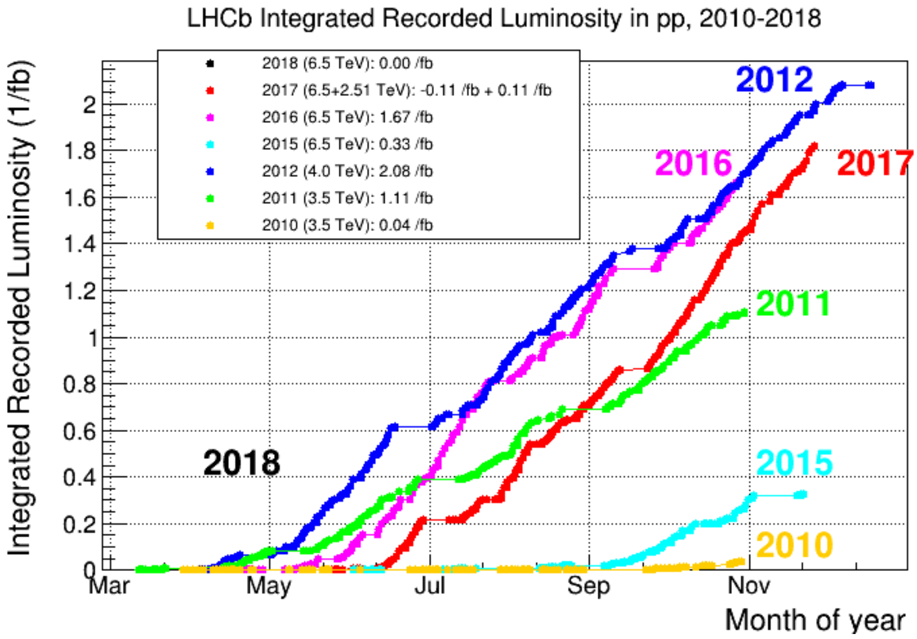


Figure 2.4: LHCb integrated luminosity in *pp* collisions, separated by year of data taking [66].

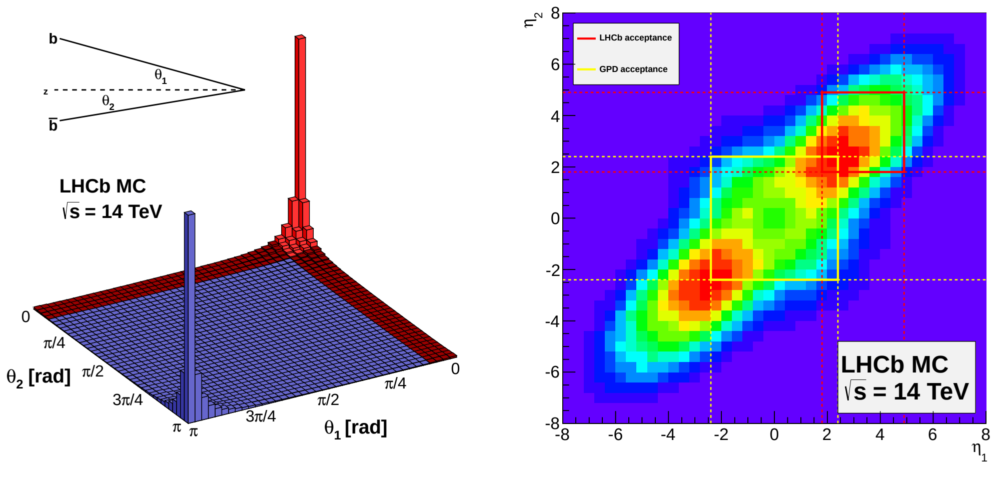
LHCb measurements

The LHCb experiment has already made relevant measurements and discoveries. Currently, the most promising evidence of new physics comes from semileptonic decays. In 2014, the LHCb measured the ratio of branching fractions $R(K) = \mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+e^+e^-)$, an observable related to the lepton flavour universality (LFU), with a result still compatible with the SM prediction within 2.6σ. Since then, several flavour measurements have shown slight deviations with respect the SM: the angular analyses of $B^0 \to K^{*0}\mu^+\mu^-$ [67–69] (3.4σ local deviation in a q^2 region) and $B_s^0 \to \phi\mu^+\mu^-$ decays [70] (3.0σ deviation), the measurement of $R(D^*) = \mathcal{B}(\bar{B}^0 \to D^{*-}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*-}\mu^-\bar{\nu}_\mu)$ [71, 72] (2.1σ, increasing to 4.0σ when combined with Belle and BaBar results), and the measurement of $R(J/\psi) = \mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$ [73] (2.0σ deviation).

Regarding new physics searches, $B_s^0 \to \mu^+\mu^-$ is a rare decay very sensitive to SM extensions. It was first observed in 2013 by LHCb [74], and subsequent measurements of the branching fraction have shown compatibility with the SM, even combining LHCb and CMS data [75, 76]. Another set of discoveries was made in the exotica sector. In 2015, the collaboration observed decays compatible with pentaquarks (bounded states) for the first time [8]. The $Z(4430)^-$ tetraquark state ($c\bar{c}d\bar{u}$) [7] was observed, and the quantum numbers of the $X(3872)$ tetraquark were determined [77]. Finally, regarding CP violation measurements, LHCb has observed CP violation for the first time in a B_s^0 decay [78], and in the charm [79] and baryon sectors [80], and has measured the ϕ_s phase in $B_s^0 \to J/\Psi\phi$ decays [81], among other measurements.

2.3 The LHCb detector

The LHCb detector [82] is a single-arm forward spectrometer optimized to perform flavour physics measurements. This layout was chosen since heavy $q\bar{q}$ pairs are produced predominantly in the forward or backward direction at high energies, as Figure 2.5 illustrates. It covers the region from 10 mrad to 250 mrad in the horizontal (non-bending) plane and to 300 mrad in the vertical (bending) plane, equivalent to a pseudorapidity between $2 < \eta < 5$. Figure 2.6 shows a side view of the detector with its components.



(a) (b)

Figure 2.5: (a) Angular distribution of the $b\bar{b}$ production. Only the forward region (in red) enters in the LHCb acceptance. (b) Comparison of the $b\bar{b}$ pseudorapidity acceptance for LHCb (red square) and other general purpose detectors (yellow square).

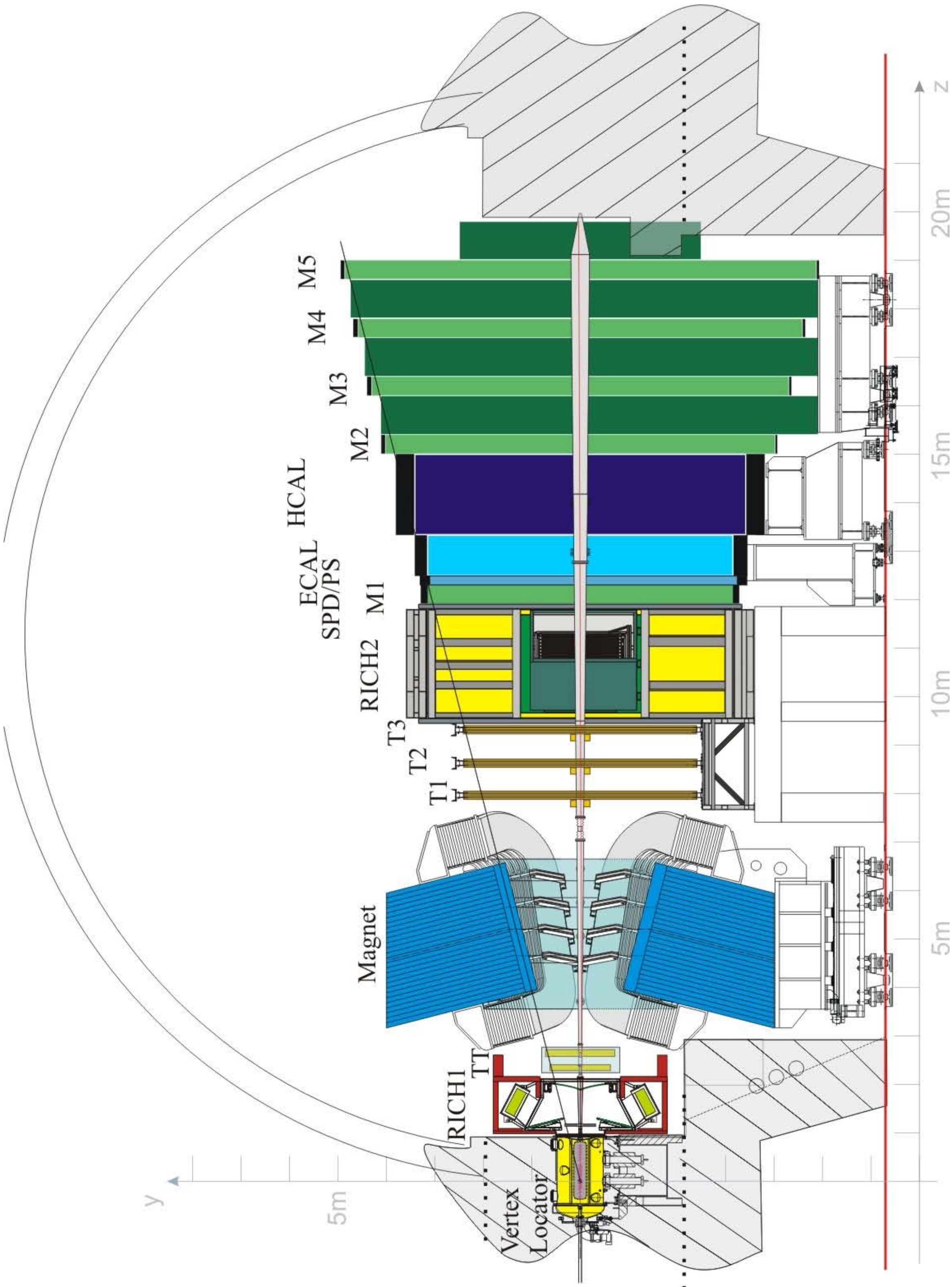


Figure 2.6: Side view of the LHCb detector [83].

The LHCb detector consists of a silicon vertex detector (VELO) surrounding the collision point, a silicon strip tracker (TT), a dipole magnet to bend charged particles, two RICH detectors for particle identification, three tracking planes (T1-T3) consisting of a silicon strip detector near the beampipe (IT) and a straw-tubes detector in the outer region (OT), a calorimter system consisting of a Scintillating Pad Detector (SPD), a pre-shower (PS), an electromangetic calorimeter (ECAL) and a hadronic calorimeter (HCAL), and a muon system with five stations (M1-M5).

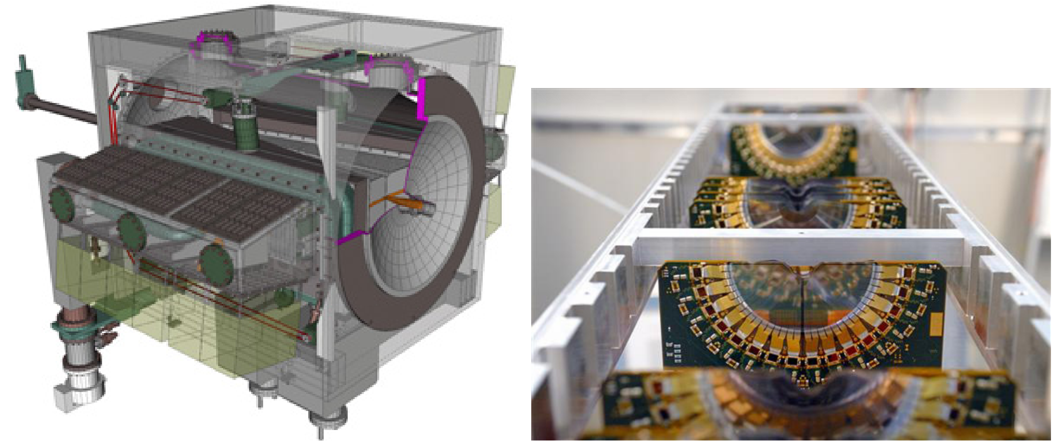
2.3.1 Tracking system

The purpose of the tracking system is to reconstruct the vertices and trajectories of charged tracks. It consists of the Vertex Locator (VELO) near the interaction point, the Tracker Turicensis (TT), and three Tracking Stations (T1-T3), with a dipole magnet between the last two.

The VELO is the closest detector to the interaction region, designed to determine precisely the trajectories of the tracks and identify displaced vertices. The system consists of a 1 meter long semistrip detector separated into two halves to cover the full azimuthal acceptance, each of them containing 21 half-circular stations as in Figure 2.7 (right), arranged along the beam direction. The stations are filled with two types of sensors placed in alternation, the ϕ- and R-sensors, for the determination of the azimuthal and radial coordinates, respectively. Each sensor is composed of 2 048 silicon strips with high radiation tolerance. The ϕ-sensor has radial strips with a stereo angle between 10° and 20°, while the R-sensor has azimuthal strips at constant radius. The whole setup is mounted inside a vaccum vessel as illustrated in Figure 2.7 (left), to keep the sensors isolated.

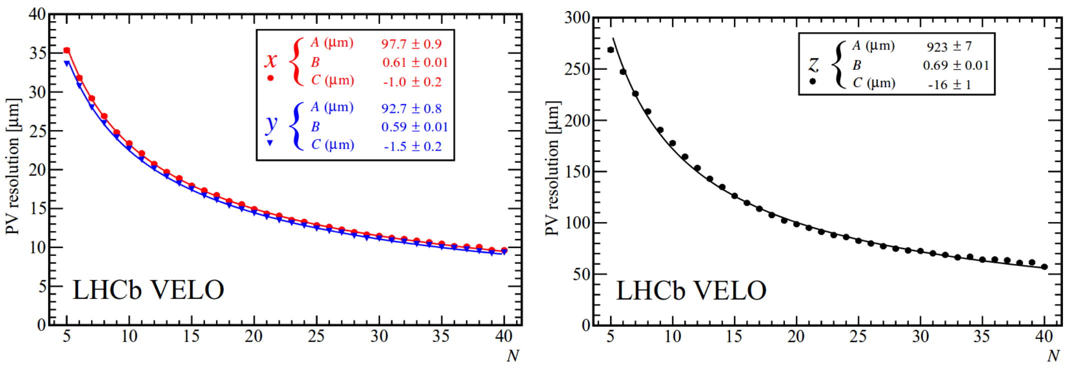
The VELO detector is designed to determine precisely the trajectories of the tracks and identify displaced vertices. This feature is essential to reconstruct impact parameters and decay lifetimes accurately. The spatial resolution of the Primary Vertex (PV) depends on the number of tracks used to reconstruct the vertex, as shown in Figure 2.8, with an average resolution of around 70 μm in the longitudinal axis and 13 μm in the transverse plane.

The dipole magnet bends the trajectory of charged particles, being essential to measure their momentum. An integrated magnetic field of 4 Tm for



(a) (b)

Figure 2.7: (a) Representation of the VELO vaccum vessel. (b) An array of VELO sensors, conforming one of the detector units.

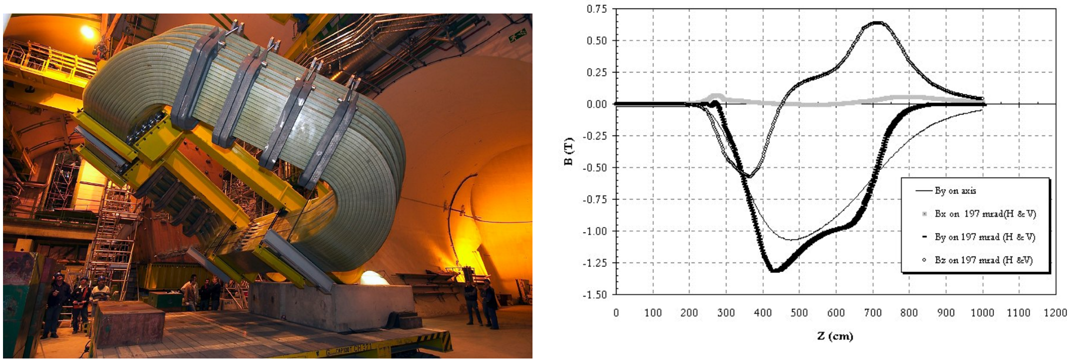


(a) PV_{xy} resolution (b) PV_z resolution

Figure 2.8: Primary vertex resolution on the transverse (a) and longitudinal (b) components, as a function of the number of tracks in 2011 data [84].

tracks crossing the magnet provides a momentum resolution of about 0.4%, for particles up to 200 GeV/c. Figure 2.9 (right) shows the three components of the magnetic field along the z-axis. In order to avoid any sort of systematic effect in CP measurements, the direction of the field magnet is inverted periodically. The magnet, shown in Figure 2.9 (left), is divided into two saddle-shaped coils placed symmetrically, each one 7.5 m long, 4.6 m wide, 2.5 m high and 27 tons of weight. The coils are arranged in 15 pancakes containing aluminum conductors, and the whole setup is put together in a 1450 tons steel frame.

Two different tracking detectors, the Tracker Turicensis (TT) and the Tracking Stations (T1-T3), are situated upstream and downstream the mag-



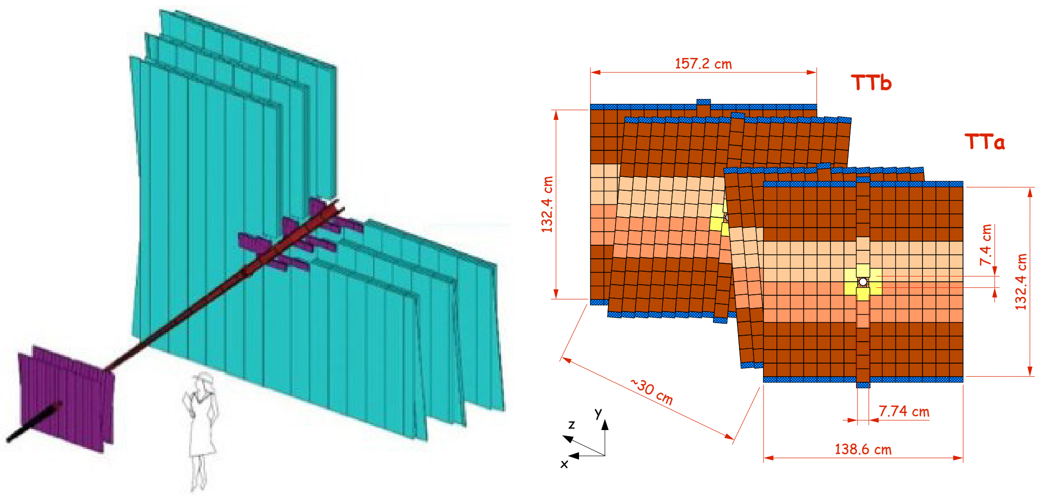
(a) (b)

Figure 2.9: (a) A picture of the magnet. (b) The three components of the magnetic field along the z-axis.

net, respectively. The Tracking Stations are divided into the Inner Tracker (IT) and the Outer Tracker (OT), covering the inner and outer acceptance of the detector. Figure 2.10 (left) shows the layout of the tracking system, while Figure 2.11 illustrates the IT and OT module display.

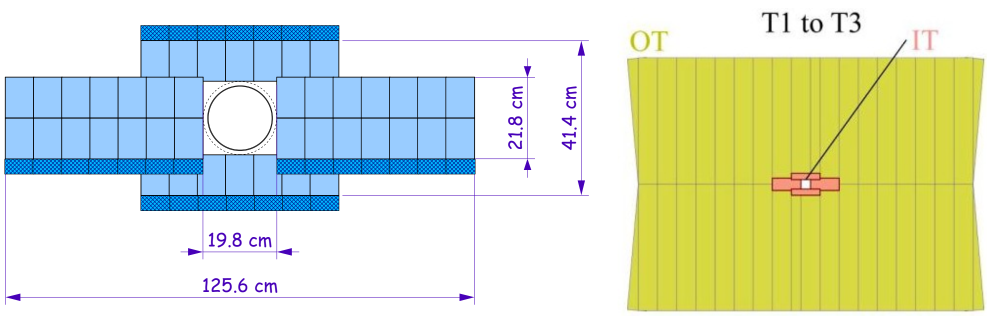
Both the TT and the IT are made of silicon microstrip sensors. The TT is a 150 cm wide and 130 cm high rectangular area located upstream the dipole magnet, covering the full acceptance. It consists of a single station, summing 145k strips and an active region of 8.4 m^2. The IT is a 120 cm wide and 40 cm high cross-shaped plane situated near the beam pipe in three stations located downstream the magnet, with 130k strips and an active area of 4.1 m^2. This area receives the highest particle density flux in LHCb, so the three separated stations help to reconstruct the track trajectories. Both silicon trackers are arranged in four layers of p-on-n strip sensors, with a strip pitch of about 200 μm. The first and last layers have vertical strips, while in the middle layers the strips are rotated 5°, as displayed in Figure 2.10 (right). The spatial resolution is around $\sim$50 μm for both trackers, which have similar readout electronics and monitoring systems.

The OT, unlike the other trackers, is a drift-time detector. It is composed of three stations surrounding the IT, covering the outer acceptance region. In this region the occupancy is lower (less than 10%), so due to budget reasons it was decided that this type of detector was enough to fulfill the experiment needs. Each station is arranged in four layers, with two monolayers of 64 drift



(a) (b)

Figure 2.10: (a) Placement of the Silicon Trackers (purple) and the Outer Tracker (blue) stations, displaying also the orientation of the modules. (b) The orientation and dimensions of the Tracker Turicensis layers.



(a) Inner Tracker (b) Outer Tracker

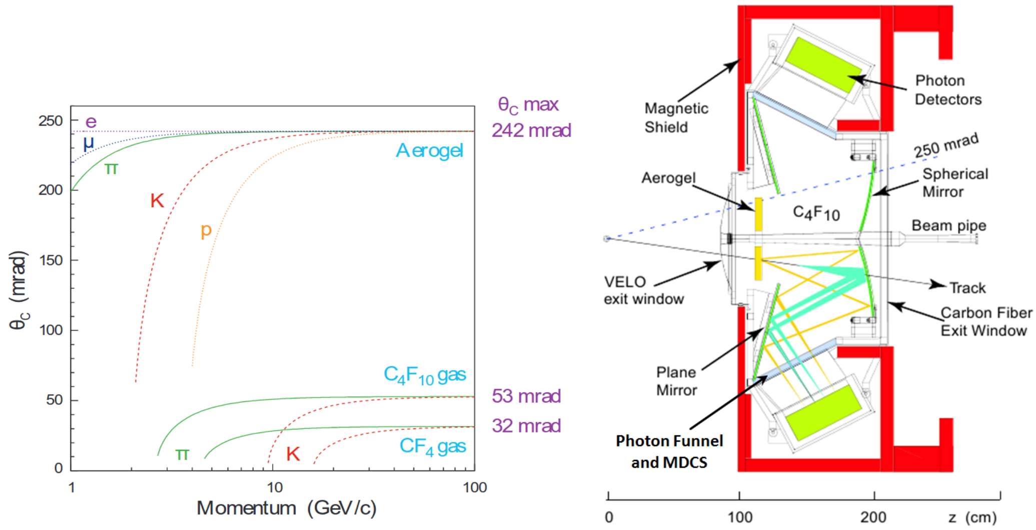
Figure 2.11: Layout of the Inner Tracker (a) and the Outer Tracker (b) modules.

straw tubes with a 4.9 mm inner diameter in each module. The orientation of the modules is the same as in the silicon trackers, with the middle layers rotated 5° and the others vertical. The tubes are filled with a gas mixture of 70% Argon and 30% CO_2, to guarantee a drift time below 50 ns and a resolution in the drift coordinate of about 200 μm. The total active area of a station is 5971 $\times$ 4850 mm^2.

2.3.2 RICH detectors

Two RICH (Ring Imaging Cherenkov) detectors are installed to identify charged particles, since being able to separate between pions, kaons and protons is a fundamental requirement to perform most LHCb measurements. One situated upstream and the other downstream the magnet, each detector is sensitive to a different momentum spectrum, covering in total a momentum range of 1-100 GeV/c. Figure 2.12 (left) shows the response of the RICH radiator gases in the momentum spectrum for different particles.

The RICH1 detector is located upstream the magnet, before the Tracker Turicensis. It is designed to identify particles in the low momentum spectrum, from 1 to 60 GeV/c, and covers the full LHCb acceptance. Silica aerogel and fluorobutane (C_4F_{10}) gas radiators are sensitive to this momentum range. The optical layout is disposed vertically as depicted in Figure 2.12 (right). The RICH2 detector is located downstream the magnet, between the last tracking station and the first muon station. The optical layout is disposed horizontally. It targets the high momentum spectrum, from 15



(a) (b)

Figure 2.12: (a) The dependency of the Cherenkov angle with the particle momentum for each of the gas radiators. Silica aerogel is better suited to identify low-momentum particles, while the fluorocarbons are more sensitive to high-momentum particles. (b) Schematic view of the RICH1 detector.

GeV/c to beyond 100 GeV/c, but covers only the acceptance where most high-momentum particles are produced, from 15 mrad to 120 mrad (horizontal plane) and 100 mrad (vertical plane). The fluorocarbon CF_4 is an ideal gas radiator to accomplish this task.

The Cherenkov light is guided outside the acceptance using a system of spherical and flat mirrors, arriving to an array of Hybrid Photon Detectors (HPDs) used to detect photons in the 200-600 nm wavelength range. The HPDs are vacuum photon detectors that guide the resulting photoelectrons onto a small silicon detector array, and they are surrounded by an iron shield and placed inside special metalic cylinders to minimize the effect of the external magnetic field. In later stages of the reconstruction, a global pattern-recognition algorithm compares the hit pattern detected by the HPDs with the patterns expected from the reconstructed tracks, and a probability is obtained. The efficiency for kaon identification is around 95% with an average pion misidentification of 5% in the 2-100 GeV/c momentum range.

2.3.3 Calorimeter system

The calorimeter system identifies photons, electrons and hadrons, and determines their energies and positions. It is divided into four subsystems: the Scintillator Pad Detector (SPD), the Preshower Detector (PS), the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL). The calorimeters measure the energy and position of the electromagnetic/hadronic showers, while the SPD/PS system improves the particle identification and the background rejection. The calorimeter system also selects high E_T particles in the hardware trigger, so a fast response was mandatory in its design. A scheme of the calorimeter system is displayed in Figure 2.13, showing also the interaction of the different particles at each subsystem.

Since the hit density is largely dependent on the distance to the beampipe, a variable lateral segmentation with different cell-size zones has been adopted, as shown in Figure 2.14. Three regions (inner, middle and outer) were chosen for the PS, SPD and ECAL, while the HCAL only has two regions (inner and outer) due to the different shower dimensions and trigger requirements. A projective geometry is followed by the subdetectors, which means that the

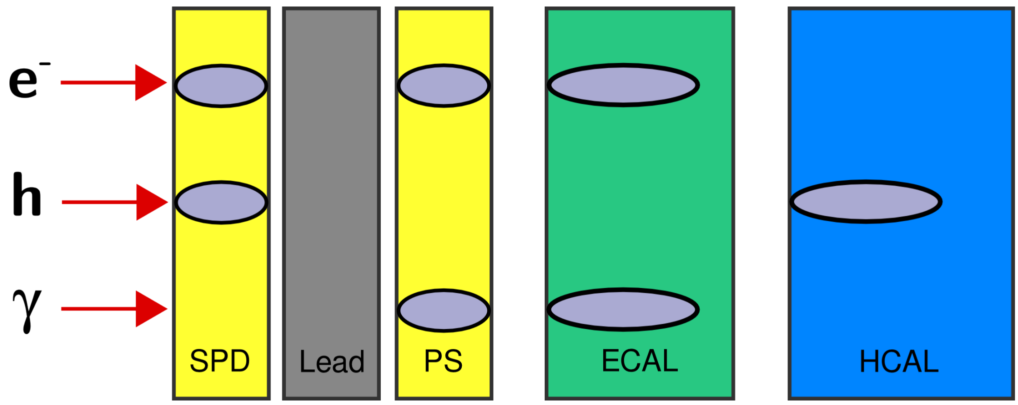


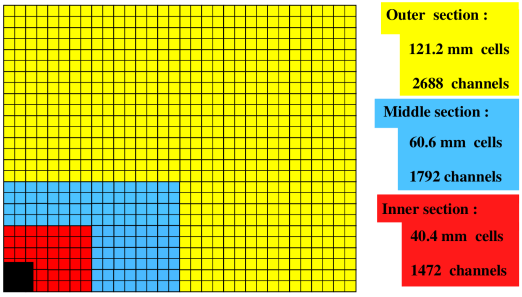
Figure 2.13: Schematic layout of the calorimeter system. The interaction of electrons, hadrons and photons is represented by grey areas if the particle produces a shower in the detector. Each particle leaves a different global signature in the calorimeter system.

cell size varies with the distance from the interaction point, so that cells of different subsystems cover the same acceptance. The PS is around 1.5% smaller than the ECAL, and the SPD is 0.45% smaller than the PS.

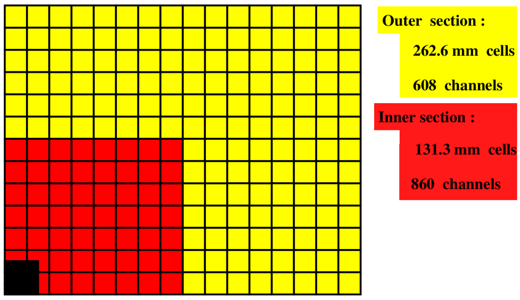
The SPD and PS are similar rectangular scintillator pads separated 56 mm, with a lead absorber of 15 mm thickness and 2.5 radiation length between them. The active area is 7.6 m wide and 6.2 m high, segmented in three regions to mimic the ECAL design. In all the four subsystems, the scintillation light is transmitted through wavelength-shifting fibres to a photomultiplier (PMT). However, the SPD/PS uses a multianode photomultiplier tubes (MAPMT), while in the ECAL and HCAL individual phototubes are set.

The SPD detects charged particles, which is useful to separate electrons from photons and neutral pions. Based on the deposited energy in the cells, a binary signal is fired to decide if there has been interaction or not. The misidentification probability is around 1%, coming from neutral particles leaving enough signal in the scintillator. The PS separates electrons from charged pions, making use of the electromagnetic shower dispersion.

The ECAL is a 7.8 m wide per 6.3 m high calorimeter located 12.5 m from the interaction point, covering from 25 mrad to the full outer acceptance of the detector to avoid the high radiation near the beampipe. It is shown in Figure 2.15. It follows a shashlik technology, alternating layers of scintillator



(a) ECAL



(b) HCAL

(c) ECAL modules

Figure 2.14: Lateral segmentation of the (a) ECAL and (b) HCAL calorimeters, showing the cell dimensions. Only one quarter of the detector is shown. (c) The three ECAL modules. The inner region has 9 cell modules of 4x4 cm^2, the middle region 4 cell modules of 6x6 cm^2 and the outer region one cell module of 12x12 cm^2 [85].

tiles and lead material (4 and 2 mm thick) crossed by wavelength-shifting fibres for the readout. A radiation length thickness of 25 X_0 is enough for the incident particles to deposit all their energy, which is a requirement for a correct reconstruction. Since the electromagnetic calorimeter is a fundamental part of the trigger decision, it was designed to provide a fast response at the expense of a modest energy resolution,

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\% \quad (E \text{ in GeV}), \tag{2.1}$$

where the first term is the statistical error and the second term takes into account systematic uncertainties from the calibration. In addition, the performance of the calorimeter cells can vary with time, having to regularly calibrate the cells to provide reliable measurements. For the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays, the invariant mass resolution of the parent b-meson pro-

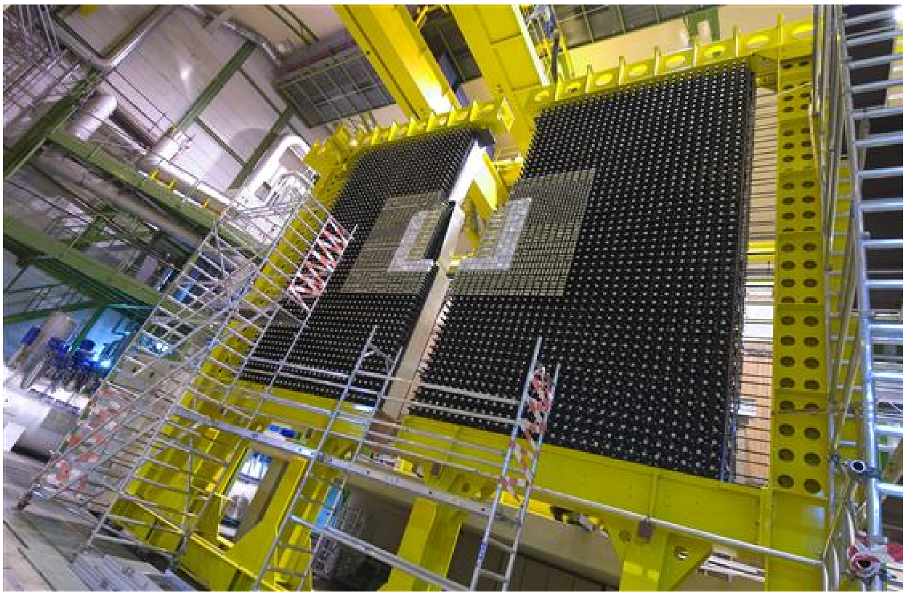


Figure 2.15: The electromagnetic calorimeter during its construction [86].

vided by the full reconstruction is about 90 MeV/c^2, with the main contribution coming from the resolution of the photon momentum. More information about the reconstruction of neutral particles is found in Section 3.1.

The HCAL is a 6.8 m wide per 8.4 m high and 1.65 m deep calorimeter located 13.33 m from the interaction point, consisting of iron absorbers and scintillator tiles arranged in parallel to the beampipe, with an active to passive material ratio of 0.18. It is segmented transversely into two regions, with the outer cells double sized in comparison with the inner cells. A radiation length thickness of 5.6 X_0 does not guarantee the deposit of the whole hadronic shower, but the design resolution requirements are much less stringent:

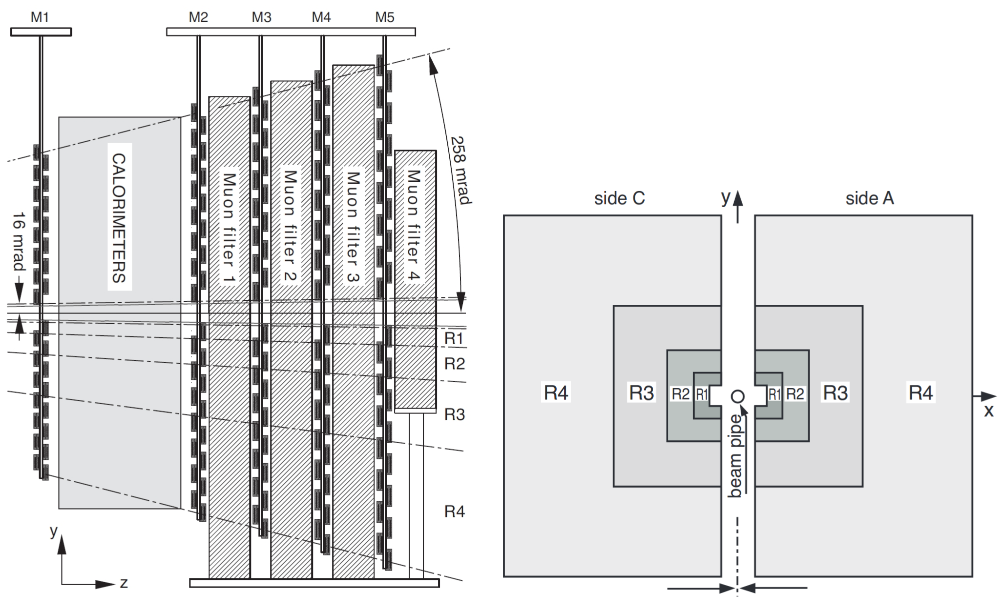
$$\frac{\sigma_E}{E} = \frac{65\%}{\sqrt{E}} \oplus 9\% \quad (E \text{ in GeV}). \tag{2.2}$$

The author of this thesis has been involved in the monitoring of the calorimeter system during the data taking, and has performed a task to improve the reconstruction of neutral particles (see Section 3.2).

2.3.4 Muon stations

The muon system identifies and measure the momentum of muons. It is the last subsystem of the LHCb detector, since muons can pass through the previous structures due to their low interaction probability. It covers an inner acceptance of 20 (16) mrad and an outer acceptance of 306 (258) mrad in the bending (non-bending) plane, and has 435 m^2 of total active area. Figure 2.16 illustrates the layout of the muon detector.

The system consists of five muon stations M1-M5 of rectangular shape placed along the beam axis, with a projective geometry to match the acceptance. M1 is placed upstream the calorimeters in order to obtain a better sensitivity on the transverse momentum, selecting high-p_T muon at trigger level. M2-M5 are located downstream the calorimeters, sandwiched between 80 cm thick iron absorbers to detect penetrating muons. M1-M3 are designed to have a high spatial resolution along the bending axis to reconstruct the track direction, and have a p_T resolution of 20%, whereas M4-M5 have limited spatial resolution but can detect penetrating muons.



(a) (b)

Figure 2.16: (a) Side view of the muon system, showing the position of the stations. (b) Layout of the stations in four regions R1-R4 with a scaling 1:2:4:8.

Moreover, each station is divided into four scaled regions R1-R4 to achieve uniform particle flux and channel occupancy, where R1 is the innermost and R4 the outermost. Multiwire proportional chambers (MWPC) are used as readout, except in the inner region of the M1 station, where triple-GEM detectors are used in order to deal with higher particle rates and ensure radiation protection. A total of 1380 chambers are contained in the detector.

The muon trigger is based on standalone muon track reconstruction and p_T measurement, and requires high-p_T muons and aligned hits on all five stations. Only muons with momentum higher than 6 GeV/c can traverse all the five stations without interaction, since the radiation length of the full system is equivalent to 20 interaction lengths. The efficiency for muon identification is 97%, with a 1-3% pion misidentification.

2.4 Trigger system

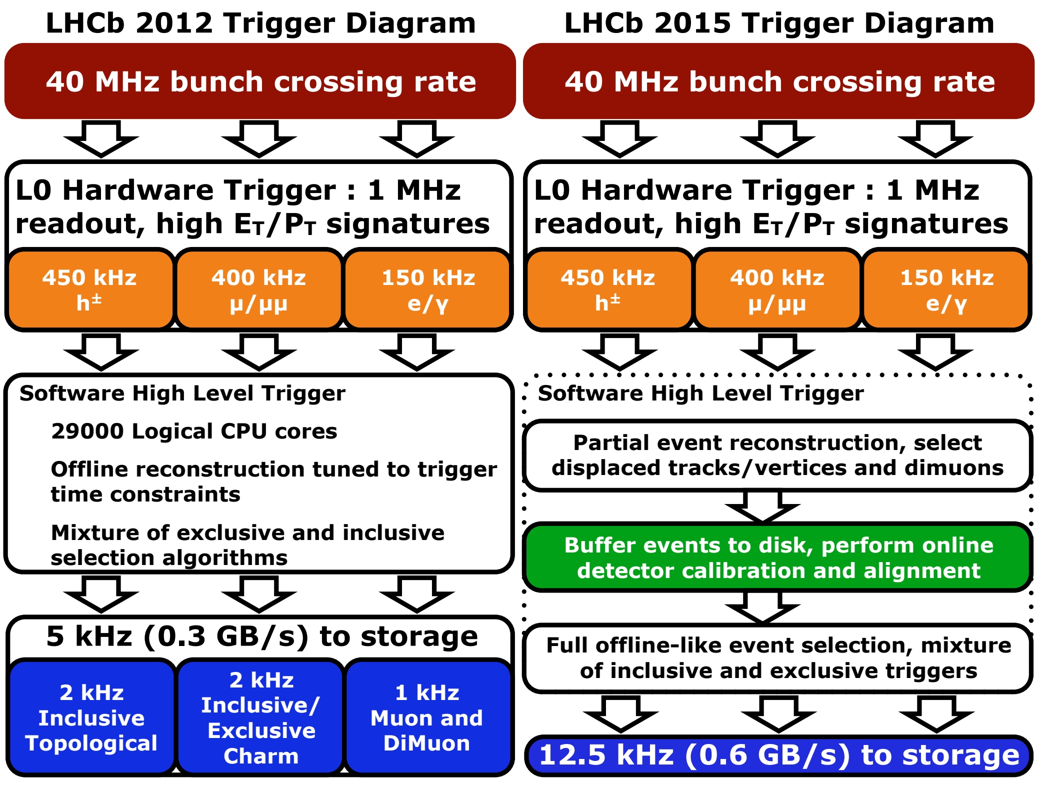
A trigger system is needed to reduce the huge amount of information produced in proton-proton collisions so it can be processed and stored. The LHCb trigger was designed in three stages: a Level-0 (L0) trigger implemented in hardware, and two high-level triggers (HLT1 and HLT2) that run in software. The total rate is reduced from the 40 MHz of the bunch crossing frequency at the interaction point to a few kHz after the trigger, corresponding to about 1 GB/s of information written in disk. Figure 2.17 shows a scheme of the trigger flow.

2.4.1 Level-0 trigger

The L0 trigger performs a basic selection of potentially interesting events by looking at the transverse energy of the particles, respect to the beam axis. It is implemented in custom boards fast enough to make a decision in under 4 μs, and reduces the bunch crossing frequency of the LHC to only 1 MHz at the output, which is the readout frequency of the detector.

It is divided in the calorimeter lines (`L0Hadron`, `L0Photon` and `L0Electron`) and the muon lines (`L0Muon`, `L0DiMuon`):

- `L0Hadron`: evaluates the highest E_T hadron candidate, constructed from the sum of the HCAL cluster energy and the energy of the ECAL



(a) Run 1 (b) Run 2

Figure 2.17: Scheme of the LHCb trigger layout, for Run 1 (a) and Run 2 (b) [87].

cluster compatible with the trajectory, if there is any.

- `L0Photon`: evaluates the highest E_T ECAL cluster, with hits in the PS and no hits in the SPD (in the cells in front of the ECAL cluster). Only the ECAL energy is used to reconstruct the photon candidate.
- `L0Electron`: same as `L0Photon`, but requiring hits in the SPD.
- `L0Muon`: looks for the muon with the highest p_T.
- `L0DiMuon`: looks for a pair of muons with the highest p_T product.

The transverse quantities are compared to a given threshold, that can be adjusted over time, to decide if it should pass the trigger. Events with more than 600 SPD hits are also removed (or more than 900 hits if fired by the di-muon line), to reject events with high number of tracks. In addition, the L0 Pile-Up (situated in the VELO) identifies the number of interactions, which is useful to retrieve the luminosity [88]. All this information arrives to the Decision Unit (DU) which evaluates the final decision.

2.4.2 High-level trigger

The High-level trigger processes the events selected by the Level-0 trigger. It is separated into two stages, HLT1 and HLT2, and runs on a dedicated computing cluster of 29 000 cores known as the Event Filter Farm (EFF).

HLT1

The HLT1 performs a partial reconstruction of the candidates by using information of the trackers and primary vertices, following an inclusive strategy to select the events. Specifically, it selects high-momentum tracks displaced from the interaction points. The rate is reduced to about 50 kHz in 2011, but was increased to 70 kHz in 2012 and to 150 kHz during Run 2. It takes advantage of the fast VELO reconstruction algorithm:

1. Reconstruct all the VELO track segments.
2. Check if they are compatible with hits in the muon chambers, or if the minimum IP respect to any interaction point is greater than a given threshold. If any of these conditions happen, the track is passed to the main tracker.
3. Match the segments against hits in the tracking stations, where a minimum p_T is required.
4. Only events with high-quality tracks are accepted.

HLT2

The HLT2 reconstructs completely the candidates, since the input rate is low enough. The output rate was reduced to 3 kHz in 2011, but was increased to 5 kHz in 2012 and 12.5 kHz during Run 2. Both inclusive (for generic decays) and exclusive (for specific decays) lines are constructed. Inclusive lines select common properties of some types of decays, for example the topological lines exploit general features of B decays, and the dimuon lines look for two reconstructed muons in $B \to J/\psi X$ decays with $J/\psi \to \mu^+\mu^-$.

Exclusive lines select a specific decay chain, with requirements similar to an offline selection. For example, the `HLT2Bs2PhiGamma` line selects event chains compatible with $B_s^0 \to \phi\gamma$ decays.

2.4.3 Trigger line categories

The trigger decisions can be splitted into three categories:

- Triggered on Signal (TOS): satisfied if the signal candidate alone fires the trigger decision.
- Triggered Independent of Signal (TIS): satisfied if the underlying event alone fires the trigger decision, so the signal candidate do not play any role.
- Triggered on Both (TOB): satisfied if the decision is neither TIS nor TOS. A mixture of the signal candidate chain and the underlying event is needed to fire the trigger.

2.5 Computing and software resources

The LHCb computing is based on a distributed multi-tier regional center model. The raw data produced after the trigger are transferred to the CERN Tier-0 centre which distribute it to the Tier-1 centers around the world, responsible of further processing of real data. Smaller Tier-2 centers help with the storage and processing of simulated data. Raw detector data has to be reconstructed and selected further in a process called *stripping*, producing the files that are actually accessible to the user to produce analysis tuples.

LHCb software

The LHCb framework was designed to run a chain of independent steps either to process the data or generate simulated events [89]. It is based on Gaudi [90], an object-oriented architecture written in `C++`, although the configuration can be handled by Python scripts. The main projects are detailed in the following.

- **Gauss** [91, 92] is used to generate simulated data. The simulation of pp collisions is handled by Pythia [93], creating the initial position and momentum of the particles. These particles are then decayed to interesting channels for the LHCb physics program by a customized EvtGen package [94], while the final state radiation is handled by PHOTOS [95]. Other event generators are also implemented for special productions. Finally, the interaction of the particles with the detector is simulated by Geant4 [96], reading from a database the detector geometry and the material data.

- **Boole** [97] simulates the detector response and the digitization of energy depositions. Although this step is considered part of the simulation, it is convenient to separate it since the response of the detectors can be improved with more data.

- **Moore** [98] runs the trigger. A copy of this software is installed on all the online farm computers to apply the HLT trigger in data. In Monte Carlo simulated data, the L0-trigger also needs to be emulated.

- **Brunel** [99] performs the event reconstruction. It takes the raw information (hits, calorimetric deposits) and transform it into clusters and tracks.

- **DaVinci** [100] is the software dedicated to analysis tasks and to process the *stripping*. The users can create tuples which are then used for offline analysis.

During this thesis, the author has been responsible for the preparation of Monte Carlo productions for the Rare Decays working group, and for the preparation of filtered Monte Carlo productions for the whole collaboration.

Analysis software

In addition, other software tools are mentioned throughout this work:

- **ROOT** [101]: a framework designed to process large amount of data, widely used by the high energy physics community. The reading and analysis of the data tuples and the creation of plots are performed with this package.

- **RooFit** [102]: included inside the ROOT framework, it is a toolkit for complex data analysis. The unbinned fits to the decay time distributions are programmed with RooFit, with the help of Minuit [103] to perform the minimization.
- **sPlot** [104]: a statistical tool to unfold data distributions. In this work, it is used to distinguish between the different components of the reconstructed mass distributions, helping in the subtraction of background events.

3

Reconstruction and selection of $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ candidates

This chapter describes the strategy to select $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ candidates. Section 3.1 introduces the calibration of the calorimeter system, and the reconstruction and the identification of neutral particles, while Section 3.2 details the work performed by the author of this thesis to improve the transverse energy profile of photons, which is used by the LHCb reconstruction software. Section 3.3 details the event selection criteria (trigger, stripping and offline selections) for $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays. Section 3.4 focuses on the background subtraction procedure through a fit to the reconstructed B mass distribution. Finally, Section 3.5 discusses about the reconstruction of the decay time variable.

3.1 Reconstruction of neutral particles at LHCb

The photon γ and neutral pion π^0 reconstruction at LHCb has been detailed in [105]. To properly reconstruct the energy of the particles, a correct calibration of the whole calorimeter system is fundamental [82]. In the SPD and PS subdetectors, the cell energy thresholds are adjusted regularly during the run periods to ensure a uniform efficiency. In the ECAL, the cell-to-cell intercalibration is performed initially by the LED system, whose light is transported to the photomultipliers. During the run periods, the intercalibration is adjusted by fitting the invariant mass distribution of $\pi^0 \to \gamma\gamma$ decays, as depicted in Figure 3.1, with a precision that reaches the 2% level.

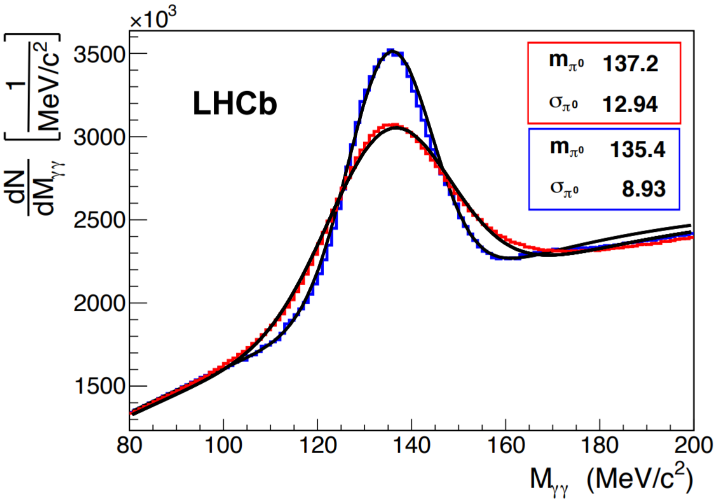


Figure 3.1: Calibration of the ECAL energy with the invariant mass distribution of $\pi^0 \to \gamma\gamma$ candidates [82]. The blue (red) curve corresponds to the fit after (before) the calibration.

The energy deposits of the ECAL cells are grouped together to form 3×3 cluster patterns around the energy local maxima, following a cellular automaton algorithm [106]. A cell that is present in several clusters shares its energy between the clusters according to the energy profile of single photons, in an iterative procedure that converges rapidly. Clusters are characterized by the total energy, the energy-weighted position and their spread in two dimensions.

A cluster is considered to be a neutral particle candidate if it is not compatible with any charged track in the event. To accomplish this, all the tracks in the event are extrapolated to the calorimeter and a two-dimensional $\chi^2_{\rm 2D}$ is constructed to evaluate the compatibility with the cluster position. Only isolated clusters with $\chi^2_{\rm 2D} > 4$ are selected as neutral particle candidates.

3.1.1 Photon reconstruction

Photons can be reconstructed as *converted* or *unconverted*. The former are photons converted into an electron-positron pair before the lead plate situated between the SPD and the PS, due to the interaction with the material upstream the calorimeter. They are reconstructed from two charged tracks, and have the advantage of a better momentum resolution compared to unconverted photons, typically a factor three lower. However, even tough 20%

of the photons are converted, only one third of them can be reconstructed, suffering also from bremsstrahlung effects. This section focuses on the reconstruction of unconverted photons.

The photon energy is calculated from the total energy of the ECAL cluster and the corresponding PS energy deposit. The energy is corrected for leakages using a simulation of isolated photons, with an observed E_γ resolution of:

$$\frac{\sigma_E}{E} = \frac{10.2 \pm 0.3\%}{\sqrt{E}} \oplus (1.6 \pm 0.1)\%. \tag{3.1}$$

The photon direction is defined from an assumed origin (the interaction point) to the 3D energy-weighted barycenter of the ECAL cluster. The barycenter position is corrected in the longitudinal direction, for the penetration depth of the photon, and in the transverse direction, to account for non-linear effects in the shower shape. All the mentioned corrections are implemented by default in the reconstruction software of LHCb.

Momentum post-calibration

However, a bias is observed in the photon momentum of radiative decays such as $B_s^0 \to \phi\gamma$ after the reconstruction, at the percent level, producing a mismatch between the reconstructed B mass peak and its expected value. This happens because the various contributions to the calorimetric energy are tuned using low-energy photons, and the balance between the different terms induces non-linear effects that are relevant in the reconstruction of high-energy photons.

The photon momentum bias is corrected by a dedicated post-calibration algorithm, as mentioned in [107], applying different correction factors that depend on the ECAL region (inner, middle, outer) and the type of photon (converted, unconverted). The correction factors are tuned so that the invariant mass of $B^0 \to K^{*0}\gamma$ candidates peaks in the expected B^0 mass [20]. The post-calibrated momentum is propagated to all the relevant kinematical quantities, including the mass and the decay time of the B candidate. The decay time resolution bias is then reduced from 20 fs, without postcalibrating the photon momentum, to a residual 5 fs after postcalibration.

3.1.2 Neutral pion reconstruction

Neutral pions with low transverse momenta can be reconstructed from a pair of photons; they are called *resolved*-π^0. However, due to the ECAL cell granularity, pions with transverse momenta greater than 2 GeV/c cannot be resolved from two well-distinguished clusters; they are called *merged*-π^0. Figure 3.2 shows the p_T distribution of resolved and merged candidates.

An algorithm was designed to reconstruct merged-π^0, splitting each single ECAL cluster into two 3x3 subclusters built around the two cells with the highest energy deposit, as depicted in Figure 3.3. The energy of the common cells is then distributed between the two subclusters according to the transverse energy profile of single photons. The process is iterative and fast converging, and its performance was tested using $D^0 \to K^-\pi^+\pi^0$ decays with an observed D^0 mass resolution around 20 MeV/c^2 and 30 MeV/c^2 for resolved and merged π^0 candidates, respectively [82].

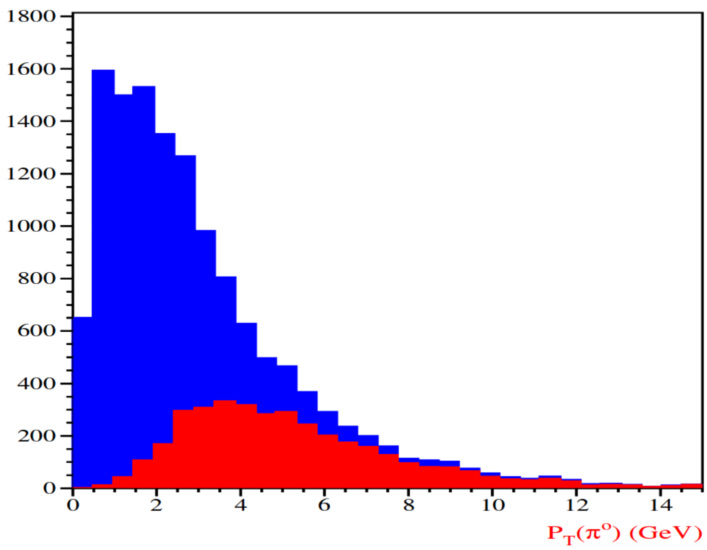


Figure 3.2: Transverse momentum distribution of neutral pions in $B^0 \to \pi^+\pi^-\pi^0$ simulated decays [105]. Resolved and merged pions are represented by the blue and red histograms, respectively. The contribution of merged pions increases with p_T.

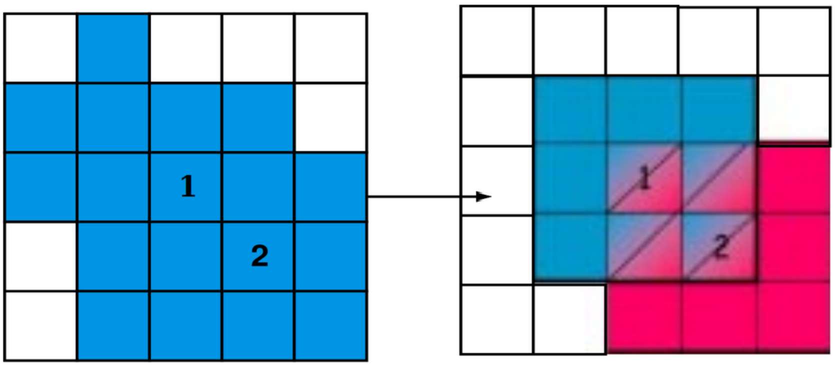


Figure 3.3: Two interleaved subclusters built from a single cluster, a procedure used in the merged-π^0 reconstruction algorithm.

3.1.3 Particle identification

To improve the identification of photons, a multivariate discriminator is constructed to distinguish between photons and merged neutral pions using a dedicated tool [108], based mainly on the cluster shape. Other discriminant variables are the energy ratio between the most energetic cell and the total cluster energy $E_{\mathrm{seed}}/E_{\mathrm{cluster}}$, and the multiplicity of PS hits in the cells in front of the ECAL cluster. Another variable, the Confidence Level (CL), distinguishes between photons, hadrons and electrons, combining information of the calorimeter and tracking systems.

3.2 Improvement of the photon energy profile

The dedicated algorithm for the reconstruction of merged-π^0 makes an extensive use of the energy profile of single photons to split the energy between the two subclusters, as explained in Section 3.1.2. Knowing accurately this profile is crucial for the algorithm to succeed. The energy profile describes the expected 2D energy spread of the cluster. It is defined as the energy fraction $E_{\mathrm{fraction}} = E_{\mathrm{cell}}/E_{\mathrm{cluster}}$, of the energy deposited in a cell respect to the total energy of the cluster, against the distance from the cell to the energy-weighted barycenter of the cluster.

The expected profile is obtained from a Monte Carlo simulation of isolated photons, also known as *particle guns*, and described empirically as a sum of exponentials with 10 parameters:

$$E_{\text{fraction}} = \begin{cases} c_1\, e^{-c_2\, x} + c_3\, e^{-c_4\, x} + c_5\, e^{-c_6\, x}, & x > 0.5 \\ 2 - c_7\, e^{-c_8\, x} - c_9\, e^{-c_{10}\, x}, & 0 \leq x \leq 0.5 \end{cases} \tag{3.2}$$

where x represents the distance cell-to-barycenter, in cell units. The parameter values were calculated in 2003, so the goal of this task was to update the parameters with the current calorimeter conditions (after Run 1 data-taking), as well as with the latest available simulations of isolated photons.

The profile has to be determined independently for clusters reconstructed in each of the calorimeter regions (inner, middle and outer), since the cell sizes are different (see Section 2.3.3), and for each photon type (converted and unconverted). A photon is categorized as converted if there is at least one hit in the SPD cells situated in front of the ECAL cluster. From a study using isolated photons, the number of converted photons is around 10% (see Table 3.1), although it depends slightly of the initial energy.

Six simulation samples with photon energies ranging from 10 to 168 GeV were used, each containing between 50 000 and 70 000 photons. However, the profile is constructed from the energy deposits on single cells (which are called energy *digits*), thus each photon provides several data points, around 10 energy digits per photon. All the information about the clusters are extracted using a custom-made script after running Brunel [99], the LHCb reconstruction software. The number of digits is listed in Table 3.1, representing the statistical power available in a given sample. Combining all the samples, the final profiles are constructed with more than 3.7 million energy digits.

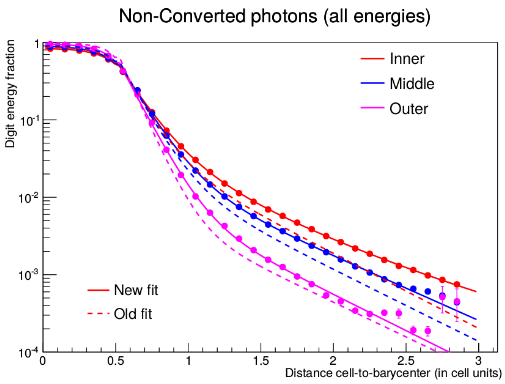
The photon profiles (all energies combined) are shown in Figure 3.4, comparing the current and previous fits to Equation 3.2; the new parameter values are listed in Table 3.2. These numbers are currently implemented in the database of the LHCb experiment, improving the neutral pion reconstruction with respect to the previous implementation. For completeness, the profiles for different initial photon energies are shown in Figures 3.5 and 3.6, in logarithmic scale.

Table 3.1: Incoming photon energy, number of energy digits, and percentage of converted photons, for each Monte Carlo sample of isolated photons.

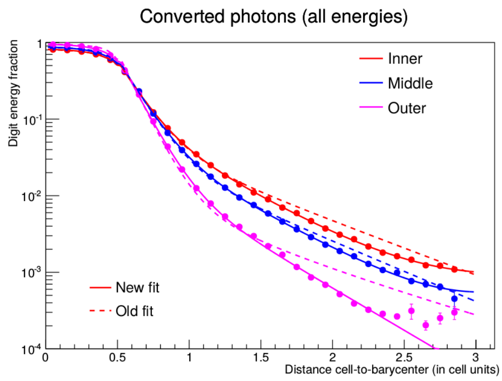
Energy (GeV)	Digits (in thousands)	Converted (%)
10	373	7.8
16.8	891	8.5
33.8	788	9.1
50	367	9.5
100	514	9.9
168	822	10.7
All	3 757	9.4

Table 3.2: Result of the fits to the improved photon energy profiles (all energies combined). They correspond to the fits represented in Figure 3.4.

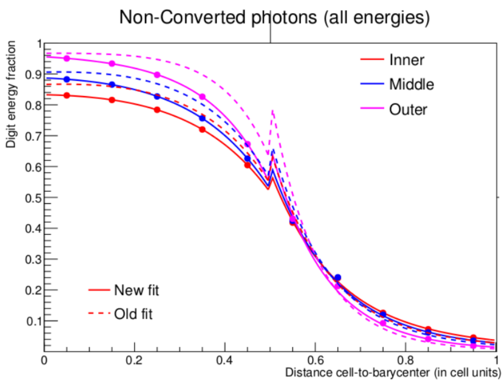
Parameter	Unconverted			Converted		
	Outer	Middle	Inner	Outer	Middle	Inner
c_1	0.0410	0.0394	0.2275	0.1482	0	0.3021
c_2	2.1509	1.9218	2.6185	2.4826	0	2.3200
c_3	43.6975	20.7799	17.7579	45.8971	23.1886	21.5929
c_4	8.4164	7.1533	7.0892	8.5348	7.4668	7.6034
c_5	0.2728	0.04234	0.0179	-0.0506	0.1710	0.0002
c_6	8.4166	1.9219	1.1905	2.4045	2.2209	-0.3830
c_7	1.0345	1.0841	1.1376	1.0412	1.0976	1.1556
c_8	-0.0182	0.0885	0.1070	0.0312	0.0993	0.0860
c_9	0.0095	0.0287	0.0296	0.0148	0.0335	0.0301
c_{10}	-7.5416	-5.4355	-5.2408	-6.6393	-5.0517	-5.0649



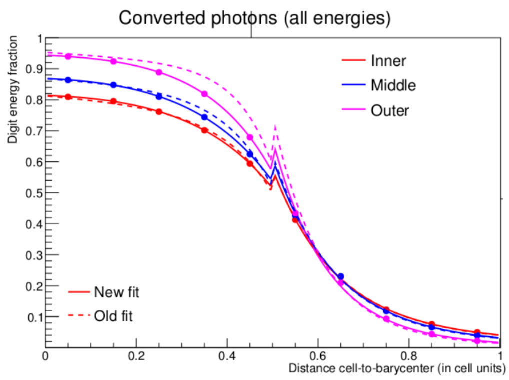
(a) Unconverted. Logarithmic scale



(b) Converted. Logarithmic scale

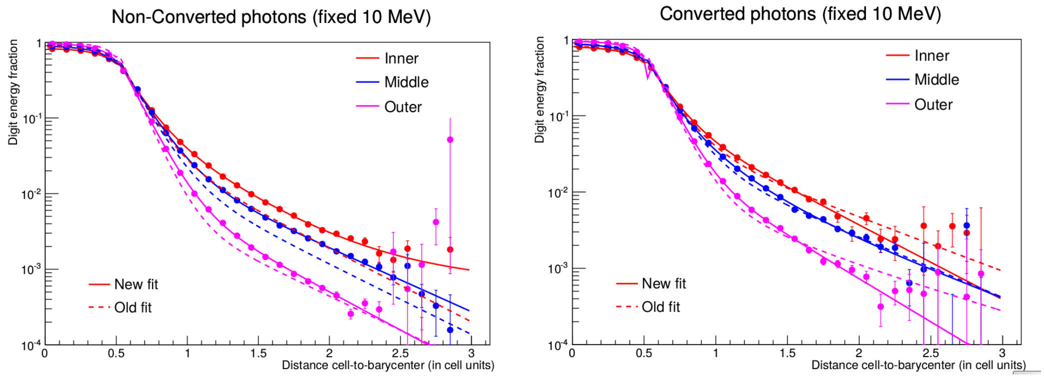


(c) Unconverted. Cartesian scale

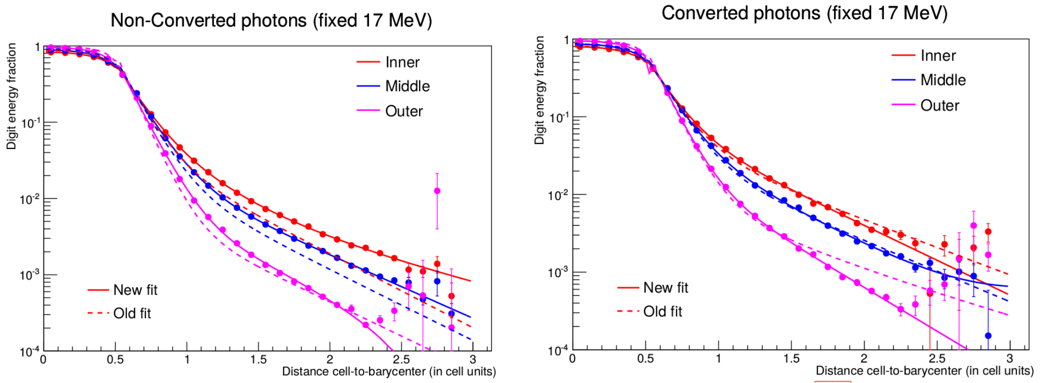


(d) Converted. Cartesian scale

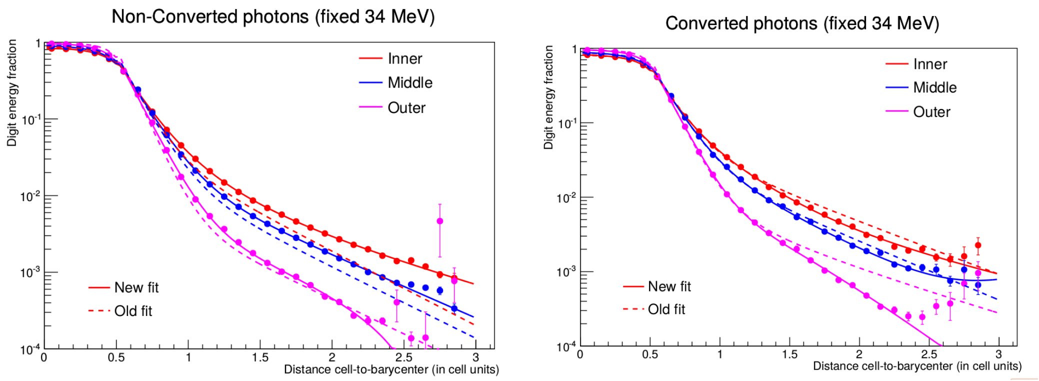
Figure 3.4: Energy profiles of unconverted (left) and converted (right) photons, all energies combined. The updated fit (solid lines) and data points are drawn over the previous fit (dashed lines). The profile is shown in logarithmic (above) and cartesian (below) scales, in the latter case zooming at the most populated region. The $x = 0.5$ discontinuity is expected following the definition of Equation 3.2.



(a) 10 GeV/c

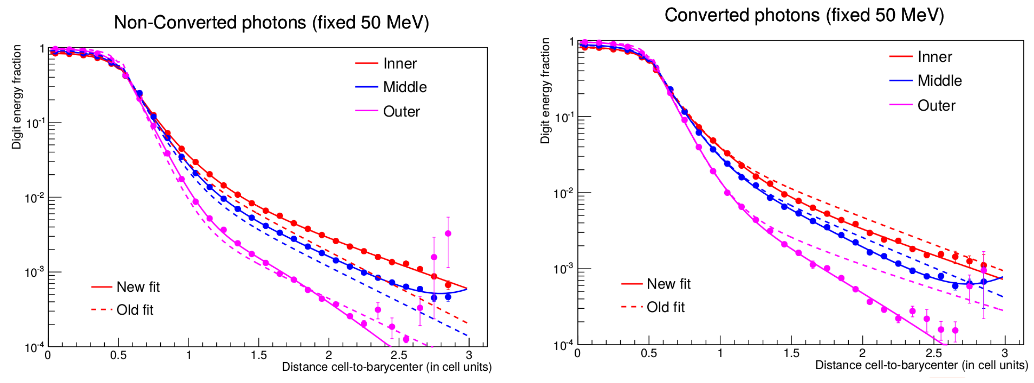


(b) 16.8 GeV/c

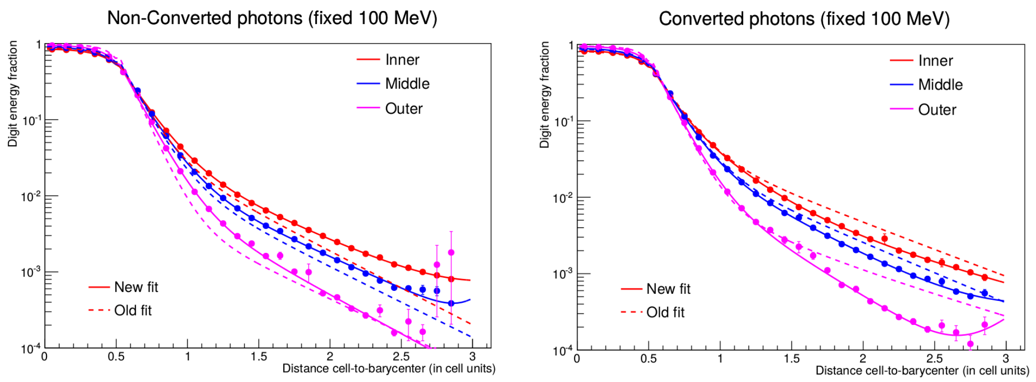


(c) 33.8 GeV/c

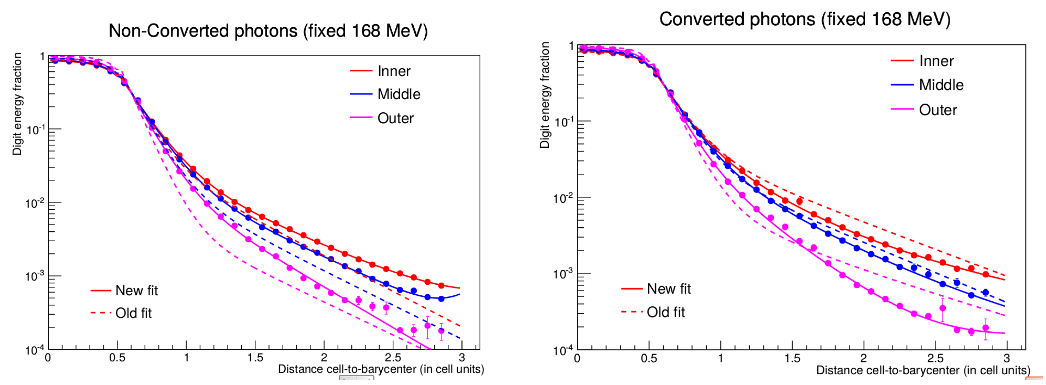
Figure 3.5: Energy profiles of photons with an incoming energy of (a) 10 GeV/c, (b) 16.8 GeV/c and (c) 33.8 GeV/c.



(a) 50 GeV/c



(b) 100 GeV/c



(c) 168 GeV/c

Figure 3.6: Energy profiles of photons with an incoming energy of (a) 50 GeV/c, (b) 100 GeV/c and (c) 168 GeV/c.

3.3 Event selection

To select $B_s^0 \to \phi(\to K^+K^-)\gamma$ and $B^0 \to K^{*0}(892)(\to K^+\pi^-)\gamma$ candidates[1] one needs to search for a highly energetic photon, together with two tracks that form a displaced vertex from all the primary vertices of the event. The reconstructed intermediate meson (ϕ or K^{*0}) has to be compatible with its expected mass, and the projection of the reconstructed B meson needs to point towards a collision vertex. The photon does not contribute to the vertex determination since it leaves no hits in the tracking system. Figure 3.7 shows a picture of the decay topology.

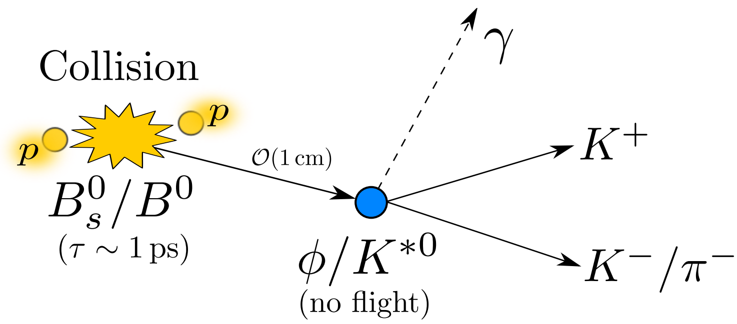


Figure 3.7: Topology of the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ radiative decays.

The sizeable lifetime of B mesons, of the order of 1 ps, combined with a high average momentum of around 100 GeV/c, produces an average flight distance of $\mathcal{O}(1\,\mathrm{cm})$ within the LHCb domains. This distance is enough to reconstruct displaced vertices, since the decay vertex resolution is typically of a few millimeters (see Figure 3.8).

The trigger, stripping and offline selections, described respectively in Sections 3.3.1, 3.3.2 and 3.3.3, have to be defined to achieve a high signal efficiency and background rejection. A cut-based offline selection is chosen instead of a multivariate-based solution to align the selections of the signal and control channels, and diminish possible systematic effects caused by the selection choice. In this section, the equivalent selections for the Run 1 and Run 2 periods are quoted, although only the former is used in the analysis.

The selection criteria exploits the fact that particles decaying from a heavy particle are produced with a higher component of the momentum

[1]From now on, the conjugate decay is implied.

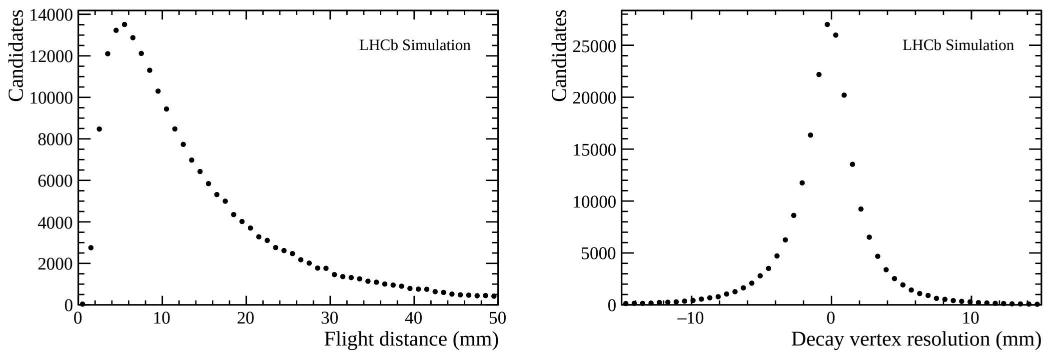
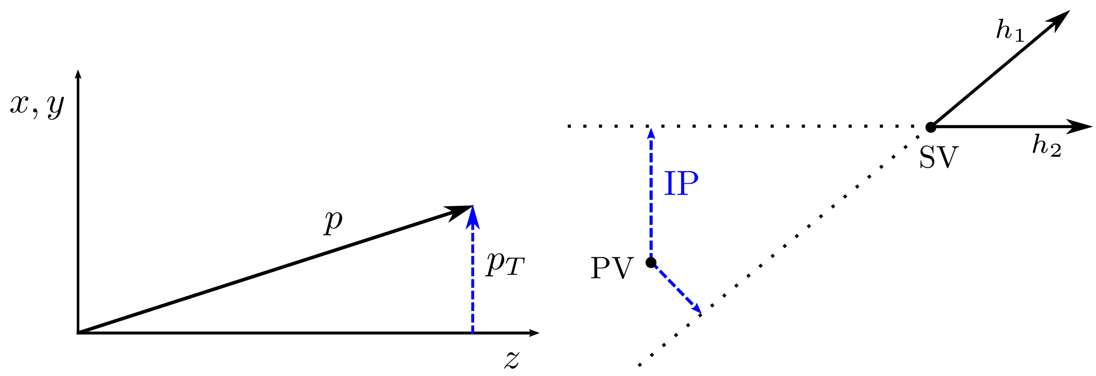


Figure 3.8: Distance travelled by the B meson in simulated $B_s^0 \to \phi\gamma$ decays (left), compared to the decay vertex resolution in the z direction (right), after the event reconstruction.

transverse to the beam axis $p_T = \sqrt{p_x^2 + p_y^2}$ compared to other fragmentation particles produced in the collision. The definition of p_T is represented in Figure 3.9 (a). The impact parameter (IP) is another useful variable to discard tracks coming from one of the primary vertices; it is constructed projecting the track backwards, and evaluating the minimum distance from the projection to the closest primary vertex. The definition of the IP variable is represented in Figure 3.9 (b). Sometimes, it is more useful to make requirements on its significance $\chi^2_{\text{IP}} = (\text{IP}/\sigma_{\text{IP}})^2$, since it takes into account the uncertainties of the track extrapolation and the vertex determination.



(a) Transverse momentum

(b) Impact parameter

Figure 3.9: Definition of the transverse momentum (p_T) and impact parameter (IP) variables.

3.3.1 Trigger

A trigger strategy has to be defined to select the signal candidates, choosing specific lines for each one of the L0, HLT1 and HLT2 trigger steps. In this analysis, every mentioned line is triggered-on-signal (TOS), to control better the effect of the selection in the data and simulation samples. In a TOS line, the candidate particles are the ones that fired the trigger, as explained in the LHCb Trigger section 2.4. Tables 3.3 and 3.4 show a schematic view of the Run 1 and Run 2 trigger strategies.

Table 3.3: Trigger strategy for $B^0_s \to \phi\gamma$ (and $B^0 \to K^{*0}\gamma$) in Run 1.

	Run 1. Standard path	Run 1. High-E_{T} path
`L0`	`L0Photon_TOS` or `L0Electron_TOS`	`L0PhotonHi_TOS` or `L0ElectronHi_TOS`
`HLT1`	`Hlt1TrackAllL0_TOS`	`Hlt1TrackPhoton_TOS`
`HLT2`	`Hlt2Bs2PhiGamma_TOS` (`Hlt2Bd2KstGamma_TOS`)	

Table 3.4: Trigger strategy for $B^0_s \to \phi\gamma$ (and $B^0 \to K^{*0}\gamma$) in Run 2.

	Run 2
`L0`	`L0Photon_TOS` or `L0Electron_TOS`
`HLT1`	`Hlt1TrackMVA_TOS` or `Hlt1TwoTracksMVA_TOS`
`HLT2`	`Hlt2RadiativeBs2PhiGamma_TOS` (`Hlt2RadiativeBd2KstGamma_TOS`)

L0 and HLT1 strategies

In radiative decays the L0 strategy focuses on the photon, requiring either the lines `L0Photon` or `L0Electron`. They select ECAL clusters with a minimum transverse energy that ranged between 2.5 and 3 GeV during 2011 and 2012, and between 2.3 and 2.7 GeV in 2015.

The selection of displaced tracks from the primary vertices is performed in the HLT1 trigger step. The specific requirements are listed in Table 3.5. For the Run 1 period, either the cut-based lines `Hlt1TrackAllL0` or `HLT1Photon` are required, a selection that is called *standard path* in the table. Good quality tracks are selected by the χ^2 of the track fit and the number of hits in the Velo and T-stations. Fragmentation tracks are rejected by the IP significance χ^2_{IP} and kinematical p and p_T requirements. The `HLT1Photon` line requires looser kinematical requirements in the hadron tracks for photons with transverse momentum higher than 4.2 GeV, a threshold implemented in the `L0PhotonHi` and `L0ElectronHi` lines. After including this *high-E_T path*, the number of events increases by an additional 20%.

In Run 2 the HLT1 lines are based on a multivariate analysis (MVA) technique, improving the efficiency of the line and making them easier to tweak. For this data period we require either `Hlt1TrackMVA` or `Hlt1Two-TrackMVA` to select a good-quality track or two tracks coming from the same vertex. Table 3.6 shows the list of requirements. The MVA algorithm is trained with four variables: the track χ^2, the flight distance χ^2 to select displaced vertices, the sum of p_T of the tracks, and the number of tracks in the event with $\chi^2_{\mathrm{IP}} < 16$. Before the MVA, a loose cut-based preselection is applied in the same variables, as listed in Table 3.6.

HLT2 strategy

Finally, a dedicated HLT2 line selects efficiently the candidates for the desired decay, based on an offline-like selection criteria [109]. They are called `Hlt2Bs2PhiGamma` and `Hlt2Bd2KstGamma` for the $B^0_s \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays, respectively[2]. These lines apply specific requirements for the exclusive channels in the full chain of the decay, as shown in Table 3.7.

The discriminant variables are the momenta p and p_T and the IP significance χ^2_{IP}, for the tracks (required to be incompatible with a primary vertex) and the B meson (required to be compatible with a PV), as well as the usual quality χ^2 requirements for tracks and vertices. The DIRA angle (θ_{dira}) of the B meson is a powerful variable for background rejection; it is the angle between the momentum of the particle and the decay length vector (the direction defined from the primary vertex to the decay vertex). Its definition is depicted in Figure 3.10. Usually, θ_{dira} takes large values for

[2]One should note that, in Run 2, the line names changed to `Hlt2RadiativeBs2PhiGamma` and `Hlt2RadiativeBd2KstGamma`, maintaining the same requirements.

Table 3.5: HLT1 trigger selection for Run 1. A logical OR of the standard and high-E_T paths is applied.

Variable	Hlt1AllL0	Hlt1Photon
p_T (MeV/c)	> 1800	> 1200
p (MeV/c)	$> 10\,000$	
χ^2/ν	< 3	
χ^2_{IP}	> 16	
IP (mm)	> 0.1	
Number of T-Hits	> 16	> 15
Number of Velo Hits	> 9	
Velo Qcut	3	4
Trigger	L0All	L0Photon or L0Electron

Table 3.6: HLT1 trigger selection for Run 2. The ghost probability cut was introduced in 2016.

Variable	Hlt1TrackMVA	Hlt1TwoTracksMVA
p_T (MeV/c)	$[1000, 25\,000]$	> 500
p (MeV/c)	–	$> 5\,000$
χ^2/ν	< 2.5	
χ^2_{IP}	> 7.4	> 4
Ghost Probability	< 0.2	–
Corrected mass (MeV/c^2)		$[1000, 10^9]$
η		$[2, 5]$
Intermediate meson p_T (MeV/c)	–	> 2000
DIRA (mrad)	–	> 0
Vertex χ^2	–	< 10
χ^2		MVA
Flight Distance χ^2		MVA
Sum p_T		MVA
Number of tracks with $\chi^2_{\text{IP}} < 16$		MVA

Table 3.7: HLT2 trigger selection criteria, required by the exclusive lines `Hlt2Bs2PhiGamma` and `Hlt2Bd2KstGamma`.

	Variable	Run 1		Run 2	
		$B_s^0 \to \phi\gamma$	$B^0 \to K^{*0}\gamma$	$B_s^0 \to \phi\gamma$	$B^0 \to K^{*0}\gamma$
Tracks	χ^2/ν	< 5		< 4	
	χ^2_{IP}	> 20		> 20	
	p_T (MeV/c)	> 500		> 500	
	p (MeV/c)	–		> 3000	
ϕ/K^{*0}	Vertex χ^2/ν	< 25	< 16	< 25	< 16
	Mass window (MeV/c^2)	20	100	20	100
	p_T (MeV/c)	–		> 1500	
B	Vertex χ^2/ν	< 25	< 16	< 16	
	p_T (MeV/c)	> 2000		> 2000	
	$\Delta M(B)$ (MeV/c^2)	1500		1500	
	χ^2_{IP}	> 12		> 12	
	DIRA (mrad)	< 0.063	< 0.045	< 0.063	< 0.045
Triggers	(L0Photon or L0Electon) and Hlt1(Two)TrackMVA				

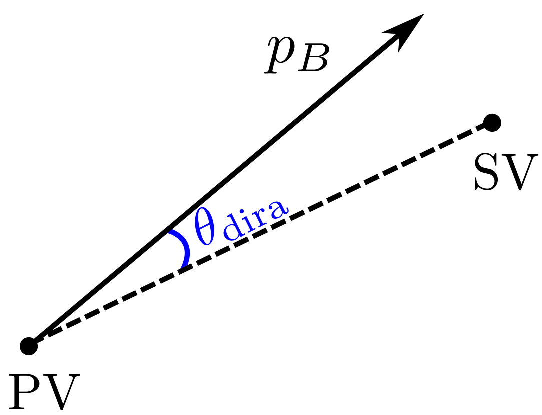


Figure 3.10: Definition of the DIRA angle (θ_{dira}) variable.

background events since the two vectors are not aligned.

The mass windows of the intermediate particles are particularized for each decay since ϕ is a narrower resonance compared to the K^{*0} meson. The mass widths for $\phi(1020)$ and $K^{*0}(892)$ are 4.247 ± 0.016 and 47.3 ± 0.5 MeV, respectively [20]. $\Delta M(\phi) < 20$ MeV$/c^2$ corresponds to 5 times the ϕ width, while $\Delta M(K^{*0}) < 100$ MeV$/c^2$ is about 2 times the K^{*0} width. The mass window of the initial meson $\Delta M(B) < 1500$ MeV$/c^2$ rejects combinatorial background while still keeping some sidebands near the peak, useful for background characterization.

3.3.2 Stripping

The stripping versions used in this analysis are the `Stripping21` for 2011 and 2012, and `Stripping24` for 2015 data sets. One of the improvements of the event selection of this thesis respect to the reference [59] is a 20% increase in signal yield in the Run 1 dataset, coming from the `Stripping21` improved selection, as explained below. The previous analysis used an exclusive line of the `Stripping20p2` version.

Radiative lines follow at present an inclusive strategy. There are lines for decays of the type $B \to X(Nh)\gamma$ with $N = 2, 3$ or 4 hadron tracks in the final state, which are combined with a photon. The tracks are reconstructed without particle identification requirements, assuming a pion mass hypothesis. An inclusive strategy is useful in the searching of new channels that have not been observed yet, or in the study of physical backgrounds.

In the present thesis we make use of the inclusive strategy with the `B2XGamma2pi` line[3], which selects two hadrons in the final state $B \to X(hh)\gamma$. The requirements are listed in Table 3.8; they are, in general, less stringent than the exclusive HLT2 lines mentioned in the previous section. The most distinguished feature is the addition of tight kinematical selection criteria for one of the tracks, as well as a higher quality χ^2 decay vertex. A requirement on the ghost probability of the tracks is also introduced. This variable, that ranges between 0 and 1, allows to reject tracks that are reconstructed from spurial combinations of hits produced from fragmentation particles.

[3] One should note that, in Run 2, the name changed to `Beauty2XGamma2pi`.

Table 3.8: Stripping selection for the inclusive two-track radiative line $B \to X(hh)\gamma$.

	Variable	Cut
Tracks	χ^2/ν	< 3
	χ^2_{IP}	> 16
	p_T (MeV/c)	> 300
	p (MeV/c)	> 1000
	$p_T^{h_1} + p_T^{h_2}$ (MeV/c)	> 1500
	GhostProb	< 0.4
Tight Track[1]	χ^2/ν	< 2.5
	p_T (MeV/c)	> 1000
	p (MeV/c)	> 5000
	IP (mm)	> 0.1
Photon	p_T (MeV/c)	> 2000
	CL	> 0
B	Vertex χ^2/ν	< 9
	χ^2_{IP}	< 9
	DIRA (mrad)	> 0
	M (MeV/c^2)	$[3280, 9000]$
HLT2 trigger	At least one TIS or TOS radiative trigger	

[1] At least one track needs to fulfill all these requirements.

3.3.3 Offline selection

A cut-based offline selection has been developed in this thesis to reject background while still retaining a high signal efficiency. This solution was preferred over a MVA selection in order to align the selections between the $B_s^0 \to \phi\gamma$ signal and $B^0 \to K^{*0}\gamma$ control channels, allowing for a better control of the reconstruction effects. Table 3.9 contains the full list of requirements.

Table 3.9: Offline selection criteria for $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$.

	Variable	Cut
PID preselection	$K^\pm$ ProbNNk	> 0.2
	$\pi^\pm$ ProbNNpi	> 0.2
	$\pi^\pm$ ProbNNk	< 0.2
Optimization	Tracks χ^2_{IP}	> 55
	γ/π^0 separation	> 0.6
	Helicity $\lvert\cos(\theta_H)\rvert$	< 0.8
	B p_T (MeV/c)	> 3000
	$\Delta M(\phi)$ ($\Delta M(K^{*0})$) (MeV/c^2)	< 15 (< 100)

The selection consists of two parts. First, a particle identification (PID) preselection is made to retain specific hadrons in the final state; two kaons in the case of the $B_s^0 \to \phi\gamma$ channel, and a kaon and a pion for $B^0 \to K^{*0}\gamma$ decays. Second, an optimized set of requirements is chosen using the most discriminant variables available to distinguish between signal and combinatorial background.

The particle identification algorithms associate the reconstructed hadronic tracks to pions and kaons. The PID preselection removes most of the physical background that comes from the misidentification of the final state hadrons. The requirements `ProbNNk` $>$ `0.2` for kaons and `ProbNNpi` $>$ `0.2 && ProbNNk` $<$ `0.2` for pions are chosen so most of the expected background from physical sources is eliminated. The construction of the particle identification variables is carried out by the `PIDCalib` package [110], using a neural network approach that combines the information of every subsystem into a single probability for each particle hypothesis.

The optimization of the selection and the study of the discriminating variables are detailed in the following.

Optimization procedure

To optimize the offline selection, one needs to maximize a figure-of-merit that depends on the number of signal S and background B events in a given region. The chosen figure-of-merit in this analysis has been the signal significance:

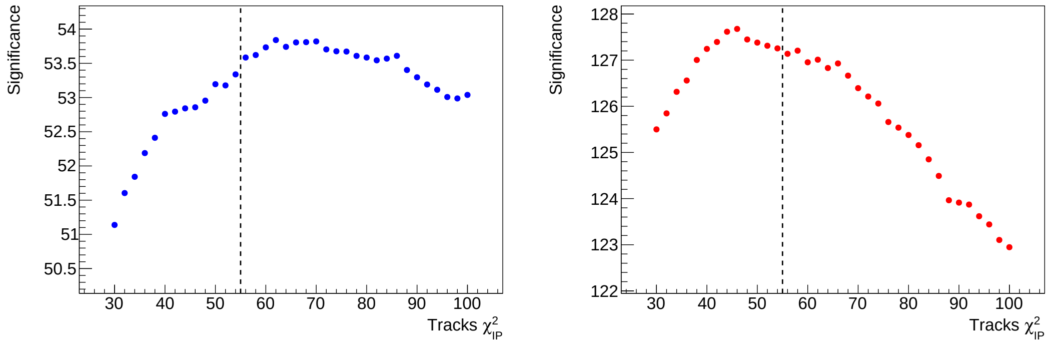
$$\mathcal{S} = \frac{S}{\sqrt{S+B}}, \tag{3.3}$$

which is a quantity usually utilized in high energy physics to maximize the observation probability of undiscovered decays (since $\sigma_{\text{stat.}}(S+B) = \sqrt{S+B}$). In our case, the best figure-of-merit would be the one that maximizes the sensitivity to the parameters to measure, taking into account not only the statistical but also the systematic uncertainty. The chosen quantity $\mathcal{S}$ is a reasonable approximation to the best figure-of-merit in measurements dominated by the statistical uncertainty, as it is the present analysis.

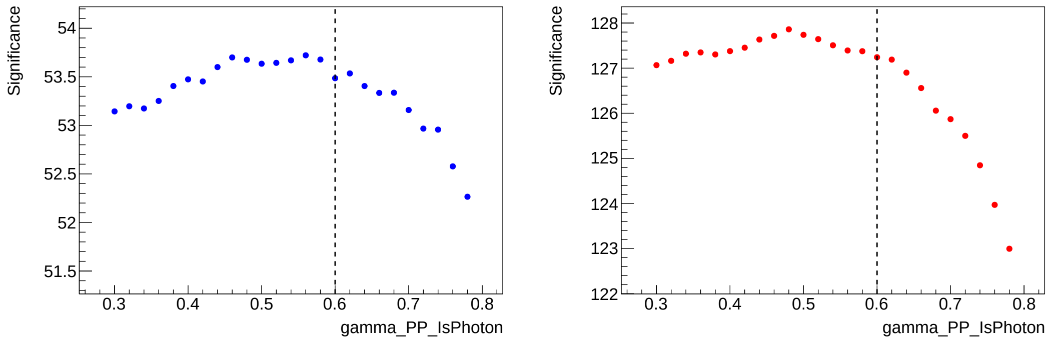
For the optimization procedure, the signal yield S is extracted from simulation and the background yield B is taken from high-mass sideband data, a region expected to be populated only by combinatorial background.

To find the optimal point, the signal significance is calculated for different selection cut values, following an iterative procedure from the most to the least discriminating variable. Figures 3.11 and 3.12 show the signal significance curves for each variable while keeping fixed the rest of the cuts. The selected values, represented in the plots as dashed lines, are chosen to be the same for the two channels in the offline selection since the maxima are close to each other. The significance increase on real data after each requirement is shown in Table 3.10, while the effect of the full offline selection can be noticed on Figure 3.13, making evident the decrease of combinatorial background under the signal peak.

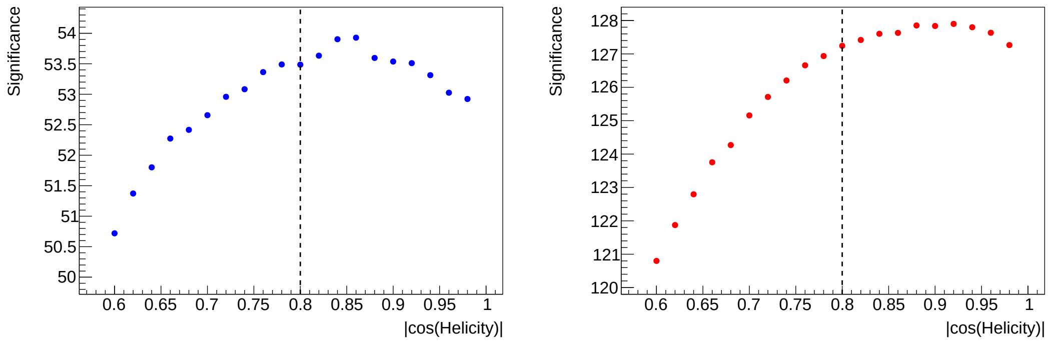
The whole optimization procedure is cross-checked extracting both the signal and background yields from data. After applying a cut, the reconstructed B mass is fitted with the sum of a Gaussian for the signal and an exponential for the combinatorial background. The signal and background



(a) χ^2_{IP} of the hadron tracks

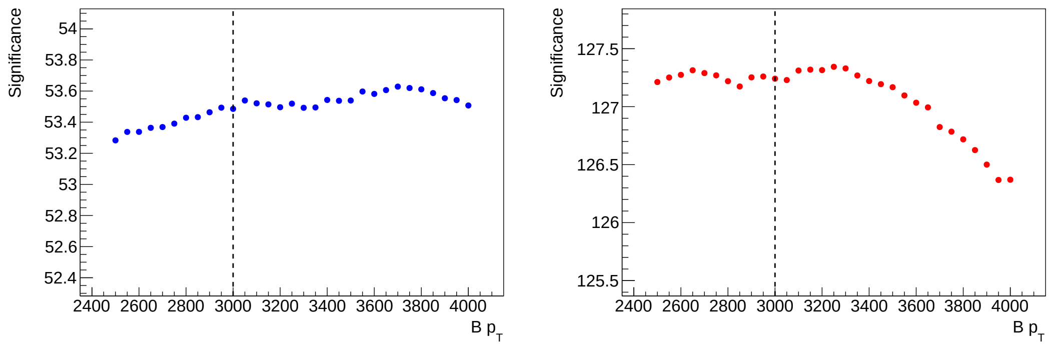


(b) γ/π^0 separation

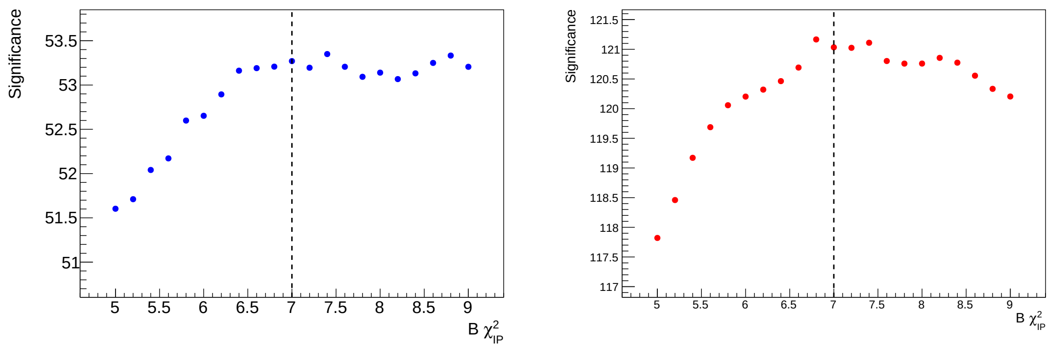


(c) Cosinus of the helicity angle

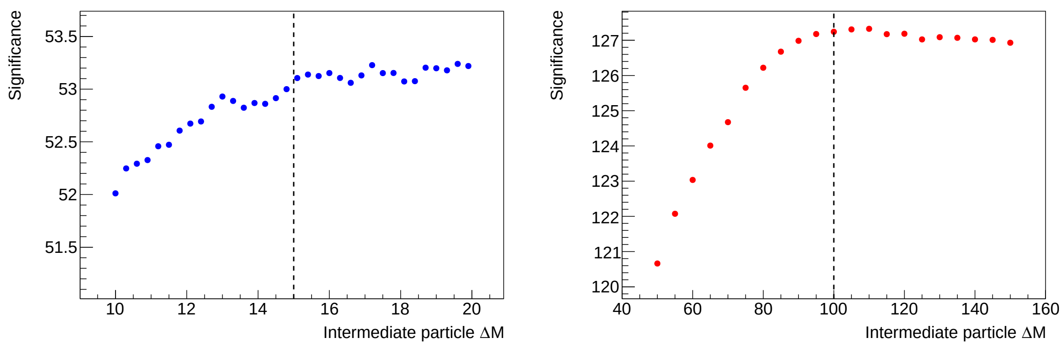
Figure 3.11: Signal significance against the cut value on the discriminant variables, where the rest of the cuts have been applied, for $B^0_s \to \phi\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) channels in the 2011 + 2012 dataset. The dashed line marks the chosen offline requirement, close to the maximum.



(a) B transverse momentum



(b) B $\chi^2_{\rm IP}$

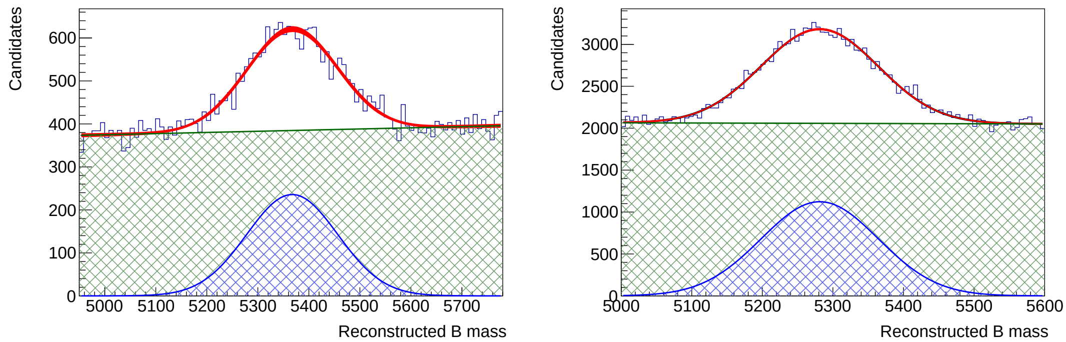


(c) Mass window of the intermediate resonance (note the different range)

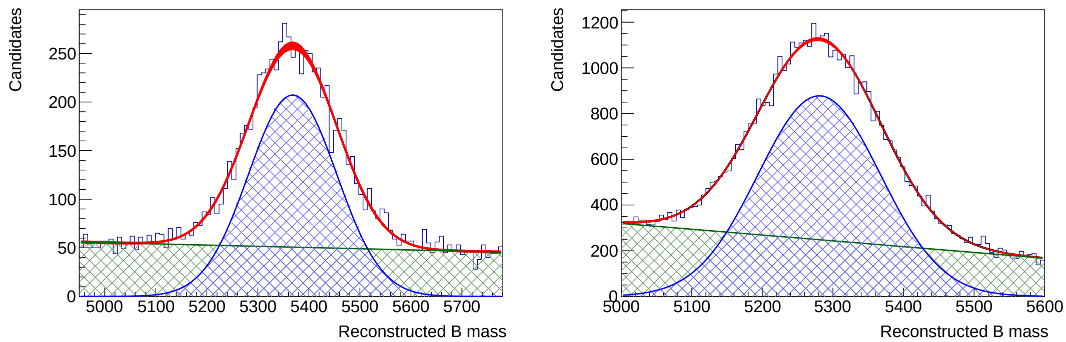
Figure 3.12: Signal significance against the cut value on the discriminant variables, where the rest of the cuts have been applied, for $B_s^0 \to \phi\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) channels in the 2011 + 2012 dataset. The dashed line marks the chosen offline requirement, close to the maximum.

Table 3.10: Significances of each requirement of the offline selection, evaluated from a fit to the B mass distribution on $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ data. Each requirement is added on top of the previous one.

Requirements	Significance $\mathcal{S}_{\phi\gamma}$	Significance $\mathcal{S}_{K^{*0}\gamma}$
Starting point	21.8	89.7
+ PID preselection	38.5	91.68
+ χ^2_{IP}	48.8	112.2
+ γ/π^0	52.0	119.0
+ $\lvert\cos\theta_H\rvert$	53.2	120.2
+ B p_T	53.7	122.7



(a) Before the offline selection. Significances: $\mathcal{S}_{\phi\gamma} = 38.5$ and $\mathcal{S}_{K^{*0}\gamma} = 91.7$.



(b) After the offline selection. Significances: $\mathcal{S}_{\phi\gamma} = 53.7$ and $\mathcal{S}_{K^{*0}\gamma} = 122.7$.

Figure 3.13: Effect of the offline selection on the reconstructed B mass distribution, for $B_s^0 \to \phi\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) channels in the 2011 + 2012 dataset. The fit is made for visualization purposes using the sum of a Gaussian for signal, in blue, and a linear function for background, in green.

yields are taken as the area of the Gaussian and the exponential, respectively, within the $1\,\sigma$ region around the Gaussian peak as provided by the fit. The optimal points are compatible with the nominal procedure.

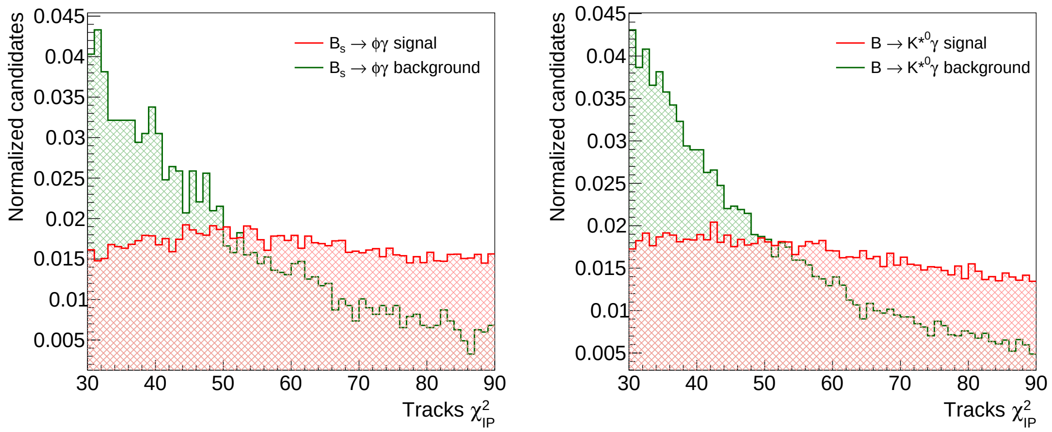
Discriminant variables

The most powerful variables to reject combinatorial background are the χ^2_{IP} of the tracks, the γ/π^0 separation and the B helicity. Other variables entering in the selection are the B transverse momentum and the mass window of the intermediate particles ϕ/K^{*0}. The discriminating power of these variables are clearly exposed in Figure 3.14, where the signal and background distributions are normalized and plotted together. The signal distributions are extracted from simulation samples, while the background corresponds to high-mass sideband data. All of the variables have similar distributions in the $B^0_s \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ samples, since the kinematical properties of these two channels are similar. Variables that affect differently the proper time acceptance in the two channels were not used in the optimization, as it is the case of the DIRA angle or the flight distance χ^2 [111].

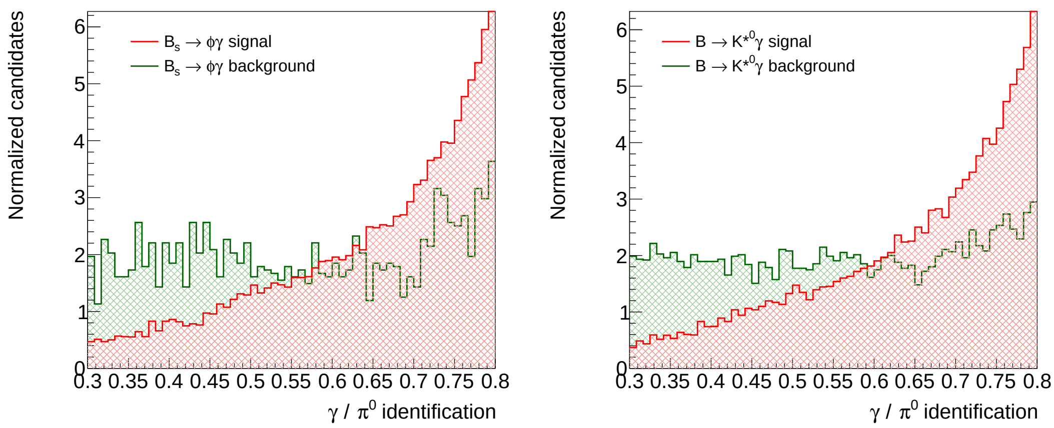
The χ^2_{IP} is the most efficient variable to remove the combinatorial background. It is the significance of the impact parameter, taking into account the uncertainties of the track projection and the primary vertex. The requirement is made on the minimum χ^2_{IP} of the two hadron tracks of the candidate, $\min(\chi^{2,h_1}_{\mathrm{IP}}, \chi^{2,h_2}_{\mathrm{IP}}) > 55$, corresponding to a significance greater than $7\,\sigma$. This tight optimal cut evidences the fact that most of the combinatorial background is coming from fragmentation tracks produced in the primary vertex, which is the reason that a requirement on the secondary vertex quality χ^2_{SV} is not useful to reject the background.

The γ/π^0 variable distinguishes between high-energy photons and neutral pions reconstructed as a single cluster. Several signatures related to the shape and energy of the cluster are translated into a MVA decision variable [108], as explained in Section 3.1.3. Events with a value higher than 0.6 on this value are selected.

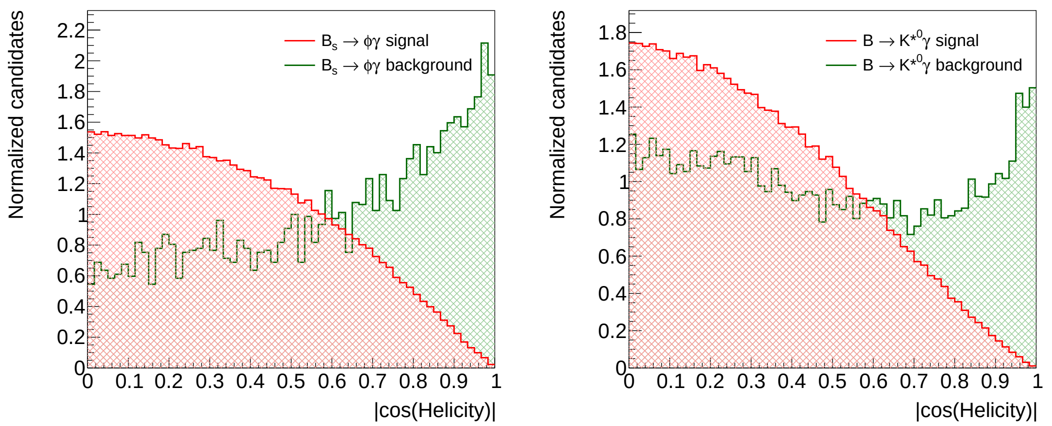
The helicity θ_H is the angle between the K^+ and the B meson in the rest frame of the intermediate resonance. The selection criteria makes a requirement on the absolute value of the cosine of the helicity angle $|\cos\theta_H|$.



(a) $\chi^2_{\rm IP}$ of the hadron tracks



(b) γ/π^0 separation



(c) Cosinus of the helicity angle

Figure 3.14: Normalized signal and background distributions of the discriminant variables used in the offline selection, for the $B^0_s \to \phi\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) channels. The signal distribution (red area) is taken from simulation, while the background (green area) corresponds to data with $m_B > 5500$ MeV/c^2.

Without reconstruction effects, a $\sin^2\theta_H$ distribution is expected since the ϕ is a $J^P = 1^-$ particle, where J is the total angular momentum and P is the parity quantum number. Background sources coming from a π^0 follow a $\cos^2\theta_H$ dependency, while the combinatorial background is expected to produce a uniform distribution. Removing events with small helicity angles $|\cos\theta_H| < 0.8$ provides the best balance of signal efficiency and background rejection. The requirements on the helicity angle and the γ/π^0 variable reject not only combinatorial background, but also part of the physical backgrounds with a π^0 in the final state.

3.4 Background subtraction

After applying the event selection criteria there is still a significant amount of background contamination, as Figure 3.13 shows. The background contribution is subtracted using the sPlot [104] technique to disentangle the different components, from a fit to the invariant B mass distribution where the distributions of the background sources are extracted using information of real and simulated data. The sPlot algorithm associates weights to each event of the sample, related to the probabilities to pertain to the different components, as explained in the following.

3.4.1 The sPlot technique

The background contribution is subtracted using sPlot [104], a statistical tool to unfold data distributions when several components (sources of events) are merged into a single sample. The unfolding is performed through a maximum likelihood fit to a set of known distributions (the discriminating variables), to obtain the unknown distributions of the other variables (the target variables). In this analysis the discriminating variable is the invariant B mass, and the target variables are the observables needed for the measurement (the decay time, the decay time uncertainty, the mistag fraction, etc.). The method estimates a per-event weight (called *sWeight*) for the signal component $w_s(m_i)$, related to the probability of an event to pertain to the signal distribution. Background events receive small or negative signal sWeights, so when applied to the corresponding events they effectively perform the background subtraction. The signal distributions are obtained after weighting the sample by the signal sWeight $w_s(m_i)$.

sPlot weights

The per-event signal sWeight is defined as:

$$w_s(m_i) = \frac{\sum_j V_{sj} F_j(m_i)}{\sum_k N_k F_k(m_i)}, \tag{3.4}$$

where $j, k = \{s, b_1, b_2, ...\}$ are the components (the signal and the different background sources), F_j are the probability density functions, N_j the yield of j-type events, and V_{sj} the covariance matrix obtained by inverting:

$$V_{sj}^{-1} = \frac{\partial^2(-\mathcal{L})}{\partial N_s \partial N_j} = \sum_i^N \frac{F_s(m_i) F_j(m_i)}{\left[\sum_k N_k F_k(m_i)\right]^2}, \tag{3.5}$$

which is equal to the second derivative of the likelihood function. The histograms of a variable weighted by the per-event sWeight, $x_i\, w_s(m_i)$, represents the true signal distribution of x, since the background contribution to the histogram cancels out on a statistical basis as a consequence of the $w_s(m_i)$ definition choice. However, this method is valid under the assumption that the target and discriminating variables are uncorrelated.

3.4.2 Background sources

Background types

Several types of background sources can be distinguished:

- The **combinatorial** background is produced by random combinations of tracks coming from the collision vertices. The mass distribution of the reconstructed mother is uniform, usually flat or linear.

- **Physical** backgrounds are those produced by specific decay channels that satisfy the requirements of the event selection. They can be separated into two categories:

  - **Partially reconstructed** backgrounds, where at least one final particle in the decay has not been reconstructed. Their mass distribution is peaked at the low-mass region (i.e., values lower than the signal peak), due to the energy loss of the missing particles.
  - **Peaking** backgrounds, where at least one final particle has been wrongly reconstructed as one of the signal particles (misidentification). They populate the signal region; their mass distribution is similar to the signal since there is no energy loss.

Relative contaminations

Any physical decay with at least two hadrons in the final state and a highly energetic γ or π^0 can potentially contribute as background. An exhaustive study was performed in reference [107] to determine the physical backgrounds with a significant contamination C_B, relative to the signal:

$$C_B = \frac{N_B}{N_S} = \frac{f_{q,B}}{f_{q,S}} \frac{\mathcal{B}(B)}{\mathcal{B}(S)} \frac{\epsilon_B}{\epsilon_S}, \tag{3.6}$$

where B and S are the background and signal decays, f_q is the probability for a quark q to hadronize into a b-hadron (q makes reference to the other quark of the hadron), $\mathcal{B}$ is the branching fraction of the decay channel and ϵ the total reconstruction efficiency for a particular process. When computing relative contaminations, the dependency on the accquired number of events and production cross-sections cancel out.

The branching fractions and hadronization probabilities are fixed to the PDG [20] averages, while the selection efficiencies are obtained from LHCb Monte Carlo full simulation, applying the signal event selection to several simulated decays. The relative contaminations are calculated in the B mass region from 4600 to 6000 MeV/c^2.

Relevant background components for the $B_s^0 \to \phi\gamma$ decay

The only partially reconstructed background with a significant contribution to the $B_s^0 \to \phi\gamma$ decay is $B^+ \to \phi K^+\gamma$, in which one of the kaons is not reconstructed. The relative contamination has been calculated to be 1.2%. The relevant peaking backgrounds are $\Lambda_b \to \Lambda^*(pK)\gamma$ (1.8% relative contamination, with the proton reconstructed as a kaon), $B^0 \to \phi(KK)\pi^0$ (1.6%, where the π^0 is reconstructed as γ) and $B^0 \to K^{*0}(K\pi)\gamma$ (0.1%, where a pion is reconstructed as a kaon).

Relevant background components for the $B^0 \to K^{*0}\gamma$ decay

The $B^0 \to K^{*0}\gamma$ reconstruction, however, is affected by several physical backgrounds to a higher degree due to the wider mass of the intermediate meson K^{*0} compared to the ϕ meson. Three groups of partially reconstructed backgrounds can be distinguished:

- $B^0 \to K^{*0}\eta(\gamma\gamma)$ in which a photon is missing, with a contribution of 2% relative to the signal.

- Decays of the type $B \to (K\pi\pi)\gamma$ with or without an intermediate resonance, where one of the pions is not reconstructed; among them, the main contributions are $B^+ \to K_1(1270)^+\gamma$ (12%), $B^0 \to K_1(1270)^0\gamma$ (7%), $B^+ \to K_1(1400)^+\gamma$ (5%), and $B^+ \to K_2^*(1430)^+\gamma$ (2%).

- Decays of the type $B^0 \to D^0\rho$, where three pions (neutral or charged) are missing, and a π^0 is reconstructed as a photon. The main contributions are coming from the decays $B^+ \to D^0(K\pi\pi^0)\rho^+(\pi\pi^0)$ (4%) and $B^+ \to D^0(K\pi\pi\pi)\rho^+(\pi\pi^0)$ (0.4%).
- $B^+ \to \rho^+\rho^0$ (0.2%): in which a charged pion is missing, the other charge pion is misidentified as a kaon, and the neutral pion is misidentified as a photon.

Among the peaking backgrounds, those with significant relative contaminations are: $\Lambda_b \to \Lambda^*(pK)\gamma$ (2%, with the proton reconstructed as a pion), $B^0 \to K^{*0}(K\pi)\pi^0$ (2%, with the π^0 reconstructed as a photon), $B^0 \to \rho^0(\pi\pi)\gamma$ (0.2%, where a pion is reconstructed as a kaon), and $B_s^0 \to \phi(KK)\gamma$ (0.2%, where a kaon is reconstructed as a pion).

3.4.3 Modelling of the signal mass distribution

The signal mass distribution is modelled by a double-sided Crystal Ball function, that behaves as a Gaussian function on its core but having larger exponential tails:

$$\mathrm{CB}(m;\mu,\sigma,\alpha_i,n_i) = \begin{cases} A_L\left(B_L - \frac{m-\mu}{\sigma}\right)^{-n_L}, & \text{for } \frac{m-\mu}{\sigma} \leq -\alpha_L \\ \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right), & \text{for } -\alpha_L < \frac{m-\mu}{\sigma} < \alpha_R \\ A_R\left(B_R + \frac{m-\mu}{\sigma}\right)^{-n_R}, & \text{for } \frac{m-\mu}{\sigma} \geq \alpha_R \end{cases} \tag{3.7}$$

where $i = \{L, R\}$, $\alpha_{L,R} > 0$ and:

$$\begin{aligned} A_i &= \left(\frac{n_i}{\alpha_i}\right)^{n_i} \exp\left(-\frac{\alpha_i^2}{2}\right), \\ B_i &= \frac{n_i - 1}{\alpha_i}. \end{aligned} \tag{3.8}$$

Long tails are needed in both sides (left L and right R), described by the $n_{L,R}$ and $\alpha_{L,R}$ parameters; the low-mass distribution is affected by energy losses in the material volume of the detector, while the high-mass distribution can be populated with ECAL pile-up deposits, artificially raising the reconstructed momentum of the photon.

The parameters that describe the shape of the tails $n_{L,R}$ and $\alpha_{L,R}$ are extracted from a fit to the simulation samples (and fixed in the data fit), while μ and σ are left to float in the data fit. Table 3.11 lists the Crystal Ball parameters for the signal.

Table 3.11: Fitted parameters of the signal and combinatorial mass components. The μ, σ (for signal) and p_0 (for combinatorial) are fitted from data, while the other parameters are extracted from simulation and fixed in the data fit.

Component	Parameter	Obtained from	$B_s^0 \to \phi\gamma$	$B^0 \to K^{*0}\gamma$
Signal	μ (MeV/c^2)	Data	5368.8 ± 1.7	5279.0 ± 0.8
	σ (MeV/c^2)	Data	85.9 ± 1.8	90.2 ± 0.9
	α_L	Simulation	2.475 ± 0.029	2.398 ± 0.034
	α_R	Simulation	1.419 ± 0.036	1.332 ± 0.033
	n_L	Simulation	0.424 ± 0.039	0.564 ± 0.054
	n_R	Simulation	7.34 ± 0.87	8.37 ± 0.96
Combinatorial	p_0 (MeV^{-1})	Data	-0.185 ± 0.030	-0.128 ± 0.060

3.4.4 Modelling of the background mass distributions

The background mass shapes are modelled with different functions depending on the background type, as listed in Table 3.12.

Table 3.12: The functions to describe the signal and background components of the mass distribution.

Component	Model
Signal	Double-sided Crystal Ball (Equation 3.7)
Combinatorial background	First-order polynomial (Equation 3.9)
Partially reco. background	ARGUS $\otimes$ Gaussian (Equation 3.10)
Peaking background	Single or double-sided Crystal Ball (Equation 3.7)

- The combinatorial background is described by a first-order polynomial:

$$P_{\text{comb}}(m; p_0) = 1 + m\, p_0. \tag{3.9}$$

The slope p_0 is fitted directly on data since the mass sideband yield is enough to characterize the distribution. The fitted value is quoted in Table 3.11.

- Peaking backgrounds are described by a single- or double-sided Crystal Ball, depending on the decay channel. The width is expressed relative to the signal $\sigma_{\text{peak}} = f_\sigma\, \sigma_{\text{signal}}$, where f_σ is the scaling parameter. For the mean, an offset is introduced $\mu_{\text{peak}} = \mu_{\text{signal}} + \Delta M$. The parameters f_σ and ΔM are extracted from simulation data, as well as α_i and n_i. All of them are fixed in the data fit. The values used in the fit are listed in Table 3.13.

- Partially reconstructed backgrounds are described by an ARGUS function convolved with a Gaussian:

$$P_{\text{partial}}(m; \Delta m_0, c, p, f_\sigma) = \text{ARGUS}(m; m_B + \Delta m_0, c, p) \otimes G(0, f_\sigma\, \sigma_0), \tag{3.10}$$

with the ARGUS defined as [112]:

$$\text{ARGUS}(m; m_0, c, p) = m\left(1 - \frac{m^2}{m_0^2}\right)^p \exp\left\{c\left(1 - \frac{m^2}{m_0^2}\right)\right\}, \tag{3.11}$$

where m_0, c and p are the cutoff mass, curvature and power, respectively, and $m \leq m_0$ (or else the function is zero). The parameters Δm_0, c, p and f_σ are extracted from simulation, while m_B is set to the mass of the initial B meson of the background component (B^+, B^0, B_s^0...). The Gaussian $G(0, f_\sigma\, \sigma_0)$ is scaled relative to the signal width σ_0 through the scaling parameter f_σ. All of the parameters are fixed in the data fit. Table 3.14 lists the parameters for each partial background.

Table 3.13: Parameters of the peaking background components, extracted from simulation.

Peaking Backgrounds	ΔM (MeV/c^2)	f_σ	α	n
$\boldsymbol{B_s^0 \to \phi\gamma}$				
$\Lambda_b^0 \to \Lambda^*\gamma$	10 ± 14	1.4 ± 0.1	1.0 ± 0.4	11 ± 20
$B^0 \to K^{*0}\gamma$	17.2 ± 9.5	1.2 ± 0.1	2.0 ± 0.4	1.0 ± 0.8
$\boldsymbol{B^0 \to K^{*0}\gamma}$				
$B^0 \to K^{*0}\pi^0$	-52 ± 16	1.2 ± 0.1	0.6 ± 0.2	3.1 ± 2.6
$\Lambda_b^0 \to \Lambda^*\gamma$	73.2 ± 4.0	1.43 ± 0.04	0.96 ± 0.08	3.2 ± 0.7
$B_s^0 \to \phi\gamma$	-11.5 ± 0.7	1.13 ± 0.01	$\alpha_L = 1.82 \pm 0.04$ $\alpha_R = 1.73 \pm 0.06$	$n_L = 1.4 \pm 0.1$ $n_R = 5.7 \pm 0.9$
$B^0 \to \rho\gamma$	78.2 ± 0.8	1.06 ± 0.01	$\alpha_L = 2.19 \pm 0.04$ $\alpha_R = 1.25 \pm 0.04$	$n_L = 0.9 \pm 0.1$ $n_R = 19.3 \pm 6.0$

Table 3.14: Parameters of the partially reconstructed backgrounds in the mass distribution, extracted from simulation.

Partial Backgrounds	Δm_0 (MeV/c^2)	f_σ	c	p
$\boldsymbol{B_s^0 \to \phi\gamma}$				
$B^+ \to \phi K^+\gamma$	$-m_K$	0.89 ± 0.04	-4.9 ± 0.6	0.3 ± 0.1
$\boldsymbol{B^0 \to K^{*0}\gamma}$				
$B^0 \to K_1(1270)\gamma$	$-m_\pi$	1.21 ± 0.03	-5.5 ± 0.4	0.13 ± 0.04
$B^0 \to K^{*0}\eta$	-52.5 ± 7.4	1.26 ± 0.04	-4.2 ± 0.4	0.00 ± 0.03
$B \to D\rho$	-418.7	2.5 ± 0.3	-3.3 ± 3.3	1.7 ± 0.8
$B^+ \to K_1(1270)^+\eta$	-418.7	2.5 ± 0.3	-3.3 ± 3.3	0.4 ± 0.5

3.4.5 Fit to data mass distributions

The data mass distribution is fitted including all the signal and background components, using an extended maximum likelihood fit:

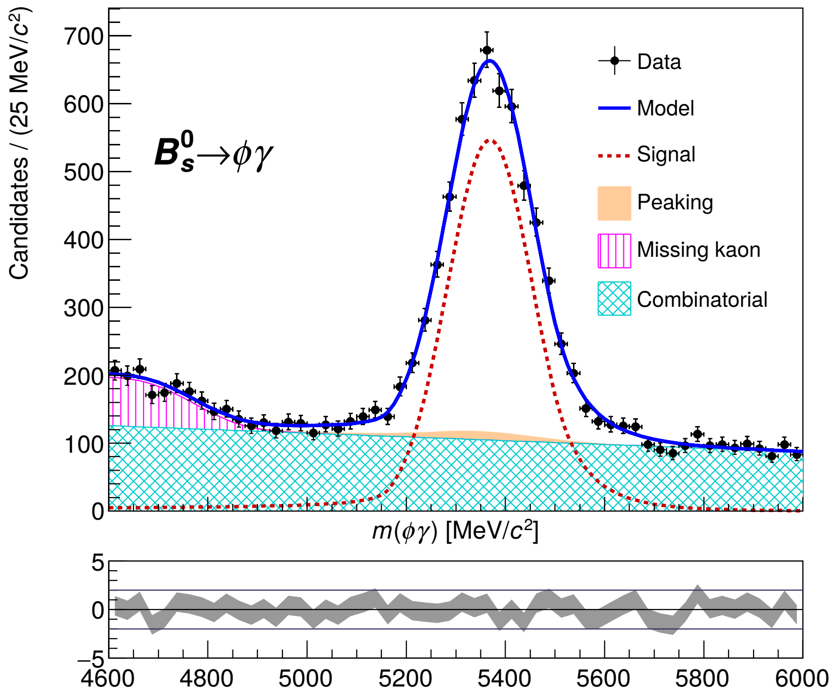
$$\begin{aligned} P(m) = N_s & \left[P_{\text{signal}}(m) + \sum_k C_k P_{k,\text{peaking}}(m) \right] \\ & + N_{\text{comb}} P_{\text{comb}} + \sum_i N_{i,\text{partial}} P_{i,\text{partial}}, \end{aligned} \tag{3.12}$$

where N_s is the signal yield, and $C_k = N_k/N_s$ is the peaking background contribution relative to the signal. The sPlot method cannot be used to subtract the peaking backgrounds since their mass distribution is similar to the signal. The fit is limited to the range [4600, 6000] MeV/c^2 to avoid border effects in the distribution coming from differences between the online (trigger) and offline selections.

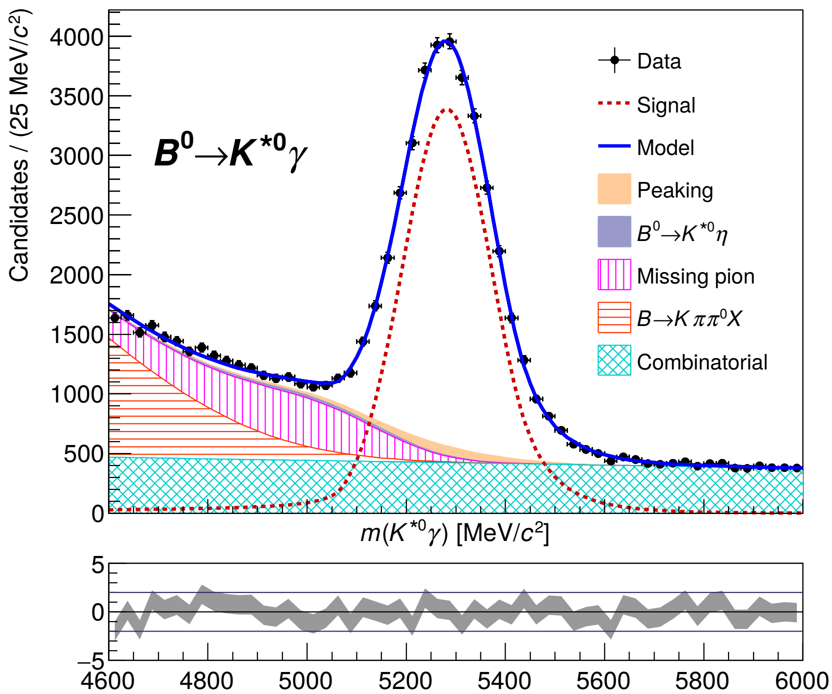
The yields of the components are left free, except for the components with mass distributions similar to the signal, which are: (1) the peaking backgrounds, whose yields are fixed to $N_s C_k$, and (2) the partial background $B^0 \to K^* \eta$ reconstructed as $B^0 \to K^{*0}\gamma$, which has a mass distribution that is peaked in the signal region since the missing particle (the photon) has zero mass. A modified version of sPlot is needed to obtain correct results with fixed yields [113].

The data mass distributions are fitted twice. First, to obtain the parameters that are left free in the model: the signal mean μ, the signal width σ and the combinatorial slope p_0, to the values shown in Table 3.11, which are then fixed in the next iteration. Second, to obtain the yields of the different components and associate the sWeights to each event, since the sPlot method demands to fix all the nuisance parameters to evaluate the weights correctly.

The fits to the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ mass distributions are shown in Figure 3.15, and the resulting relative yields for the signal and background components are listed in Table 3.15. The signal component accounts for about 80% of the total yield under the signal region ($1.5\,\sigma$ around the peak) in the two channels. The combinatorial is the main background source, contributing to 15-20% of the yield, while peaking backgrounds contribute only to 2% of the events. Partially reconstructed backgrounds make a total



(a) $B_s^0 \to \phi\gamma$ mass fit



(b) $B^0 \to K^{*0}\gamma$ mass fit

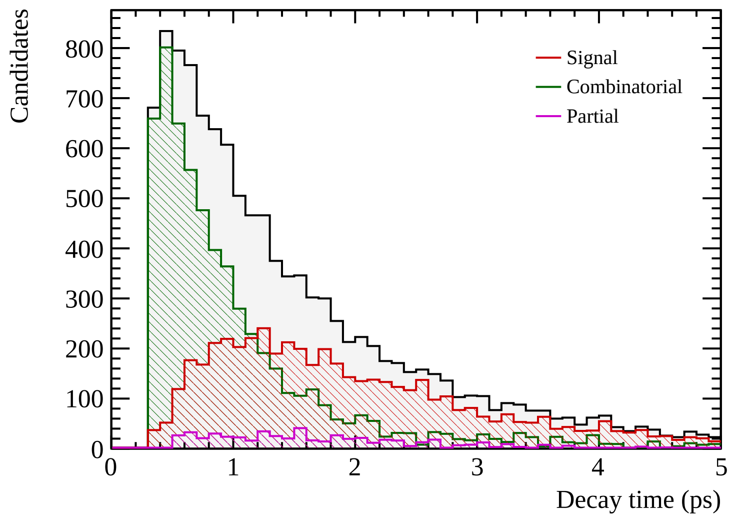
Figure 3.15: Fit to the B mass distributions of $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays, for the stripping 21 data samples. The yield of the signal component is $5\,266 \pm 104$ for $B_s^0 \to \phi\gamma$ and $33\,816 \pm 373$ for $B^0 \to K^{*0}\gamma$.

of 3% for $B^0 \to K^{*0}\gamma$ decays but are negligible for $B_s^0 \to \phi\gamma$ decays. The obtained signal yields correspond to $5\,266 \pm 104$ for $B_s^0 \to \phi\gamma$ and $33\,816 \pm 373$ for $B^0 \to K^{*0}\gamma$.

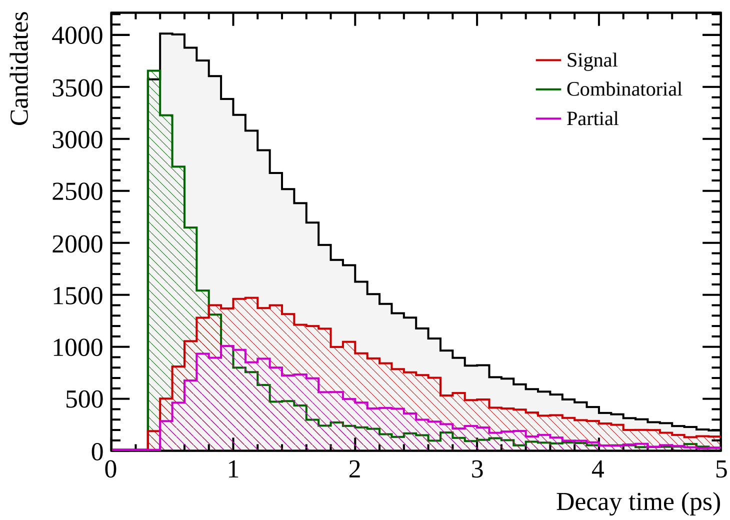
Figure 3.16 compares the decay time distribution of the signal, combinatorial and partial components, after applying the sPlot procedure in $B_s^0 \to \phi\gamma$ data. The comparison between background-subtracted data and simulation is shown in Appendix A.1, for several variable distributions, showing a general agreement.

Table 3.15: Relative yields of the signal and background components in the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ mass distributions, in the signal region ($1.5\,\sigma$ around the peak).

Component		$B_s^0 \to \phi\gamma$	$B^0 \to K^{*0}\gamma$
Signal		$(80.11 \pm 0.55)\%$	$(83.15 \pm 0.20)\%$
Combinatoria bkg.		$(20.45 \pm 0.56)\%$	$(14.58 \pm 0.20)\%$
Partial bkgs.	$B^+ \to \phi K^+\gamma$	0%	–
	$B \to K^{*0}\pi\gamma$	–	$(2.43 \pm 0.08)\%$
	$B^0 \to K^{*0}\eta$	–	$(0.41 \pm 0.06)\%$
	$B \to K\pi\pi\pi^0 X$	–	$(0.34 \pm 0.03)\%$
Peaking bkgs.	$\Lambda_b \to \Lambda^*(pK)\gamma$	$(1.17 \pm 0.14)\%$	$(0.79 \pm 0.05)\%$
	$B^0 \to \phi\pi^0$	$(0.82 \pm 0.12)\%$	–
	$B^0 \to K^{*0}\gamma$	$(0.07 \pm 0.03)\%$	–
	$B^0 \to K^{*0}\pi^0$	–	$(1.13 \pm 0.06)\%$
	$B^0 \to \rho\gamma$	–	$(0.09 \pm 0.01)\%$
	$B_s^0 \to \phi\gamma$	–	$(0.18 \pm 0.02)\%$



(a) $B_s^0 \to \phi\gamma$



(b) $B^0 \to K^{*0}\gamma$

Figure 3.16: Decay time distribution of (a) $B_s^0 \to \phi\gamma$ and (b) $B^0 \to K^{*0}\gamma$ candidates for the signal and background components, after applying the sPlot method on the mass distribution.

3.5 Decay time reconstruction

3.5.1 Decay time algorithm

The proper time t of a decaying particle can be calculated from its travelled distance L in the laboratory frame (decay length), corrected by a relativistic factor that is equivalent to the ratio of its mass and momentum:

$$ct = \frac{L}{\gamma\beta} = L\frac{m}{|\vec{p}|}. \tag{3.13}$$

At LHCb, the decay length vector is reconstructed from the decay vertex and the primary vertex positions $\vec{L} = \vec{x}_{\mathrm{DV}} - \vec{x}_{\mathrm{PV}}$.

An algorithm called `PropertimeFitter` calculates the proper time and its uncertainty. It minimizes a χ^2 quantity in an iterative procedure, varying the input values within their uncertainties to align the flight and momentum vectors $\vec{L}$ and $\vec{p}$, by minimizing $\vec{L} \times \vec{p}$. Further explanation of the algorithm can be found in [114, 115]. The mass of the particle used by the fitter can be fixed to the latest world-average value, instead of using the reconstructed quantity. This is called the *mass-constrained* option, and it is the chosen approach for the analysis, since it improves the decay time resolution (the width from 45 fs to 40 fs, and the mean from 8 fs to 5 fs) and decreases the correlation of the decay time with the reconstructed mass (to almost zero correlation), a condition required for a correct behaviour of the sPlot background subtraction procedure.

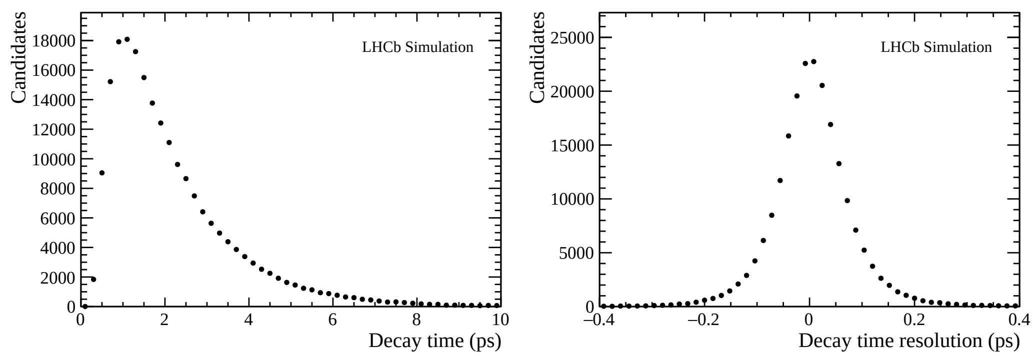


Figure 3.17: Decay time distribution (left) and resolution (right) in simulated $B_s^0 \to \phi\gamma$ decays, after applying the event selection criteria.

3.5.2 Decay time reconstruction effects

The decay time distribution from simulated $B_s^0 \to \phi\gamma$ candidates is shown in Figure 3.17 (left). Due to reconstruction effects, its shape is not an exponential-like distribution expected from the decay rate expression (Equation 1.28). To account for this fact, the distribution has to be modelled by introducing a decay time acceptance $A(t)$:

$$F(t) = \left[\Gamma(t^{'}) \otimes R(t,t^{'})\right] A(t). \tag{3.14}$$

where t is the reconstructed decay time and $t^{'}$ the truth decay time, $\Gamma(t^{'})$ the time-dependent decay rate and $R(t,t^{'})$ the decay time resolution.

Decay time acceptance

The decay time acceptance describes the inefficiencies produced by the reconstruction and the event selection in the decay time distribution. The determination of the acceptance is crucial in the measurement of the $\mathcal{A}^{\Delta}$ observable, since the term $\cosh(\Delta\Gamma_s\, t/2)$ is spread over the decay time distribution. The description of the decay time acceptance function is detailed in Sections 4.2 and 5.1, within the context of the untagged and tagged analyses, respectively.

Candidates with small decay times ($t < 1$ ps) are harder to distinguish from background events and, therefore, their efficiency is lower after the event selection. On the other hand, candidates with large decay times ($t > 4$ ps) are also affected by inefficiencies since the tracking system was optimized to select tracks from the collision point. The L0 and HLT1 trigger selections are the main sources of inefficiencies at high decay times, which were studied applying the selection step-by-step in simulation samples [107]. Figure 3.18 shows a fit to simulated $B_s^0 \to \phi\gamma$ decays, where the acceptance shape can be appreciated. In fits to the time-dependent decay rate, the acceptance normalization is not relevant since it is absorbed as an overall factor.

Decay time resolution

The decay time resolution accounts for the detector effects in the reconstruction of t, representing the change from the physical decay time to the

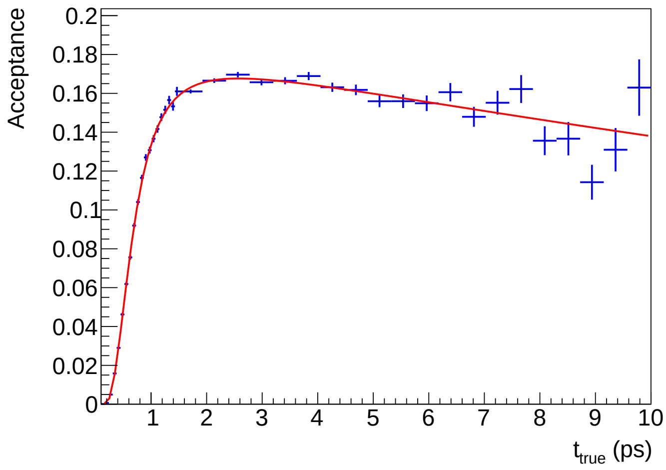


Figure 3.18: Decay time acceptance of simulated $B_s^0 \to \phi\gamma$ decays, after the event selection. The y-axis shows the ratio between the reconstructed and the generated decay time distributions. The fit will be detailed in Section 4.2.

reconstructed one. In simulation, it is defined as $\Delta t = t_{\text{reco}} - t_{\text{true}}$, the residual of the reconstructed and truth decay times. The determination of the resolution is crucial in the measurement of the $\mathcal{C}$ and $\mathcal{S}$ observables, since the terms $\cos(\Delta m_s\, t)$ and $\sin(\Delta m_s\, t)$ oscillate quickly. The description of the decay time resolution function is detailed in Section 5.3 within the context of the tagged analysis, since it is not relevant for the untagged analysis.

The decay time resolution of radiative decays is 70 fs on average, following a non-Gaussian distribution, with an average bias of 5 fs, as shown in Figure 3.17 (right). This resolution is about four times narrower than the oscillation period of the B_s^0.

4 Measurement of $\mathcal{A}^{\Delta}_{\phi\gamma}$ in $B^0_s \to \phi\gamma$ decays

This chapter describes the measurement of the $\mathcal{A}^{\Delta}_{\phi\gamma}$ parameter in $B^0_s \to \phi\gamma$ decays, the only mixing-induced parameter accessible in this channel without tagging the flavour of the initial meson. The work performed in this thesis has lead to the first measurement of $\mathcal{A}^{\Delta}_{\phi\gamma}$, using the Run 1 dataset of LHCb [59].

Section 4.1 presents a summary of the analysis strategy, as well as a list of the used datasets for real and simulated data. Section 4.3.1 details the probability density function used to describe the decay time distribution, considering acceptance and resolution effects. The description of the fit strategy is presented in Section 4.3, along with their validation using toy Monte Carlo experiments. The study of the expected sensitivity is performed in Section 4.4, as well as further cross-checks, while Section 4.5 describes the systematic effects relevant for the measurement. Finally, Section 4.6 exposes the final results to real data.

4.1 Analysis strategy

- A fit to the decay time is performed simultaneously to the signal channel $B^0_s \to \phi(\to K^+K^-)\gamma$ and the control channel $B^0 \to K^{*0}(892)(\to K^+\pi^-)\gamma$ (and its conjugate decay). The latter is used to extract the decay time acceptance.

- The candidates are reconstructed, selected and background-subtracted based on the procedure described in Chapter 3.

- The analysis does not need flavour tagging.

- Two independent fit strategies were developed: (1) an unbinned maximum likelihood fit to both channels, and (2) a binned fit to the ratio of yields between the two channels.

- The description of the decay time acceptance is crucial for the $\mathcal{A}^{\Delta}$ measurement. On the other hand, the full description of the decay time resolution is not needed. The acceptance parameters are extracted from $B^0 \to K^{*0}\gamma$ data, while the small difference in acceptance between the two channels, ΔA, is obtained from simulation samples.

- All the procedure was performed maintaining blinded the $\mathcal{A}^{\Delta}$ value.

4.1.1 Datasets

An integrated luminosity of 3 fb^{-1} collected by the LHCb experiment in 2011 and 2012 is used in the analysis, corresponding to 1 fb^{-1} of pp collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV in 2011 and 2 fb^{-1} at 8 TeV in 2012, as stated in Table 4.1. The data-taking is produced in two states of the magnet polarity, up and down, to avoid potential systematic effects introduced by the magnetic field direction. On average, 1.79 visible interactions per bunch-crossing were recorded.

Table 4.1: Integrated luminosity of Run 1 samples, splitted by year (2011 and 2012) and magnet polarity (up and down).

Year	$\sqrt{s}$ (TeV)	Polarity	Integrated luminosity (pb^{-1})
2011	7	Down	569.2 ± 19.9
2011	7	Up	415.2 ± 14.5
2012	8	Down	1015.9 ± 35.6
2012	8	Up	1033.6 ± 36.2

4.1.2 LHCb full-simulation samples

About 9 million of simulated events, generated with at least one $B_s^0 \to \phi(K^+K^-)\gamma$ decay chain, are reconstructed and selected with the same criteria as real data with Run 1 conditions. A simulation sample of the same size is used also for $B^0 \to K^{*0}(K^+\pi^-)\gamma$ decays; they correspond to the `Sim08d` simulation samples produced by LHCb. The simulated candidates are required to be signal by applying the truth-matching condition `B_BKGCAT == 0` || `B_BKGCAT == 50`, where the former means pure signal and the latter takes into account decays where the charged hadron emitted a hard photon. Around 90 000 signal candidates survive the reconstruction and event selection.

4.2 Decay time acceptance

An introduction to the decay time acceptance was given in Section 3.5. This section focuses on the description of the acceptance function in the context of the untagged analysis. Two parametrizations of the acceptance function are considered: one that requires four parameters, $A^{(4\text{par})}$, and another that requires only two, $A^{(2\text{par})}$, which are detailed in Section 4.2.1. The extraction of the acceptance difference between the signal and control channel, ΔA, is presented in Section 4.2.2.

4.2.1 Parametrizations

4-parameter acceptance

The 4-parameter acceptance function is modelled by the following power-law empirical expression:

$$A^{(4\text{par})}(t; a, n, t_0, \delta\Gamma) = \frac{[a(t-t_0)]^n}{1+[a(t-t_0)]^n}\, e^{-\delta\Gamma\, t}, \qquad \text{for } t > t_0. \tag{4.1}$$

The exponential term (parameter $\delta\Gamma$) characterizes the acceptance at large decay times, while the rest of the expression shapes the low decay-time acceptance (parameters a, n and t_0). The efficiency is zero for decay times less than t_0, acting as the starting point. Figure 4.1 (b) shows a fit to the acceptance of simulated $B_s^0 \to \phi\gamma$ decays, explicitely separating the low and high components of the function.

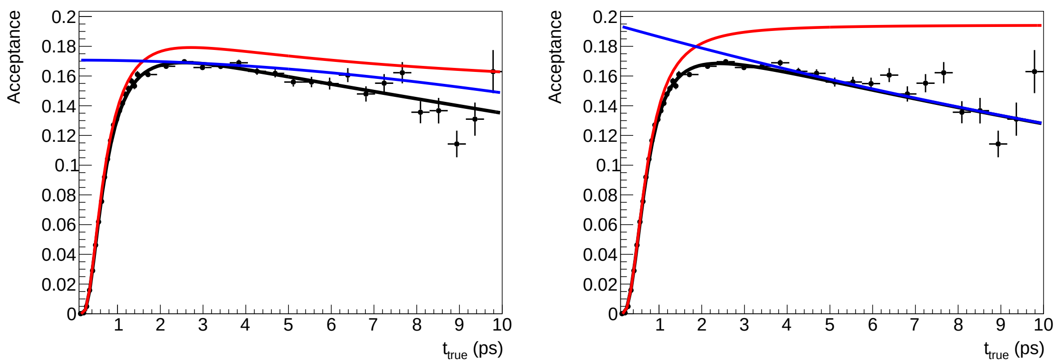
2-parameter acceptance

The 2-parameter acceptance function is described by:

$$A^{(2\text{par})}(t; b, c) = \frac{t^{b/t}}{\cosh{(ct)}}, \qquad \text{for } A^{(2\text{par})} > 0, \tag{4.2}$$

otherwise the function is evaluated as zero. The b and c parameters characterize the low and high acceptance, respectively. The Equation 4.2 can be divided by $e^{b/e}$ to ensure that all values fall under one, so it can be interpreted as an efficiency, although it is not necessary for the fit procedure. Figure 4.1 (a) shows a fit to the acceptance of simulated $B^0_s \to \phi\gamma$ decays, explicitely separating the low and high components of the function.

The low decay-time term, $t^{b/t}$, was found by fitting the acceptance distribution to hundreds of different parametrizations using the *zunzun* site [116], dedicated to online curve fitting. The function $t^{b/t}$ provided the lowest sum of residuals from the list of 2- and 3-parameter functions.



(a) $A^{(2\text{par})}$ (b) $A^{(4\text{par})}$

Figure 4.1: Low (red) and high (blue) components of the acceptance function, superimposed over the acceptance distribution of $B^0_s \to \phi\gamma$ simulated data.

4.2.2 Acceptance from full-simulation

The acceptance parameters of the simulated datasets are extracted from a simultaneous maximum likelihood fit to the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ samples to the following probability density function:

$$F(t^{'}) = \Gamma(t^{'})\, A(t^{'}), \tag{4.3}$$

where $t^{'}$ is the truth decay time (without resolution effects) and $\Gamma(t^{'})$ is the decay rate, whose parameters are fixed to the generated values listed in Table 4.2 corresponding to the `sim08` production of the LHCb simulation.

Table 4.2: Parameters used in the generation of the LHCb full-simulation. They correspond to the `sim08` production [117].

Parameter	Value	Parameter	Value
τ_d	1.519 ps	$\mathcal{A}^{\Delta}$	0
τ_s	1.512 ps	Δm_s	17.8 ps^{-1}
$\Delta\Gamma_s$	0.1097 ps^{-1}	Δm_d	0.507 ps^{-1}

The results are shown in Table 4.3. The $B^0 \to K^{*0}\gamma$ acceptance ($A_{K^*\gamma}$) is quoted directly, while the $B_s^0 \to \phi\gamma$ parameters ($A_{\phi\gamma}$) are quoted relative to the $B^0 \to K^{*0}\gamma$ ones:

$$\Delta A^{(i)} = A^{(i)}_{\phi\gamma} - A^{(i)}_{K^*\gamma}, \tag{4.4}$$

where $A^{(i)}$ represent each of the acceptance parameters. The results are also quoted for the probability density function defined as:

$$F(t) = \Gamma(t)\, A(t), \tag{4.5}$$

where t is the reconstructed decay time. The correlation between the ΔA parameters are taken into account to extract the systematic uncertainties due to the limited MC statistics, in Section 4.5.

Table 4.3: Acceptance parameters from simulation, for $B^0 \to K^{*0}\gamma$ and the difference between channels. The results are quoted with the PDF defined with respect to the t (reconstructed, Equation 4.5) and $t^{'}$ (truth, Equation 4.3) decay times.

	Acceptance $A_{K^*\gamma}$			Acceptance difference $\Delta A = A_{\phi\gamma} - A_{K^*\gamma}$	
Param.	$F(t^{'})$	$F(t)$	Param.	$F(t^{'})$	$F(t)$
a	1.873 ± 0.038	1.898 ± 0.037	Δa	0.014 ± 0.058	0.063 ± 0.054
n	2.149 ± 0.094	2.151 ± 0.092	Δn	0.08 ± 0.14	0.00 ± 0.13
t_0	0.178 ± 0.013	0.185 ± 0.013	Δt_0	-0.013 ± 0.020	0.001 ± 0.018
$\delta\Gamma$	0.0324 ± 0.0031	0.0313 ± 0.0030	$\Delta\delta\Gamma$	-0.0019 ± 0.0043	-0.0013 ± 0.0042
b	0.7225 ± 0.0048	0.7359 ± 0.0048	Δb	-0.0233 ± 0.0066	-0.0185 ± 0.0067
c	0.0516 ± 0.0047	0.0511 ± 0.0047	Δc	0.0022 ± 0.0064	0.0020 ± 0.0064

4.3 Fit strategies

This section focuses on the description of the decay time PDF, introduced in 4.3.1, and the explanation of the two fit strategies: the unbinned maximum likelihood fit (Section 4.3.2) and the ratio fit (Section 4.3.3).

4.3.1 Full PDF description

The decay time distribution is modelled multiplying the physical decay rate by an acceptance function:

$$F(t) = \Gamma(t)\, A(t). \tag{4.6}$$

Explicitly, for the $B^0_s \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays, the full PDF is described by:

$$\begin{aligned} F_{B^0_s\to\phi\gamma}(t) &= e^{-\Gamma_s t}\Big[\cosh\left(\Delta\Gamma_s\, t/2\right) - \mathcal{A}^{\Delta}\sinh\left(\Delta\Gamma_s\, t/2\right)\Big]\, A_{\phi\gamma}(t), \\ F_{B^0\to K^{*0}\gamma}(t) &= e^{-\Gamma_d t}\, A_{K^*\gamma}(t), \end{aligned} \tag{4.7}$$

where the time-dependent physical decay rate is the one derived in Section 1.5.1, summing the contributions of the B and $\bar{B}$ decay rates. In an untagged analysis, the cosinus and sinus terms of the decay rate cancel out, since they carry opposite signs in decays with B_s^0 and $\bar{B}_s^0$. The convolution is removed completely from the modelling by using the acceptance defined in Section 4.2, which absorbs the resolution bias effects. The resolution width do not play a relevant role in the $\mathcal{A}^{\Delta}$ measurement, as it will be shown in the systematic uncertainty section (5.6).

The acceptance parameters are extracted from the control decay $B^0 \to K^{*0}\gamma$ in the data fit, while the difference between channels is fixed from simulation:

$$A_{\phi\gamma}^{\rm data} = A_{K^{*0}\gamma}^{\rm data} + \Delta A_{\phi\gamma-K^{*0}\gamma}^{\rm MC}, \tag{4.8}$$

to the values shown in Table 4.3, defined from the $F(t^{'})$ definition of Equation 4.3. The determination of the acceptance is crucial in the measurement of $\mathcal{A}^{\Delta}$, since the $\sinh\left(\Delta\Gamma_s\, t/2\right)$ function is spread over the full range of the decay time distribution.

Input parameters from external measurements

The input parameters, entering in the time-dependent decay rate, are fixed in the data fit to the latest world-averages provided by the HFLAV group [39]. These parameters are the lifetimes τ_d and τ_s, the mass diferences Δm_d and Δm_s, and the decay rate widths $\Delta\Gamma_d$ and $\Delta\Gamma_s$. Table 4.4 lists the parameters taken from external sources. Their uncertainties, along with the correlation $\rho(\Gamma_s, \Delta\Gamma_s)$, are included in the fit as Gaussian constraints, as explained in detail in Section 4.3.

Table 4.4: Input parameters from external measurements, obtained from the HFLAV Summer-2017 average [39].

Parameter	World-average
τ_d	1.518 ± 0.004 ps
τ_s	1.5094 ± 0.0041 ps
$\Delta\Gamma_s$	0.090 ± 0.005 ps^{-1}
$\rho(\Gamma_s, \Delta\Gamma_s)$	-0.082
Δm_s	17.757 ± 0.021 ps^{-1}
Δm_d	0.5065 ± 0.0019 ps^{-1}

4.3.2 Unbinned maximum likelihood fit

A weighted maximum likelihood fit is performed to the decay time distribution, taking advantage of the sFit method. The fit is implemented using the RooFit framework [102] mentioned in Section 2.5, while the internal RooFit implementation of the MINUIT [103] package is used to perform the minimization. A toy Monte Carlo study is performed to validate the fit results, where different values of $\mathcal{A}^{\Delta}$ were generated.

Maximum likelihood

The maximum likelihood estimation is a widely known technique [118, 119] to estimate a set of unknown parameters θ given a set of individual observations x_i. The likelihood function is the probability to observe the dataset $\{x_i\}$ as a function of θ, constructed as the product of the individual probabilities:

$$\mathcal{L}(\theta) = \prod_{i}^{n} P(x_i|\theta). \tag{4.9}$$

The best estimator of the parameters θ are the values that maximize this function. In practice, algorithms usually minimize the negative logarithm of the likelihood:

$$-\ln \mathcal{L}(\theta) = -\sum_{i}^{n} \ln P(x_i|\theta), \tag{4.10}$$

since the summation is much easier to compute than the product and the result is more manageable, but otherwise the estimators are equivalent.

sFit and weighted likelihood

The likelihood function can be further modified to accomodate the usage of per-event weights, multiplying the logarithm of each event probability by the weight:

$$-\ln \mathcal{L}(\theta) = -\sum_{i}^{n} w_i \ln P(x_i|\theta). \tag{4.11}$$

In our case, w_i correspond to the sWeights calculated after the sPlot background-subtraction procedure, as explained in Section 3.4. The sFit method [120] states that the minimization of the Equation 4.11 provides unbiased estimators of θ because the total background contribution to the likelihood function cancels out statistically. The sFit method is an extension of the sPlot technique that simplifies the fitting procedure in the presence of background, since there is no need to model explicitly the decay time distributions of the different background sources. However, it produces underestimated uncertainties for θ; they are estimated as if all the events in the sample were signal, neglecting the dilution effects of the background subtraction. To reproduce the uncertainty correctly, the log-likelihood function has to be scaled [121]:

$$-\ln \mathcal{L}(\theta) = -\alpha \sum_{i}^{n} w_i \ln P(x_i|\theta). \tag{4.12}$$

The global factor α depends on the sum of weights:

$$\alpha = \frac{\sum_i^n w_i}{\sum_i^n w_i^2} \simeq \frac{N_s}{N_s^{\rm eff}}, \tag{4.13}$$

which corrects for the average statistical power per signal event. The average statistical power accounts for the effective signal yield of the sample, which in the limit of a large data sample is equivalent to $N_s^{\rm eff} = \sum_i^n w_i^2$ [122, 123].

Gaussian-constrained parameters

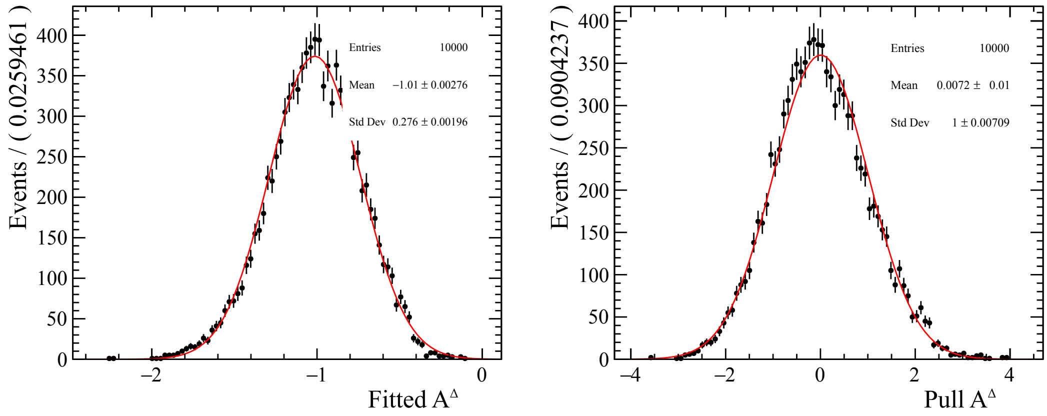
A Gaussian-constrained option can be activated in the fit to include the effect of the parameters that are fixed in the fit but known within a given precision. This option has been used to calculate the effect of the uncertainties of the external measurements, listed in Section 4.3.1, as well as the systematic uncertainty introduced by the acceptance parameters, detailed in Section 4.5. The Gaussian-constrained option adds additional terms to the log-likelihood expression, varying the input parameters y within their uncertainties:

$$-\ln \mathcal{L}(\theta) = -\sum_{i}^{n} \ln P(x_i|\theta) - \ln G(y, \sigma_y). \tag{4.14}$$

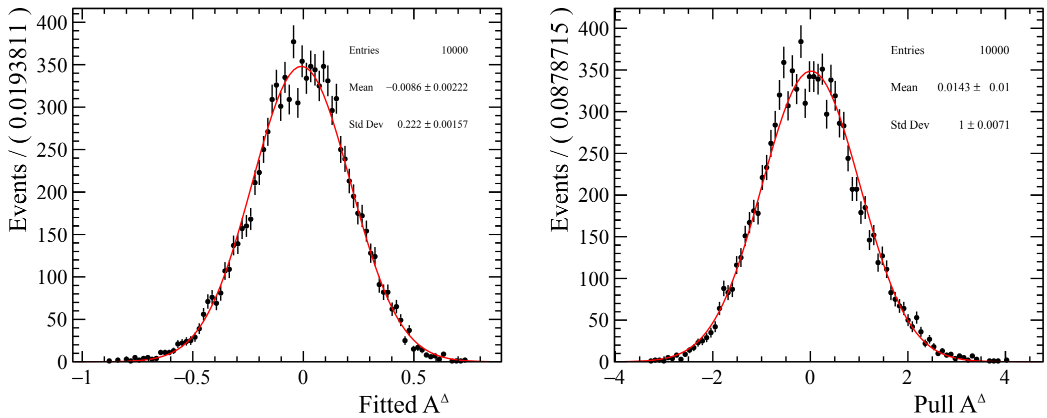
Multivariate Gaussian functions are used to consider the correlation between parameters.

Validation of the unbinned fit

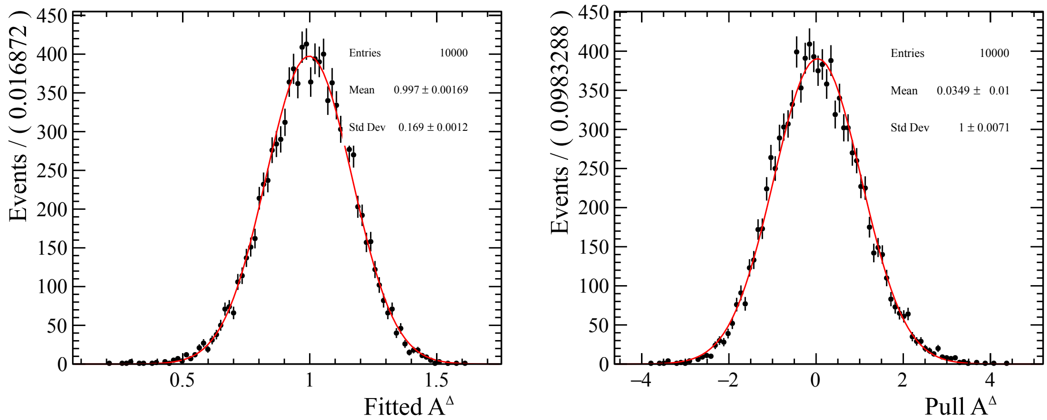
The validity of the fit has been tested with toy Monte Carlo experiments. Sets of 1 000 toys with a yield of $N_{B_s^0\to\phi\gamma} = 5000$ events are produced (without control channel events), where the decay-time distribution is fitted according to the PDF described in Section 4.3.1 and generated according to the



(a) $\mathcal{A}^{\Delta}_{\text{gen}} = -1$



(b) $\mathcal{A}^{\Delta}_{\text{gen}} = 0$



(c) $\mathcal{A}^{\Delta}_{\text{gen}} = 1$

Figure 4.2: Toy Monte Carlo experiments generated with three different values, $\mathcal{A}^{\Delta}_{\text{gen}} = \{-1, 0, 1\}$, performed with the unbinned fit, including acceptance and resolution effects. The dispersion of $\mathcal{A}^{\Delta}$ fitted values is shown at the left-hand side of the figure, while the pull distribution is shown at the right-hand side. A Gaussian fit is superimposed.

effect to be tested. The validation relies on the shape of the pull distribution:

$$\text{Pull} = \frac{\mathcal{A}^{\Delta}_{\text{fit}} - \mathcal{A}^{\Delta}_{\text{gen}}}{\sigma(\mathcal{A}^{\Delta}_{\text{fit}})}. \tag{4.15}$$

In a correctly behaving fit, the mean of the distribution should be centered at zero, and if the uncertainties are estimated correctly, the width of the distribution should be equal to unity.

Figure 4.2 shows the resulting distributions of pure-signal toys, including acceptance and resolution effects. The expected pull distribution is recovered when the generation and fitting is performed with the same PDF, showing no biases and a correct uncertainty estimation for the individual fits.

4.3.3 Binned ratio fit

A least-squares minimization fit is performed to the binned ratio of the $B^0_s \to \phi\gamma$ over $B^0 \to K^{*0}\gamma$ decay time distributions, using a custom-made framework [124] written in the `C++` programming language developed during the present thesis. It allows to fit binned distributions in general, either single distributions, ratios, or simultaneous fits. The framework makes use of the ROOT [101] library to manipulate the data and its implementation of the MINUIT [103] package to perform the minimization. The fit is verified with toy Monte Carlo experiments for different binning configurations.

The expected ratio yield

The estimator of $\mathcal{A}^{\Delta}$ in this case is obtained by minimizing the least squares quantity:

$$\chi^2 = \sum_i^{\text{bins}} \left[\frac{R_i^{\text{expected}} - R_i^{\text{observed}}}{\sigma_i^{\text{observed}}} \right]^2, \tag{4.16}$$

where R_i is the ratio of yields in the bin i. The *observed* quantities are extracted directly from the data samples, whereas the calculation of an unbiased *expected* yield R_i^{expected} is the motivation of this section. To compute the expected ratio yield in the i-th bin, the probability density functions are integrated within the bin boundaries t_i^{low} and t_i^{high}:

$$R_i^{\text{expected}} = \frac{\int_{t_i^{\text{low}}}^{t_i^{\text{high}}} P(B^0_s \to \phi\gamma)\, dt}{\int_{t_i^{\text{low}}}^{t_i^{\text{high}}} P(B^0 \to K^{*0}\gamma)\, dt}. \tag{4.17}$$

The explicit expression becomes, using Equation 4.7:

$$R_i^{\text{expected}} = \frac{\int_{t_i^{\text{low}}}^{t_i^{\text{high}}} A_{B_s^0\to\phi\gamma}(t)\, e^{-t/\tau_s} \left[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) - \mathcal{A}^{\Delta} \sinh\left(\frac{\Delta\Gamma_s t}{2}\right)\right]\, dt}{\int_{t_i^{\text{low}}}^{t_i^{\text{high}}} A_{B^0\to K^{*0}\gamma}(t)\, e^{-t/\tau_d}\, dt}, \tag{4.18}$$

where $A_{\text{decay}}(t)$ are the acceptances of the given decay. The same strategy can be used to fit a single channel, for example the $B^0 \to K^{*0}\gamma$ sample:

$$S_i^{\text{expected}} = \int_{t_i^{\text{low}}}^{t_i^{\text{high}}} A_{B^0\to K^{*0}\gamma}(t)\, e^{-t/\tau_d}\, dt. \tag{4.19}$$

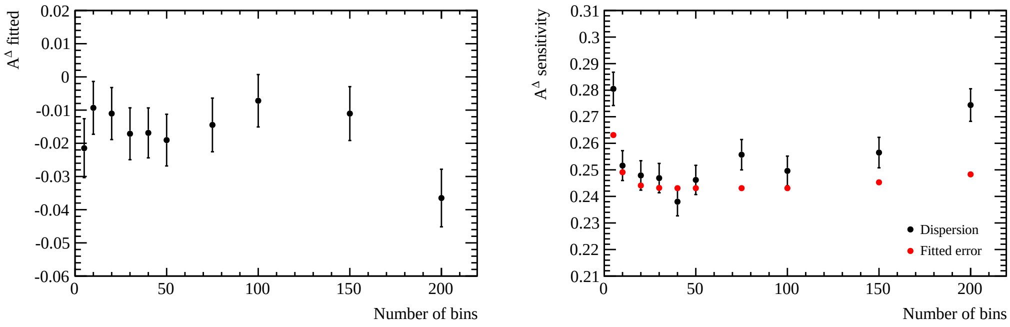
To facilitate the calculations, one could think of evaluating the PDF at the center of the bin t_i^{center}, instead of computing the integral. This approximation works fine for narrow bins. However, the difference is significant for wider bins, eventually leading to biases in the fit. From toy Monte Carlo studies, a bias in the $\mathcal{A}^{\Delta}$ parameter of about 0.2 was observed when simply using the center of the bin. Hence, the integral computations are maintained in the fit.

Binning scheme

Another relevant aspect to consider in the ratio fit is the choice of the binning scheme. To avoid low-statistics effects from the least squares minimization, a variable-sized bin width was constructed to ensure the same number-of-events per bin (equal-yield binning). This practice is convenient when the distribution is scarcely populated in some regions [125, 126], as it is the case in the $B^0_s \to \phi\gamma$ decay-time distribution at large times, where the tail of the exponential shape dominates, and at very small decay times due to selection inefficiencies.

The binning scheme with equal yields is constructed from Equation 4.7, after fixing $\mathcal{A}^{\Delta}$ to zero. Since the PDF shape is dominated by the exponential and acceptance functions, the sensitivity loss of making such assumption is negligible. A percentile mesh is constructed, using the `GetQuantiles` routine of the `TH1` class of ROOT to compute the percentiles of a histogram distribution. The binning is chosen so that the increase in percentile is equal between consecutive bins, effectively creating bins with equal areas.

The robustness of the adaptive binning strategy is assessed from Monte Carlo pseudoexperiments by changing the number of bins, as shown in Figure 4.3. A small bias of -0.01 is observed in the dispersion mean of the



(a) Fitted $\mathcal{A}^{\Delta}$ mean.

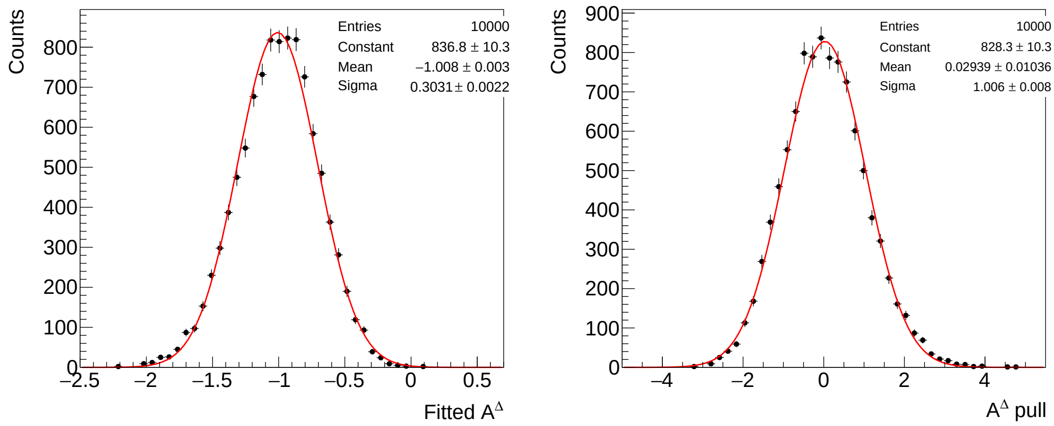
(b) $\mathcal{A}^{\Delta}$ uncertainty.

Figure 4.3: Toy Monte Carlo experiments to validate the ratio fit, indicating the (a) mean and (b) uncertainty of the fitted $\mathcal{A}^{\Delta}$ values against the number of bins on the equal-yield binning. Each point represent a set of 1 000 pseudoexperiments, generated with $\mathcal{A}^{\Delta}_{\text{gen}} = 0$, $N_{B^0_s \to \phi\gamma} = 5000$ and $N_{B^0 \to K^{*0}\gamma} = 29\,000$.

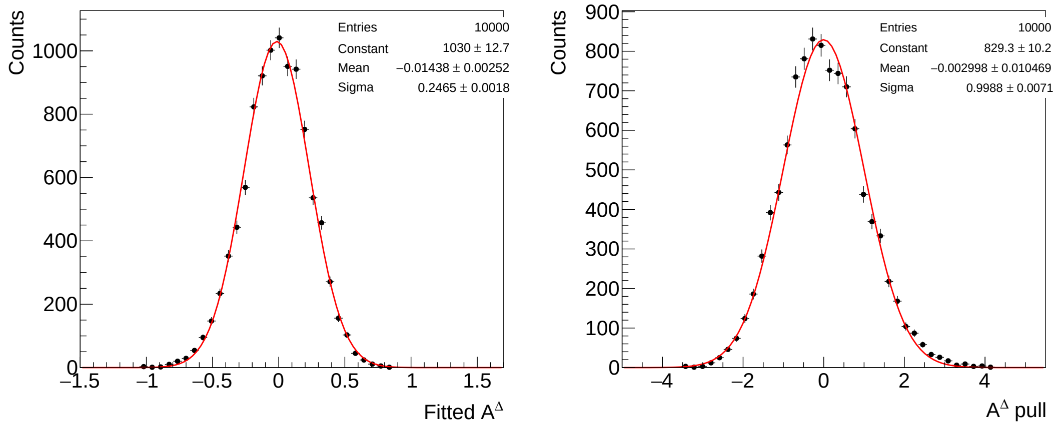
fitted $\mathcal{A}^{\Delta}$ values, as shown in Figure 4.3 (a), due to the asymmetry of the $\mathcal{A}^{\Delta}$ dispersion distribution. The sensitivity to $\mathcal{A}^{\Delta}$ is a function of the number of bins n, which is maximized in the range $n = [20, 100]$, as can be seen in Figure 4.3 (b). Binning schemes with $n < 10$ or $n > 100$ are not advisable, since the dispersion of the fitted $\mathcal{A}^{\Delta}$ values is not compatible with the fitted uncertainty $\sigma(\mathcal{A}^{\Delta})$. With too few bins ($n < 10$), the histogram cannot reproduce the shape of the distribution due to oversmoothing. With too many bins ($n > 100$), low-statistics effects appear. An intermediate number of bins $n = 30$ guarantees a fit with no biases and a correctly evaluated uncertainty lying on the plateau of maximum sensitivity.

Validation of the ratio fit

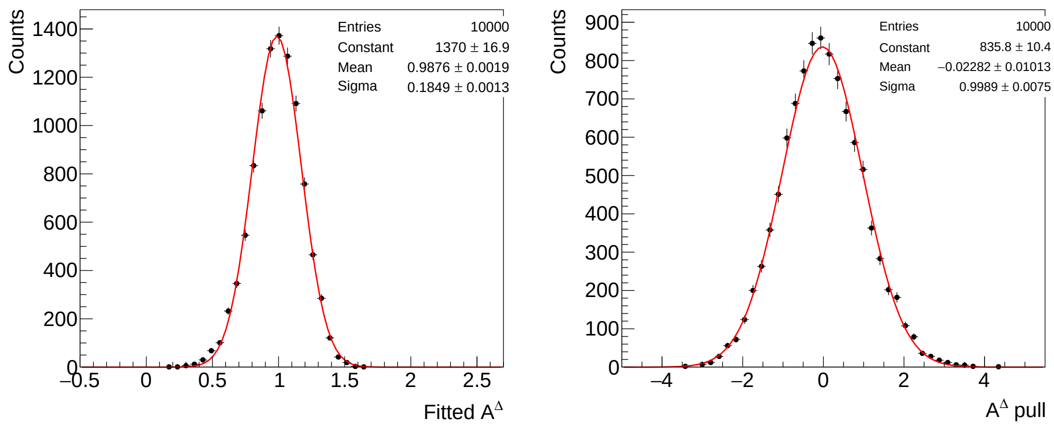
The ratio fit is tested with toy Monte Carlo experiments, fixing the number of bins of the adaptive binning to $n = 30$ and the yields to $N_{B^0_s \to \phi\gamma} = 5000$ and $N_{B^0 \to K^{*0}\gamma} = 29\,000$. Sets of 10 000 simulated experiments were generated for three different generation values $\mathcal{A}^{\Delta} = \{-1, 0, 1\}$, obtaining the dispersion and pull distributions of the fitted $\mathcal{A}^{\Delta}$ values as it is shown in Figure 4.4. The expected pull distributions are recovered, with offsets and widths compatible with 0 and 1, respectively, demonstrating a correct



(a) $\mathcal{A}^{\Delta}_{\text{gen}} = -1$



(b) $\mathcal{A}^{\Delta}_{\text{gen}} = 0$



(c) $\mathcal{A}^{\Delta}_{\text{gen}} = 1$

Figure 4.4: Simulated experiments generated with three different $\mathcal{A}^{\Delta}_{\text{gen}}$ values, performed with the ratio fit. The dispersion of $\mathcal{A}^{\Delta}$ fitted values is shown at the left-hand side, while the pull distributions are shown at the right-hand side of the figure. A Gaussian fit is superimposed.

behaviour for the individual fits. The fit is also tested using a likelihood function for the minimization instead of the nominal least squares, showing no significant differences.

Comparison with the unbinned ML fit

Compared to an unbinned likelihood strategy, a binned solution is computationally faster. The CPU time spent by the algorithm is proportional to the number of bins (total number of integrals to compute) regardless of the sample size, whereas the unbinned fit gets slower with the sample size since the likelihood function has to be calculated for every candidate. This property allows a faster production of toy Monte Carlo experiments, useful to check the validity of the results, control the behaviour of the fit and study systematic effects with great precision. In addition, the statistical dilution introduced by the background subtraction procedure is correctly taken into account in the uncertainties of each bin yield. However, it provides a slightly smaller sensitivity compared to the unbinned fit, due to the inevitable loss of information that occurs when binning a data sample. For $\mathcal{A}^{\Delta} = 0$, the statistical degradation is calculated to be $\Delta\sigma_{\mathcal{A}^{\Delta}} = 0.02$, a factor of magnitude smaller compared to the expected statistical uncertainty.

Dispersion between fit strategies A study is performed to estimate the expected $\mathcal{A}^{\Delta}$ dispersion between the unbinned fit and the ratio fit strategies. 1 000 toy Monte Carlo experiments are generated, and fitted with the two strategies. For the case $\mathcal{A}^{\Delta} = 0$, the $\mathcal{A}^{\Delta}$ dispersion between the two methods is evaluated to be Gaussian distribution with a width of 0.14.

4.4 Sensitivity studies and cross-checks

4.4.1 Dependency $\sigma_{\mathcal{A}^{\Delta}}$ vs $\mathcal{A}^{\Delta}$

This section aims to study the dependency of the uncertainty $\sigma_{\mathcal{A}^{\Delta}}$ with the $\mathcal{A}^{\Delta}$ input value and with the number of signal events. The statistical uncertainties are evaluated using simulated experiments, generating an experiment and then fitting it with the same decay time PDF using the unbinned fit strategy. All the studies presented in this section are produced with 10 000 pseudoexperiments and $N_{B^0_s \to \phi\gamma} = 5000$ signal events, while the number of $B^0 \to K^{*0}\gamma$ events is set to $N_{B^0 \to K^{*0}\gamma} = 6\, N_{B^0_s \to \phi\gamma}$.

Five sets of toys are generated for different values of $\mathcal{A}^{\Delta}_{\text{gen}} = \{-1, -0.5, 0, 0.5, 1\}$. The $\mathcal{A}^{\Delta}$ and pull distributions are shown in Figure 4.2, showing an unbiased mean and a correctly evaluated uncertainty in all the scenarios. As it is observed from the figures, the uncertainty $\sigma_{\mathcal{A}^{\Delta}}$ depends on the $\mathcal{A}^{\Delta}$ value used to generate the pseudoexperiments. The sensitivity is higher for positive generated values, and lower for negative ones, following a linear dependency $\sigma_{\mathcal{A}^{\Delta}} = p_0 + p_1\mathcal{A}^{\Delta}$ that has been explicitly evaluated in Figure 4.5.

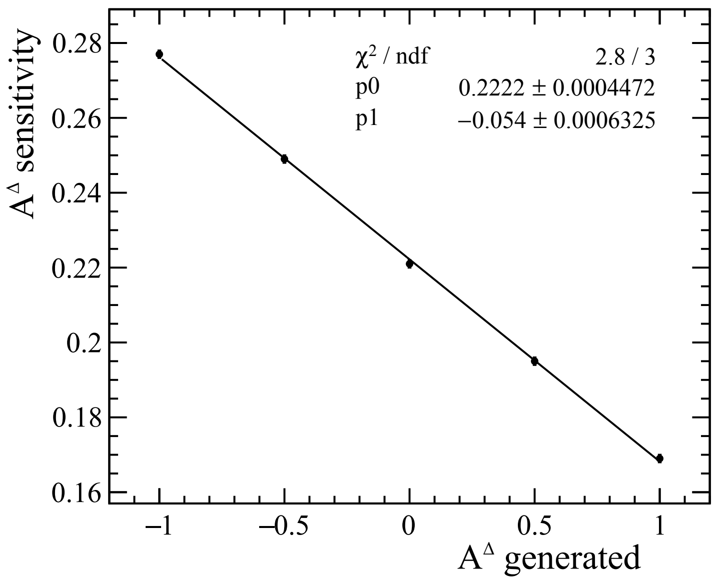


Figure 4.5: Uncertainty $\sigma_{\mathcal{A}^{\Delta}}$ as a function of the generated $\mathcal{A}^{\Delta}_{\text{gen}}$ value, evaluated from simulated experiments with $N = 5000$. The uncertainty depends linearly with the generated value.

The uncertainty not only depends on $\mathcal{A}^{\Delta}$, but also on the number of generated signal events $N_{B_s^0 \to \phi\gamma} = 5000$. To correctly determine the dependency with both $\mathcal{A}^{\Delta}$ and the statistics of the data sample, different values of the pair (N_{events}, $\mathcal{A}^{\Delta}_{\text{gen}}$) are generated. The statistical uncertainty of $\mathcal{A}^{\Delta}$ in the unbinned fit strategy can be described by:

$$\sigma_{\text{signal stat.}}(\mathcal{A}^{\Delta}) = \frac{15.63 - 3.82\,\mathcal{A}^{\Delta}}{\sqrt{N}}. \tag{4.20}$$

4.4.2 Contributions to the uncertainty

The total sensitivity of the measurement is a sum of several contributions:

$$\sigma_{\mathcal{A}^{\Delta}} = \sqrt{\sigma^2_{\text{signal stat.}} + \sigma^2_{\text{sweights}} + \sigma^2_{\text{external}} + \sigma^2_{\text{systematics}}}. \tag{4.21}$$

- $\sigma_{\text{signal stat.}}$ is the statistical uncertainty from the signal statistics. It is estimated with simulated experiments in this section, ranging from 0.17 to 0.28 within the considered scenarios (different $\mathcal{A}^{\Delta}$ generated values). The contribution introduced by the control channel $B^0 \to K^{*0}\gamma$ is calculated to be 0.13.

- σ_{sweights} is the dilution introduced by the ${}_s\mathcal{P}lot$ background subtraction. This contribution is estimated to be $\sigma_{\text{sweights}} \simeq 0.22$, and it is included in the statistical uncertainty:

$$\sigma_{\text{stat.}} = \sqrt{\sigma_{\text{signal stat.}} + \sigma_{\text{sweights}}}. \tag{4.22}$$

  In the unbinned fit this contribution is estimated by fitting without the correcting factor α in the likelihood function. In the ratio fit it is estimated by deactivating directly the sWeight uncertainty correction, meaning that the histogram option `TH1::Sumw2` is not called before the fit. Both methods agree in the size of the contribution.

- σ_{external} is the contribution from the input parameters taken from external measurements, known within a given uncertainty (see Table 4.4). It is quoted separately from the statistical and systematic uncertainties in the final result. In the unbinned fit strategy, σ_{external} is estimated for each parameter by introducing a Gaussian-constraint term in the likelihood function, while in the ratio fit strategy they are calculated by generating and fitting simulated toy experiments.

- $\sigma_{\text{systematics}}$: several sources of systematic uncertainties are considered. For the breakdown of the different contributions see Section 4.5.

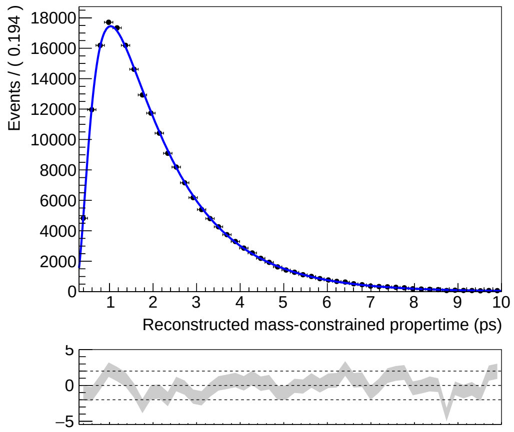
4.4.3 Cross-check with full-simulation samples

The procedure is also validated by fitting $\mathcal{A}^{\Delta}$ in the LHCb simulation samples, along with the acceptance parameters. The configuration is the same as in data, fitting the reconstructed decay time simultaneously in $B^0_s \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ while fixing the acceptance difference $\Delta A_{\phi\gamma-K^*\gamma}$. The input parameters of the decay rate are fixed to those used in the production, listed in Table 4.2.

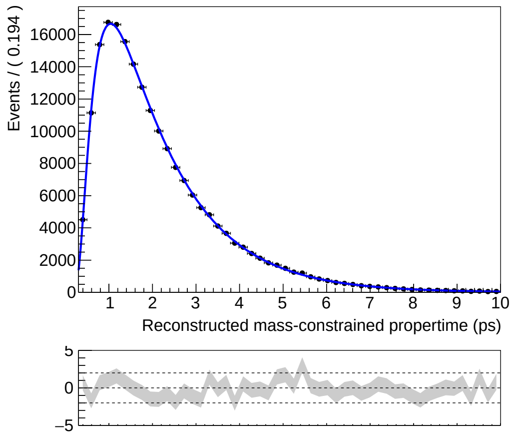
The result of the fit is represented in Figure 4.6. The fitted $\mathcal{A}^{\Delta}$ value is compatible with zero within the statistical uncertainty, as shown in Table 4.5 for the two acceptance descriptions. The statistical uncertainty of 0.05 for $N = 90\,000$ simulated signal candidates is compatible with the expectation estimated from Equation 4.20. The same study is repeated, but fitting the true decay-time distribution and modifying the acceptance difference accordingly (not including resolution effects). A value of $\mathcal{A}^{\Delta} = 0.03 \pm 0.05$ ($\mathcal{A}^{\Delta} = 0.00 \pm 0.05$) is found for the 2-parameter (4-parameter) acceptance, results which are also compatible with the generated value.

Table 4.5: Fit to the LHCb full-simulation samples with the unbinned fit strategy, using both the $A^{(2\text{par})}$ and $A^{(4\text{par})}$ acceptance definitions. The quoted acceptance parameters correspond to $B^0 \to K^{*0}\gamma$, while for $B^0_s \to \phi\gamma$ the acceptance is fixed to $A_{K^{*0}\gamma} + \Delta A_{\phi\gamma-K^*\gamma}$ (see Table 4.3 and Section 4.2.2).

Parameter	Simulation, $A^{(2\text{par})}$	Simulation, $A^{(4\text{par})}$
$\mathcal{A}^{\Delta}$	0.05 ± 0.05	-0.01 ± 0.05
b	0.7367 ± 0.0034	–
c (ps^{-1})	0.0481 ± 0.0045	–
a	–	1.898 ± 0.027
n	–	2.154 ± 0.064
t_0	–	0.185 ± 0.009
$\delta\Gamma$	–	0.0313 ± 0.0023



(a) $B_s^0 \to \phi\gamma$ simulation



(b) $B^0 \to K^{*0}\gamma$ simulation

Figure 4.6: Fit to the decay-time distributions of simulation samples.

4.5 Systematic uncertainties

This section is devoted to the study of the systematic effects that could modify the final result. Depending on the effect, the systematic contribution can be calculated in one of the following ways: by repeating the fit with a given variation and considering the observed change in $\mathcal{A}^{\Delta}$ as a systematic uncertainty, or by including the uncertainties of the input parameters as a Gaussian constraint in the likelihood fit.

The results are presented in Table 4.6, while the relevant contributions are explained in detail in the following.

4.5.1 External measurements

The input parameters listed in Section 4.3.1 are taken from world averages [39], and hence they have an uncertainty associated to them. The contribution from the external measurements is quoted separately from the rest of the systematic uncertainties in the final result.

To account for these effects, the input parameters are fixed to their central values, while the Gaussian-constrained option is activated in the unbinned fit, as explained in Section 4.3.2. The correlation $\rho(\Delta\Gamma_s, \Gamma_s)$ between the parameters is included in a multi-Gaussian description, as shown in Figure 4.8 (a). The $\mathcal{A}^{\Delta}$ uncertainty with and without the Gaussian-constrained

Table 4.6: Summary of the systematic uncertainties in the untagged $\mathcal{A}^{\Delta}$ measurement, evaluated with the unbinned fit strategy.

	Systematic source	$\sigma(\mathcal{A}^{\Delta})$
External measurements	$(\Gamma_s, \Delta\Gamma_s)$	0.067
	Γ_d	0.040
Decay time	Acceptance: MC statistics	0.067
	Acceptance: modelling	0.020
	Resolution: modelling	0.020
Background subtraction	Mass modelling: signal	< 0.01
	Mass modelling: combinatorial	0.034
	Mass modelling: partial	0.093
	Peaking backgrounds	0.036
	Mass-time correlation	0.110
Total	$\sigma_{\text{ext.}}$	0.078
	$\sigma_{\text{syst.}}$	0.170

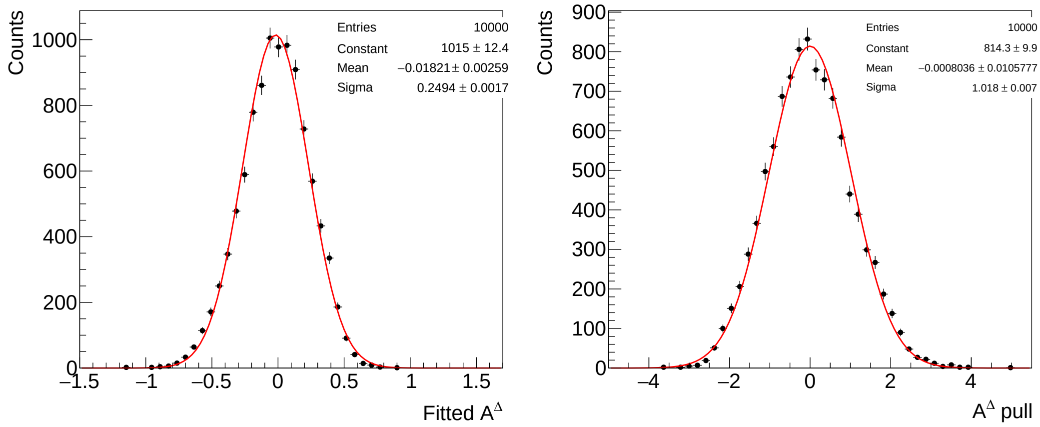
option is evaluated and subtracted in quadrature $\sqrt{\sigma^2_{\text{with}} - \sigma^2_{\text{without}}}$ to obtain the individual contributions.

The individual contribution of Γ_d and $(\Delta\Gamma_s, \Gamma_s)$ has been evaluated to be 0.040 and 0.067, respectively, making a total of:

$$\sigma_{\text{ext.}}(\mathcal{A}^{\Delta}) = 0.078 \tag{4.23}$$

for the external measurements. The uncertainties are symmetrized for simplicity, although a certain degree of asymmetry (similar to the statistical uncertainty) is observed.

As a cross-check, the external systematic uncertainties are calculated using an alternative method that involves the ratio fit. Several thousands of toy experiments are produced with and without generating randomly the external parameters, generated with the $\mathcal{A}^{\Delta}$ value fitted in data (while kept blinded). The generation of the random parameters is performed using a custom-made algorithm which uses the Cholesky decomposition algorithm, implemented in the ratio code [124]. The result of this method is presented in Figure 4.7. The total dispersion of the fitted values are subtracted in quadrature to obtain the systematic contribution, calculated to be around 0.04. This number is compatible with the numbers obtained with the unbinned fit.



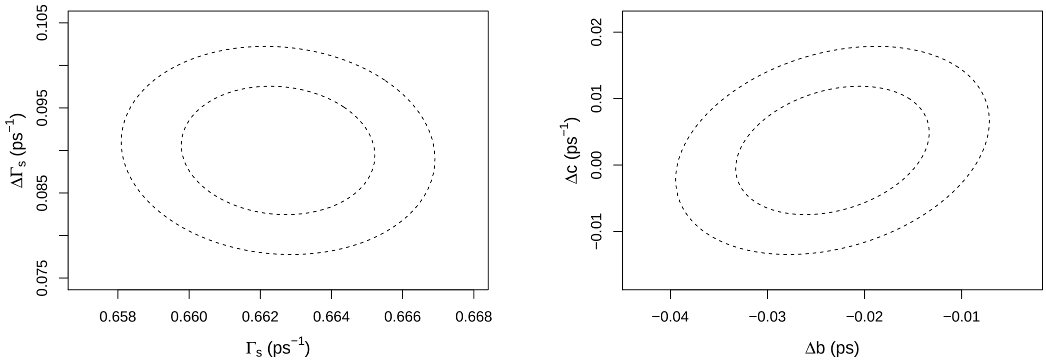
(a) $\mathcal{A}^{\Delta}$ dispersion (b) $\mathcal{A}^{\Delta}$ pull

Figure 4.7: Toy Monte Carlo experiments performed with the ratio fit, varying the external parameters within their uncertainty in the generation. It corresponds to the case $\mathcal{A}^{\Delta}_{\text{gen}} = 0$.

4.5.2 Decay time acceptance

Two sources of systematic uncertainties arise from the description of the decay time acceptance: the limited statistics of simulated events, and the modelling of the distribution. The first one is evaluated through a Gaussian-constrained fit, in the same way as the external measurements. The parameters of the acceptance difference ΔA, extracted from simulation and previously shown in Section 4.2.2, are varied within their uncertainties taking into account the correlations, given by $\rho(\Delta b, \Delta c) = 0.28$ for the $A^{(2\text{par})}$ acceptance. Figure 4.8 shows the correlation plot between Δb and Δc. The contribution of this source of systematic uncertainty is evaluated to be 0.067.

The second contribution is related to the modelling of the acceptance distribution, evaluating the possibility that the acceptance function does not reproduce the distribution perfectly. It is evaluated by fitting $\mathcal{A}^{\Delta}$ with the $A^{(2\text{par})}$ and $A^{(4\text{par})}$ acceptances. The deviation $\Delta\mathcal{A}^{\Delta} = 0.02$ is taken as the



(a) Correlation $\rho(\Gamma_s, \Delta\Gamma_s) = -0.08$ (b) Correlation $\rho(\Delta b, \Delta c) = 0.28$

Figure 4.8: Correlation of the parameters entering in the Gaussian-constrained fit to extract the systematic uncertainty for (a) the inputs from the external measurements and (b) the acceptance difference parameters from simulation. The ellipses correspond to the contours of one and two standard deviations.

systematic uncertainty, a factor three smaller than the systematic associated with the limited statistics of the simulation.

Cross-check: eliminate the ΔA correction

To test the effect of the fixed acceptance difference ($\Delta A = A_{\phi\gamma} - A_{K^*\gamma}$) in the measurement, $\mathcal{A}^{\Delta}$ is fitted in data with and without the ΔA constraints:

- Using the $A^{(4\text{par})}$ acceptance, the value changes $\Delta\mathcal{A}^{\Delta} = -0.11$ (with Δa contributing -0.14 and $\Delta\delta\Gamma$ contributing $+0.03$)
- Using the $A^{(2\text{par})}$ acceptance it changes $\Delta\mathcal{A}^{\Delta} = -0.09$ (with Δb contributing -0.07 and Δc contributing -0.02).

Note that this correction is present in the fit, so this effect is not taken as a systematic effect.

4.5.3 Decay time resolution

To determine the effect of the decay time resolution, a set of toy Monte Carlo experiments generated with different resolution configurations is run, but fitted with the nominal configuration (without resolution). The $\mathcal{A}^{\Delta}$ dispersion and pull distributions are shown in Figure 4.9. Generating with a resolution width does not produce any effect on the result ($\Delta\mathcal{A}^{\Delta} < 0.01$) but the resolution mean does ($\Delta\mathcal{A}^{\Delta} = -0.01$ for each $\Delta\mu = 1$ fs). However, the resolution bias is a reconstruction effect that is absorbed completely in the acceptance $A(t)$ defined as a function of the reconstructed decay time.

The fit is repeated using the acceptance difference defined as a function of the reconstructed decay time, $\Delta A(t)$, instead of the nominal with the truth decay time, $\Delta A(t^{'})$. The observed variation, $\Delta\mathcal{A}^{\Delta} = +0.02$, is taken as the systematic uncertainty, although it represents an upper limit of the resolution systematic.

4.5.4 Background subtraction

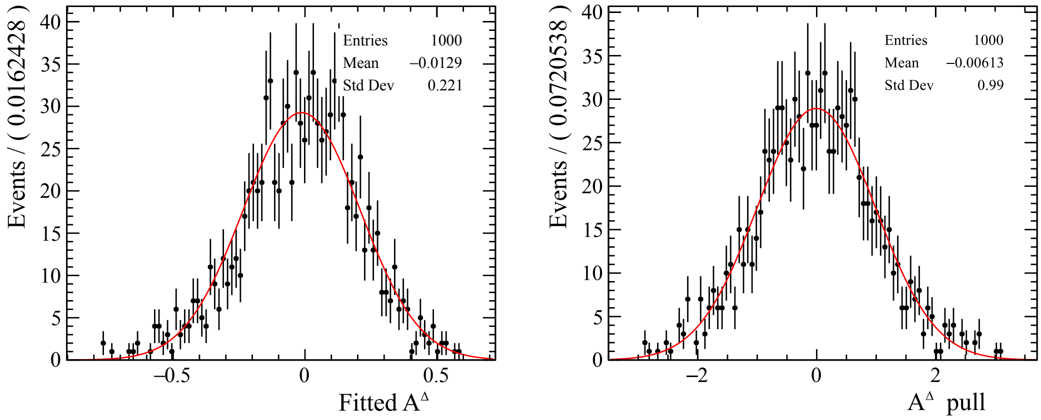
The assumptions made in the background subtraction procedure (explained in Section 3.4) can systematically influence the final result. Several sources of systematic effects are considered: the modelling of the different mass components (signal, combinatorial and partially reconstructed backgrounds), the contribution of the peaking backgrounds in the decay time distribution, and the possible correlation between the mass and the decay time variable in background decays.

Mass modelling

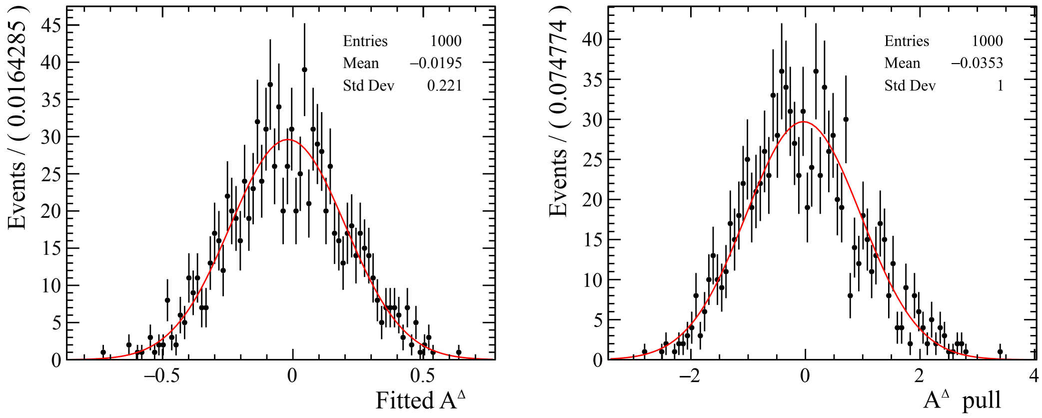
This effect accounts for the assumption of the functions used to describe the mass distributions:

- For the signal component, the background subtraction procedure is repeated using an asymmetric Apollonios function, instead of the nominal double-tailed Crystal Ball. The difference between the two methods, $\Delta\mathcal{A}^{\Delta} = 0.011$, is taken as the systematic uncertainty.

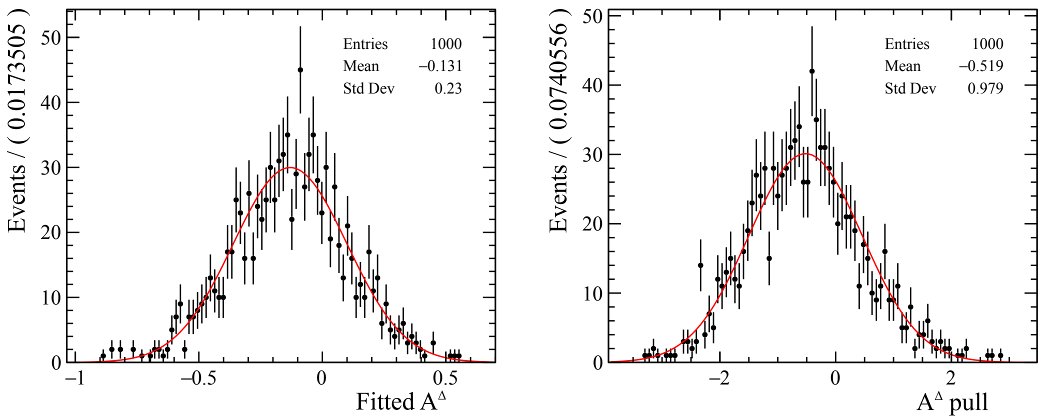
- For the combinatorial component, an exponential function is used instead of the nominal linear polynominal. The observed deviation,



(a) No resolution: $R(0,0)$ fs



(b) Resolution width only: $R(0,60)$ fs



(c) Resolution mean and width: $R(10,60)$ fs

Figure 4.9: Distribution of $\mathcal{A}^{\Delta}$ fitted values (left) and its pull distribution (right), obtained from toy Monte Carlo experiments generated with $\mathcal{A}^{\Delta}_{\rm gen} = 0$ and different resolution configurations: (a) without resolution, (b) including resolution in the generation $\mu = 0$ fs, $\sigma = 60$ fs, and (c) including a biased resolution in the generation $\mu = 10$ fs, $\sigma = 60$ fs.

$\Delta\mathcal{A}^{\Delta} = -0.029$, is taken as the systematic uncertainty.

- For the partially reconstructed backgrounds, the procedure is different for each of them:
  - For the $B^+ \to \phi K^+\gamma$ partial in $B_s^0 \to \phi\gamma$ decays, the background subtraction is repeated extracting the mass lineshape from a full-simulation of $B^0 \to \phi K_s^0\gamma$ decays, instead of the nominal channel. A deviation $\Delta\mathcal{A}^{\Delta} = 0.046$ is observed.
  - For the $B \to K\pi\pi^0 X$ partial in $B^0 \to K^{*0}\gamma$ decays, the mass lineshape is extracted from a full-simulation of $K_1\gamma$ decays, instead of a mixture of $K_1\gamma$ and $B \to D\rho$ decays. A deviation $\Delta\mathcal{A}^{\Delta} = -0.075$ is observed.
  - For the *missing pion* partial in $B^0 \to K^{*0}\gamma$ decays, the mass lineshape is extracted from a full-simulation of $B^+ \to K^*(1410)^+(\to K^{*0}\pi^+)\gamma$ decays, instead of $B \to K_1(1270)(\to K^{*0}\pi)\gamma$ decays. A deviation of $\Delta\mathcal{A}^{\Delta} = -0.030$ is observed.
  - For the partial $B^0 \to K^{*0}\eta$ partial in $B^0 \to K^{*0}\gamma$ decays, the component is totally removed (which will be absorbed by the other components), instead of fixing its yield. A deviation of $\Delta\mathcal{A}^{\Delta} = -0.006$ is observed.

  The total contribution from partially reconstructed decays is:

$$\sigma_{\text{partial}}(\mathcal{A}^{\Delta}) = 0.093.$$

Peaking backgrounds

In the nominal sPlot mass fit, the peaking background contribution is treated as part of the signal mass component. Therefore, the decay time contribution of the peaking backgrounds is not subtracted, and although the yield fraction of this type of background is only around 2% of the signal yield, they could induce an effect on the final result. The systematic effect is evaluated fitting simultaneously the Monte Carlo samples of $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$, adding a peaking background contribution which is $\times 2$ the expected yield in data. The observed deviation $\Delta\mathcal{A}^{\Delta} = 0.036$ is taken as the systematic contribution.

Mass-time correlation

The assumption that the reconstructed mass and the decay time variable are uncorrelated is made during the sPlot procedure. These assumption is accurate for the signal and combinatorial components, but a residual correlation is present in the partially reconstructed backgrounds.

The systematic effect is calculated using the Monte Carlo samples. The sPlot procedure is applied to a mixture of the signal and partial components constructed with the yields observed in data. The deviation observed in the fit between the original mixture (which conserves the correlations) and a mixture in which the events of the partial sample are shuffled (breaking the correlation), is taken as the systematic contribution, $\Delta\mathcal{A}^{\Delta} = -0.11$.

4.5.5 Possible improvements

The systematic uncertainty could be reduced by:

- Increasing the Monte Carlo statistics of the signal, control and background channels. This would improve the decay time acceptance determination, and the description of the expected mass distributions for background subtraction.

- Reduce the amount of combinatorial and physical background candidates. This can be achieved by: (1) tightening the cut-based offline selection, or even using a Boosted Decision Tree (BDT) to select the candidates, and (2) performing a veto to the physical backgrounds that produce the larger effects in the systematic contribution.

4.6 Results

The results are presented in Table 4.7, for the unbinned and ratio fit strategies. The parameters of the acceptance for the control channel, $A_{K^{*0}\gamma}$, are also quoted. The unbinned and ratio fits are shown in Figures 4.10 and 4.11, respectively, where the pull distributions do not show any particular pattern.

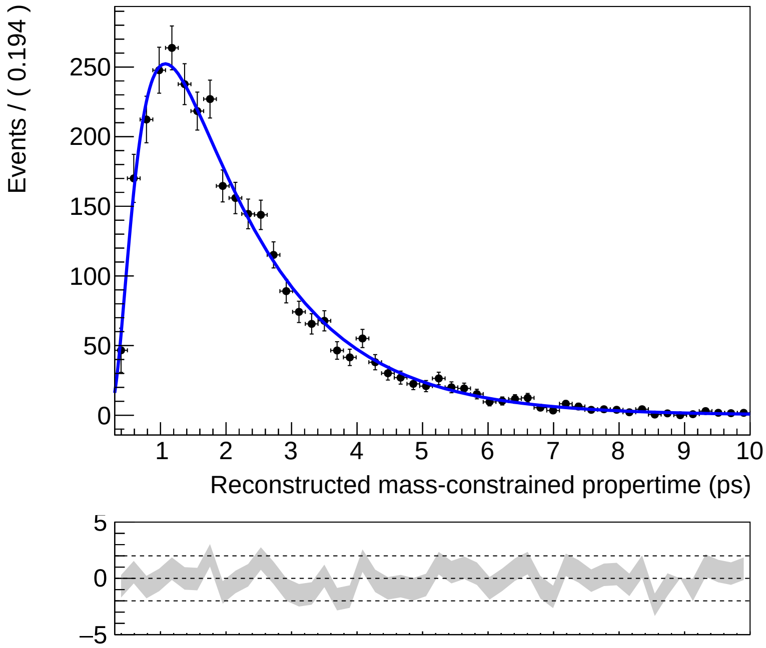
Table 4.7: Results of the $\mathcal{A}^{\Delta}$ measurement, in a simultaneous fit to $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ data using the unbinned and ratio fit strategies. The acceptance values correspond to $A_{K^{*0}\gamma}$.

Parameter	Unbinned fit	Ratio fit
$\mathcal{A}^{\Delta}$	$-0.555^{+0.331}_{-0.359}$	$-0.473^{+0.321}_{-0.343}$
a (ps^{-1})	1.88 ± 0.08	–
n	1.80 ± 0.16	–
t_0 (ps)	0.25 ± 0.02	–
$\delta\Gamma$ (ps^{-1})	0.052 ± 0.009	–

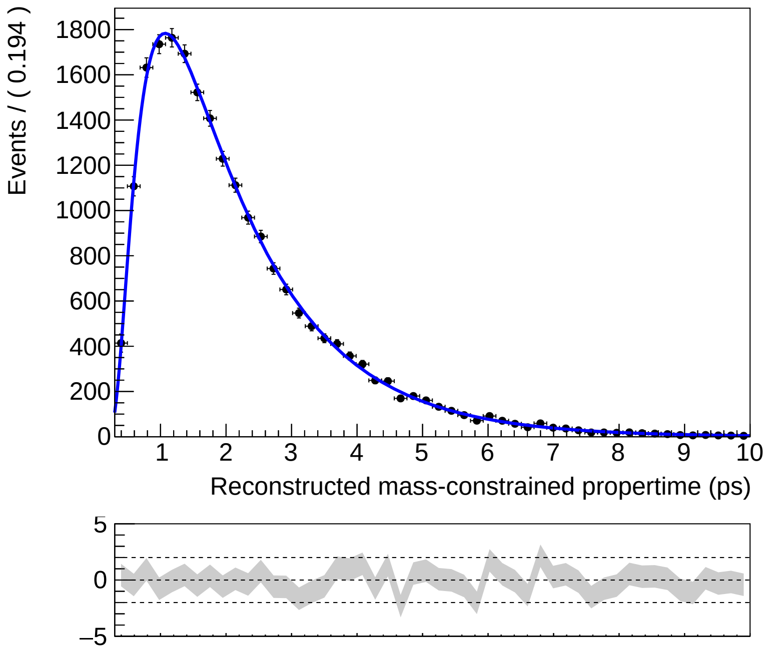
Considering the systematic uncertainties, the result for the unbinned fit can be expressed as:

$$\mathcal{A}^{\Delta}_{\phi\gamma} = -0.555\,^{+0.331}_{-0.359}(\text{stat.}) \pm 0.170(\text{syst.}) \pm 0.078(\text{ext.}), \tag{4.24}$$

where the statistical uncertainty contributions can be splitted in: 0.22 from the signal channel statistics, another 0.22 from the background subtraction, and 0.13 from the control channel statistics. At this point, the measurement is completely dominated by the statistical uncertainty.



(a) $B^0_s \to \phi\gamma$



(b) $B^0 \to K^{*0}\gamma$

Figure 4.10: Simultaneous unbinned fit to the decay time distributions of (a) $B^0_s \to \phi\gamma$ and (b) $B^0 \to K^{*0}\gamma$ decays.

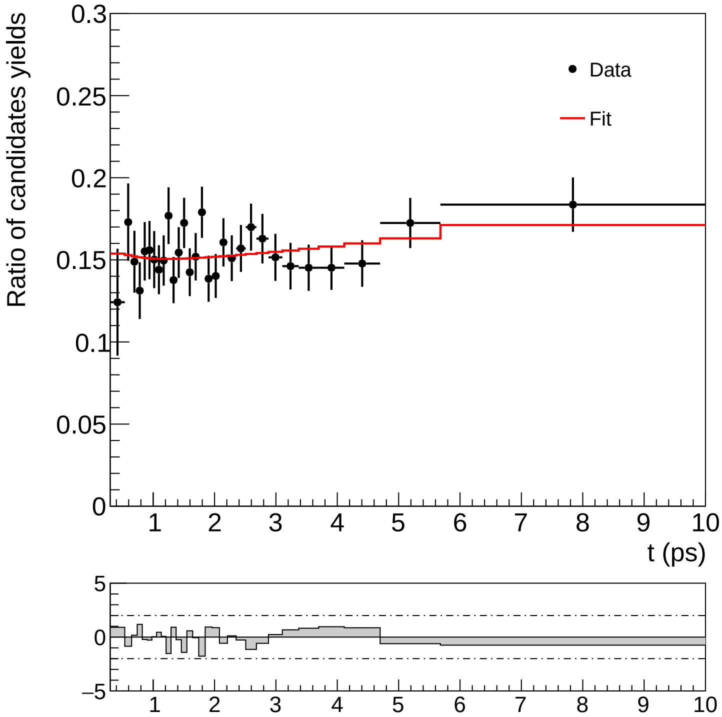


Figure 4.11: Fit to the ratio of $B_s^0 \to \phi\gamma$ over $B^0 \to K^{*0}\gamma$ decay time distributions. The best estimation is represented by the red line.

5

Measurement of $S_{\phi\gamma}$ and $C_{\phi\gamma}$ in $B_s^0 \to \phi\gamma$ decays

This chapter describes the procedure for the measurement of the observables $S_{\phi\gamma}$ and $C_{\phi\gamma}$, from a simultaneous fit to the time-dependent decay rates of $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays. It is the first time these observables are measured, and the first radiative analysis including flavour tagging information at LHCb.

The fundamental aspects of the analysis strategy are summarized in Section 5.1. Two key elements in this measurement are the flavour tagging and the decay time resolution, described in Sections 5.2 and 5.3, respectively. Section 5.4 details the fit strategy and the probability density function of the decay time, while Section 5.5 presents the sensitivity study to the measured observables. Section 5.6 discusses and enummerates the different systematic uncertainties, and Section 5.7 presents the preliminary results of the measurement.

5.1 Analysis strategy

- The datasets for real and simulated data are the same as the ones described in the untagged analysis chapter (see Section 4.1), as well as the reconstruction and selection of candidates, and the background subtraction procedure (see Chapter 3).

- Flavour tagging information is included in the samples. The output of several tagging algorithms (named OSComb and SSKaonNNet) is calibrated and combined into a single decision. An effective tagging efficiency of $\epsilon_{\text{power}} = 4.58\%$ is achieved for $B_s^0 \to \phi\gamma$ decays.

- The complete description of the decay time resolution is needed to perform the fit:
  - A double Gaussian function is used to describe the decay time resolution, employing the per-event decay time uncertainty σ_t.
  - The uncertainty σ_t is scaled by a global factor, obtained from Monte Carlo simulated data, to match the expected resolution.
  - The resolution is controlled with real data by using: (1) prompt-ϕ candidates and (2) $B^0 \to J\psi K^*$ decays where the J/ψ is faked as a photon.
- An unbinned maximum likelihood fit is performed simultaneously to the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ samples, incorporating flavour tagging and decay time resolution information:
  - As in the untagged analysis, the $B^0 \to K^{*0}\gamma$ data sample is used to constrain the decay time acceptance function. The acceptance difference between the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decay channels is fixed from Monte Carlo simulation.
  - The mistag probability ω and the decay time uncertainty σ_t are taken from data, and treated as per-event observables.
  - The PDF of the decay time is constructed from the convolution of the physical decay rate with an acceptance-resolution function (Equation 5.23), which consists of a resolution function multiplied by a binned acceptance function. The fit performs the convolution and normalization integrals analytically.
  - The mixing-induced and CPV observables ($\mathcal{A}^{\Delta}$, $\mathcal{C}$, $\mathcal{S}$) and the decay time acceptance parameters (b and c) are the free parameters of the fit.
- All the analysis strategy was performed maintaining blinded the value of the measured observables.

5.2 Flavour Tagging

There are several algorithms available at LHCb to identify the flavour of the initial B meson on an event [127–130]. They are called taggers, and exploit the fact that the heavy quarks produced in the proton-proton collision are created in $q\bar{q}$ pairs, each hadronizing into a b-hadron. The flavour-tagging algorithms can be categorized into same-side (SS) and opposite-side (OS), depending on whether they use the information of the signal or of the accompanying b-hadron, respectively. A sketch of the available flavour taggers is depicted in Figure 5.1.

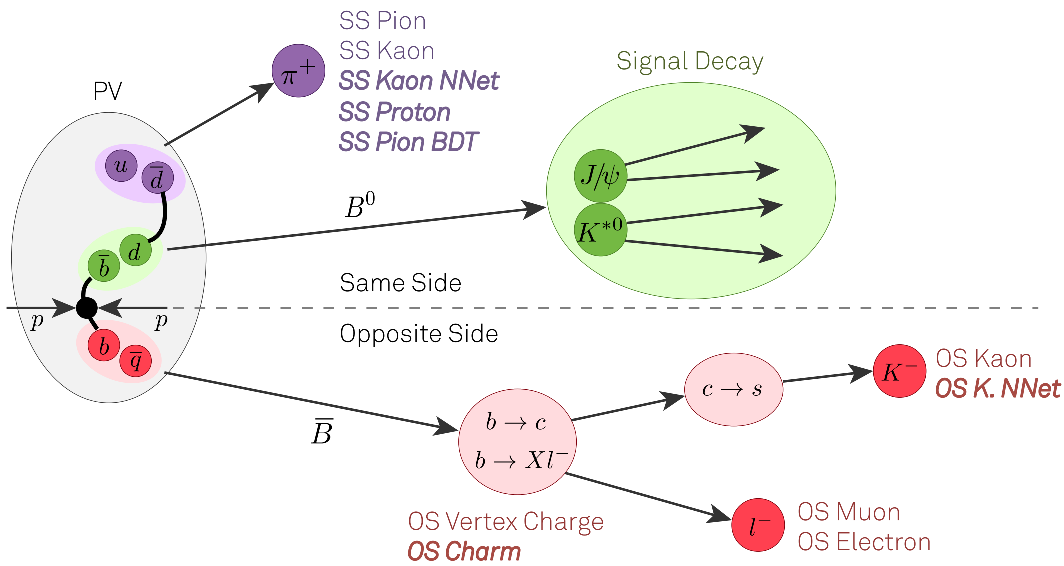


Figure 5.1: Overview of the existing flavour tagging algortihms at LHCb [131]. Above, same-side (SS) taggers look for the signatures of the signal decay. Below, opposite-side (OS) taggers look for the signatures of the decay products of the other b-hadron of the event.

5.2.1 Tagging information

Each tagger provides a decision about the flavour of the b-hadron, q, and a probability of the decision to be wrong, namely the mistag probability ω. The flavour decision can take the values 1 (for B), -1 (for $\bar{B}$), or 0 (no decision, also named *untagged*). Based on these decisions, the tagging efficiency of a sample is the number of tagged events over the total:

$$\epsilon_{\text{tag}} = \frac{N_{\text{tagged}}}{N_{\text{tagged}} + N_{\text{untagged}}}. \tag{5.1}$$

The mistag probability ω is a continuous variable that can range from 0 to 0.5. It appears due to the misidentification of the reconstructed particles, or to the particle-antiparticle oscillations of the opposite b-hadron. For a given sample, the mistag fraction can be calculated as the number of incorrect over tagged events:

$$\omega = \frac{N_{\text{incorrect}}}{N_{\text{tagged}}}. \tag{5.2}$$

The real performance of a tagger can be calculated with a single quantity, the tagging power:

$$\epsilon_{\text{power}} = \epsilon_{\text{tag}}\,\mathcal{D}^2 = \epsilon_{\text{tag}}\,(1-2\omega)^2\,, \tag{5.3}$$

representing the effective statistical reduction of a sample due to the tagging efficiency and the mistag probability. The quantity $\mathcal{D} = 1 - 2\omega$ is called the dilution factor. A tagger performs better with a high tagging efficiency and a low mistag probability.

Opposite-side (OS) taggers

Opposite-side taggers determine the flavour of the accompanying b-hadron by looking at their decay products. Signatures of a given flavour are the charge of the lepton produced in semileptonic decays (electron OSe, or muon OSμ), the charge of the kaon produced in the $b \to c \to s$ decay chain (OSK), or the charge of the detached vertices (OSVtx). The taggers OSe, OSμ, OSK and OSVtx are based on a neural network (NN) trained on simulated $B^+ \to J/\Psi K^+$ decays. The Flavour Tagging (FT) group of the LHCb combines all the mentioned taggers into a single decision: OSComb [127].

Same-side (SS) taggers

In the fragmentation process, the signal B meson is produced together with a charged hadron: a kaon (pion) for B_s^0 (B^0 or B^+), whose charge depends on the particle-antiparticle nature of the meson at the hadronization. Same-side taggers try to identify the charge of this additional track, exploiting the correlation of the particles produced in the hadronization.

In this analysis, $B_s^0 \to \phi\gamma$ decays are tagged with the SSKaonNNet algorithm [129], a neural network which searches for the kaon produced in the fragmentation process. The multivariate algorithm is trained on simulated

$B_s^0 \to D_s^- \pi^+$ decays in two steps: the first step (NN1) separates between fragmentation and detached kaons, while the second step (NN2) calculates the decision and the mistag probability using the output of NN1.

5.2.2 Calibration of the taggers

The predicted mistag probability, η, is the best estimation of a tagging algorithm of the real mistag probability, ω, for a given event. The estimated η has to be calibrated to match the real ω, considering a linear expression that depends on two parameters, p_0 and p_1:

$$\omega(\eta) = \langle \eta \rangle + p_0 + p_1(\eta - \langle \eta \rangle), \tag{5.4}$$

where $\langle \eta \rangle$ is the average predicted mistag probability of the sample. A perfectly calibrated tagger would have $p_0 = 0$ and $p_1 = 1$.

The purpose of this section is to present the calibration parameters for $B_s^0 \to \phi\gamma$ decays, in both data and simulation samples. Although only the Run 1 dataset is used in this analysis, the calibration parameters for Run 2 data are also calculated, corresponding to the 2015 and 2016 data-taking period. Table 5.1 lists the calibration parameters of the different taggers for Run 1 and Run 2, where the first uncertainty is statistical and the second is the portability systematic. A portability systematic is assigned to account for possible kinematical differences between the sample used to calibrate the tagger and the sample where it is actually applied. As it can be observed, the calibration parameters for Run 1 and Run 2 are compatible within their uncertainties. Figures 5.2, 5.3 and 5.4 show the mistag calibration curves using $B^+ \to J/\Psi K^+$ (for OS), $B_s^0 \to D_s^- \pi^+$ (for SS), and $B^0 \to K^{*0}\gamma$ (for OS), respectively.

Table 5.1: Calibration parameters of the taggers for the $B_s^0 \to \phi\gamma$ decay. For Run 1, the $\langle\eta\rangle$ parameter is extracted from the sample in which the calibration is applied to. The $B_s^0 \to D_s^-\pi^+$ sample is reweighted to match the kinematical distributions of the radiative channels.

Tagger	Calibration channel	$\langle\eta\rangle$	p_0	p_1	$\rho(p_0, p_1)$
	Run 1				
OSComb	FT group [132]	sample-dependent	$0.0062 \pm 0.0019 \pm 0.0040$	$0.982 \pm 0.007 \pm 0.034$	0.14
SSKaonNNet	FT group [132]	sample-dependent	$0.005 \pm 0.004 \pm 0.003$	$0.976 \pm 0.071 \pm 0.057$	0.02
	Run 2				
OSComb	$B^+ \to J/\Psi K^+$	0.3827	$0.0054 \pm 0.0011 \pm 0.0010$	$0.887 \pm 0.012 \pm 0.044$	0.12
OSComb	$B^0 \to K^{*0}\gamma$	0.3758	$0.011 \pm 0.011 \pm 0.005$	$0.813 \pm 0.121 \pm 0.012$	0.02
SSKaonNNet	$B_s^0 \to D_s^-\pi^+$	0.4230	$0.008 \pm 0.006 \pm 0.003$	$1.062 \pm 0.085 \pm 0.025$	0.08
SSKaonNNet	$B_s^0 \to D_s^-\pi^+$ reweighted	0.4222	$0.009 \pm 0.007 \pm 0.002$	$1.127 \pm 0.092 \pm 0.008$	0.08

Calibration for Run 1 data

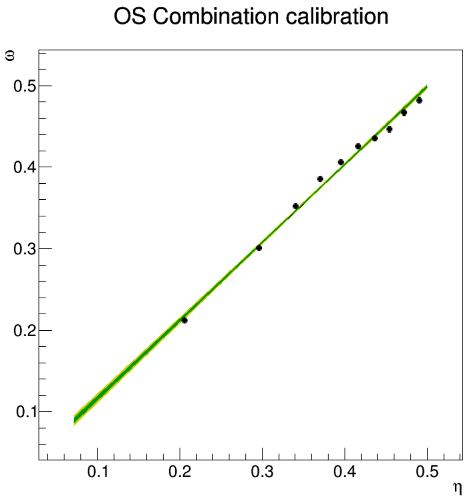
For Run 1 data, the calibration parameters for the OSComb and SSKaonNNet taggers are provided by the Flavour Tagging (FT) group of the LHCb [132]. The parameters are shown in Table 5.1.

- The calibration of the OS taggers is performed with the self-tagged $B^+ \to J/\Psi K^+$ and $B^+ \to D^0\pi^+$ decays, the neutral $B^0 \to J/\psi K^*$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ decays, and $B_s^0 \to D_s^-\pi^+$ and $B_s^{**} \to B^+K^-$ which involve B_s meson decays.
- The calibration of the SSKaonNNet is performed on self-tagged $B_s^0 \to D_s^-\pi^-$ decays, together with flavour-specific $B_{s2}^*(5840)^0 \to B^+K^-$ decays.

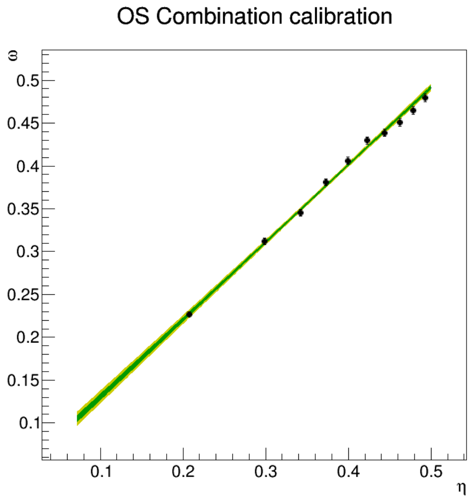
Calibration for Run 2 data

For Run 2 data, the calibration has to be extracted from other data samples similar to the signal, since the FT group does not provide the values anymore. In this analysis, the calibration is done using the EspressoPerformanceMonitor (EPM) tool [133] of LHCb, which was created for flavour tagging studies. The results are shown in Table 5.1.

- The calibration parameters for the OS tagger are extracted from a sample of 1.2 million $B^+ \to J/\Psi K^+$ decays. The calibration plot is shown in Figure 5.2. A TOS-trigger event selection are applied in $B^+ \to J/\Psi K^+$ candidates.
- The SS tagger is calibrated with a $B_s^0 \to D_s^-\pi^+$ sample, as it is shown in Figure 5.3. A reweighting is applied in order to match the kinematical distributions of the radiative channels.

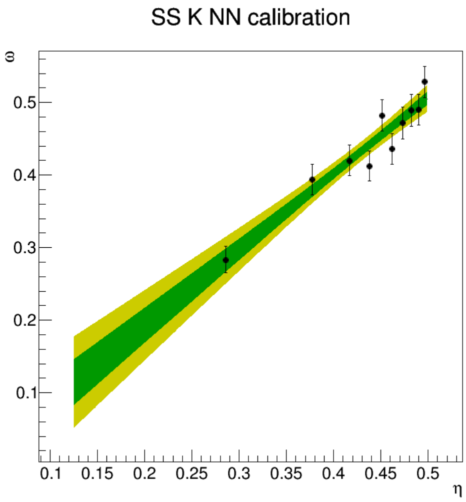


(a) Run 1

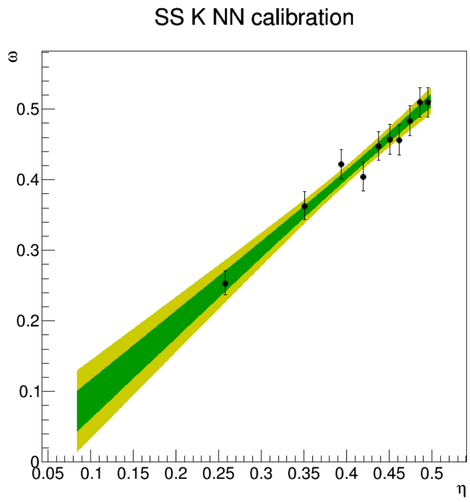


(b) Run 2

Figure 5.2: Calibration of the OSComb tagger using $B^+ \to J/\Psi K^+$ decays for Run 1 (left) and Run 2 (right). The green and yellow bands represent the 1σ and 2σ contours, respectively.



(a) Run 1



(b) Run 2

Figure 5.3: Calibration of the SSKaonNNet tagger using $B_s^0 \to D_s^- \pi^+$ decays for Run 1 (left) and Run 2 (right). The green and yellow bands represent the 1σ and 2σ contours, respectively.

Verification of the OS tagger calibration with $B^0 \to K^{*0}\gamma$

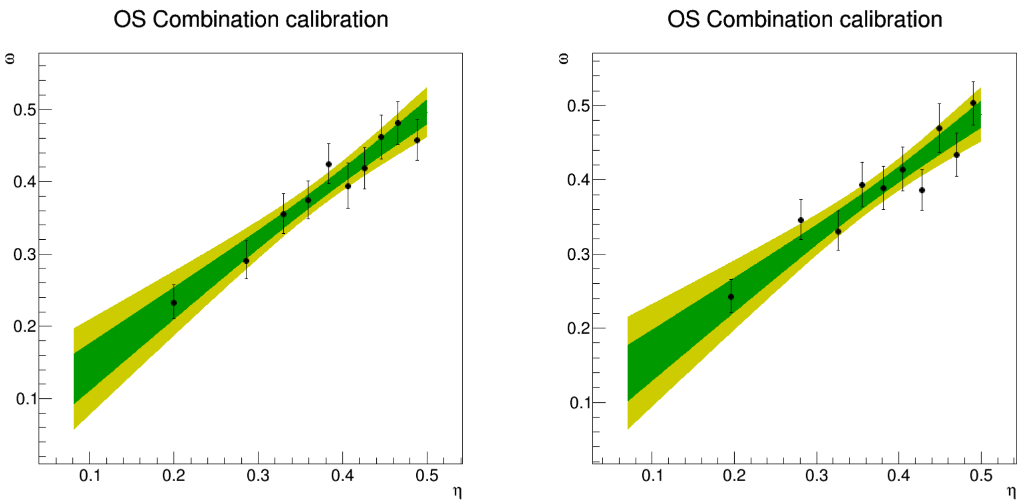
Additionally, the calibration of the OS tagger has been extracted using $B^0 \to K^{*0}\gamma$ decays, as shown in Figure 5.4 and Table 5.1. This is a decay with similar kinematical and topological properties as the signal channel $B_s^0 \to \phi\gamma$, with also a photon in the final state. The calibration parameters obtained with $B^0 \to K^{*0}\gamma$ and $B^+ \to J/\Psi K^+$ decays are compatible within statistical uncertainties. However, the available number of $B^0 \to K^{*0}\gamma$ candidates (34k) is much lower than in the $B^+ \to J/\Psi K^+$ channel (1.2 million candidates), obtaining parameter uncertainties one order of magnitude smaller when using the latter.

Mixing asymmetry The time-dependent mixing asymmetry of flavour-specific $B^0 \to K^{*0}\gamma$ decays can serve as a validation of the flavour tagging calibration. It is defined as:

$$A_{\text{mix}}(t) = \frac{N_{\text{unmixed}}(t) - N_{\text{mixed}}(t)}{N_{\text{unmixed}}(t) + N_{\text{mixed}}(t)} = (1 - 2\omega)\cos(\Delta m_d\, t), \qquad (5.5)$$

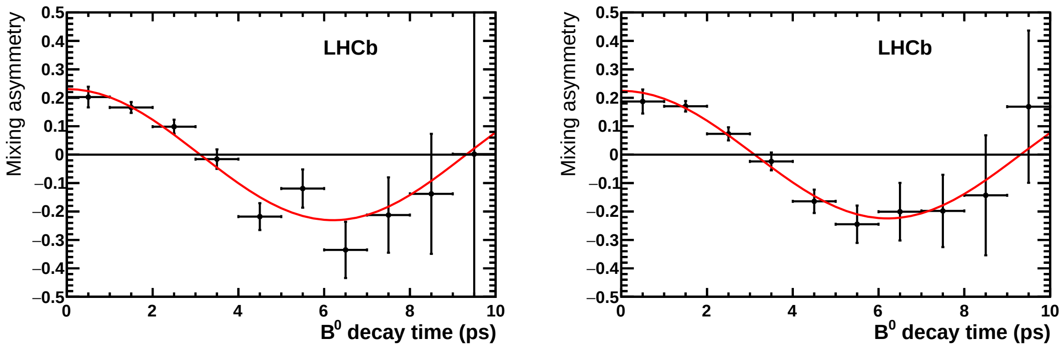
where N_{mixed} (N_{unmixed}) is the number of tagged events which have (not) oscillated at a decay time t, and $\mathcal{D} = (1-2\omega)$ is the dilution of the asymmetry amplitude induced by the mistag rate.

Figure 5.5 shows the mixing asymmetry in $B^0 \to K^{*0}\gamma$ decays using the OSComb tagger in data, in the first application of a tagging algorithm on a radiative channel at LHCb. As it can be observed, the data points are compatible with the curve constructed from Equation 5.5.



(a) Run 1 (b) Run 2

Figure 5.4: Calibration of the OSComb tagger using $B^0 \to K^{*0}\gamma$ decays for Run 1 (left) and Run 2 (right). The green and yellow bands represent the $1\,\sigma$ and $2\,\sigma$ contours, respectively.



(a) Run 1 (b) Run 2

Figure 5.5: Time-dependent mixing asymmetry in $B^0 \to K^{*0}\gamma$ decays [131], using the calibrated OSComb tagger for Run 1 (left) and Run 2 (right).

5.2.3 Combination of the taggers

The SSKaonNNet and OSComb taggers are combined [134] following a similar strategy as in the measurement of the CP-violating phase ϕ_s using $B_s^0 \to J/\Psi h^- h^+$ [135] and $B_s^0 \to D_s^+ D_s^-$ [136] decays. The taggers are combined based on the considered *best* tagger for each event: given the mistag probabilities $\omega_{\rm OS}$ and $\omega_{\rm SS}$, the best tagger would be the one with the lowest mistag probability, $\omega_{\rm best} = \min(\omega_{\rm OS}, \omega_{\rm SS})$, and the *other* tagger corresponds to $\omega_{\rm other} = \max(\omega_{\rm OS}, \omega_{\rm SS})$. The combination is based on the following logic:

- If one of the taggers provides a decision ($q \neq 0$), but the other does not ($q = 0$), then keep the tagger with the non-zero decision:

$$(q_{\rm best} \neq 0, q_{\rm other} = 0) \Rightarrow \begin{cases} q_{\rm comb} = q_{\rm best}, \\ \omega_{\rm comb} = \omega_{\rm best}. \end{cases} \tag{5.6}$$

- If both taggers provide a decision, and they *agree*, then combine the mistag rates as:

$$(q_{\rm best} = q_{\rm other}) \Rightarrow \begin{cases} q_{\rm comb} = q_{\rm best} = q_{\rm other}, \\ \omega_{\rm comb} = \dfrac{\omega_{\rm best}\,\omega_{\rm other}}{\omega_{\rm best}\,\omega_{\rm other} + (1-\omega_{\rm best})(1-\omega_{\rm other})}, \end{cases} \tag{5.7}$$

  where $p(q_{\rm comb}) = \omega_{\rm best}\,\omega_{\rm other}$ and $p(-q_{\rm comb}) = (1-\omega_{\rm best})(1-\omega_{\rm other})$ are the probabilities of being $q_{\rm comb}$ and $-q_{\rm comb}$, respectively, so that $\omega_{\rm comb}$ is correctly normalized and can be interpreted as a probability.

- If both taggers provide a decision, but they do *not agree*, then combine the mistag rates as:

$$(q_{\rm best} = -q_{\rm other}) \Rightarrow \begin{cases} q_{\rm comb} = q_{\rm best} = -q_{\rm other}, \\ \omega_{\rm comb} = \dfrac{\omega_{\rm best}\,(1-\omega_{\rm other})}{\omega_{\rm best}\,(1-\omega_{\rm other}) + (1-\omega_{\rm best})\omega_{\rm other}}, \end{cases} \tag{5.8}$$

  where $p(-q_{\rm other}) = (1-\omega_{\rm other})$ is the estimated mistag probability of the *other* tagger to be $-q_{\rm other}$.

The mistag probability distribution of the combination is shown in Figure 5.6 for the $B_s^0 \to \phi\gamma$ data sample, along with the individual distributions of the calibrated OS and SS taggers.

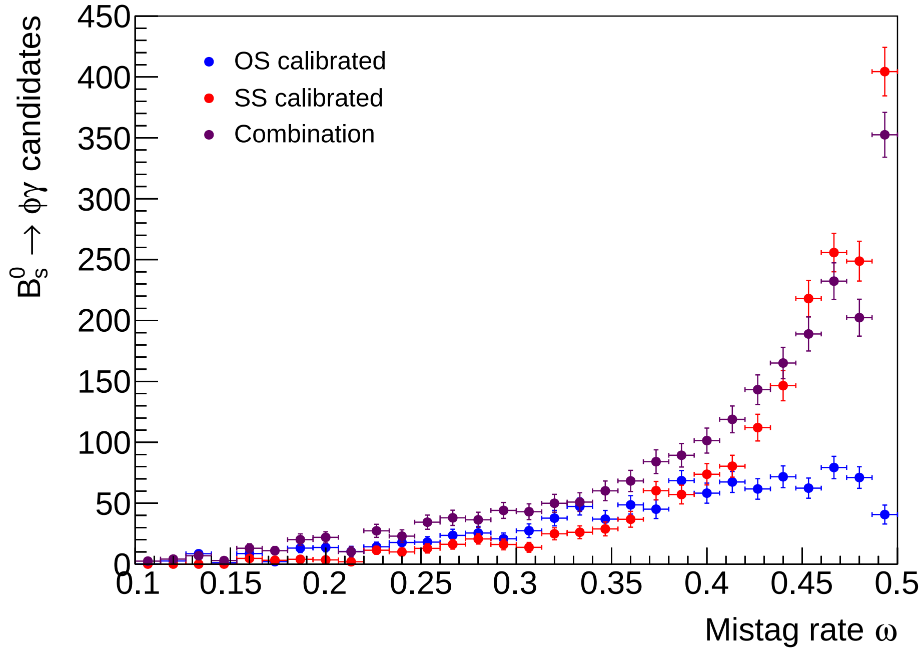


Figure 5.6: Mistag probability distribution of the OS and SS taggers, and its combination, for tagged events $q_{\text{tagger}} \neq 0$ in $B_s^0 \to \phi\gamma$ 2011+2012 data.

5.2.4 Performances of the taggers

The tagging power ϵ_{power} determines the performance of a tagger. It depends on the tagging efficiency ϵ_{tag} and the squared dilution factor $\mathcal{D}^2 = (1-2\omega)^2$ (Equation 5.3). In a weighted sample, the tagging efficiency is calculated as:

$$\epsilon_{\text{tag}} = \frac{\sum_i^{\text{tagged}} W_i}{\sum_i^{\text{total}} W_i}, \tag{5.9}$$

where W_i is the per-event sWeight. On the other hand, the average squared dilution corresponds to:

$$\langle \mathcal{D}^2 \rangle = \frac{\sum_i^{\text{tagged}} (1-2\omega_i)^2 \, W_i}{\sum_i^{\text{tagged}} W_i}, \tag{5.10}$$

which depends on the per-event mistag probability ω_i. The tagging power is then calculated from these two quantities.

Table 5.2 shows the performance of the OS and SS taggers (with and without applying the calibration), and the combination of the two taggers. The effective mistag rate ω_{eff} is calculated from the average squared dilution:

$$\omega_{\text{eff}} = \frac{1-\sqrt{\langle \mathcal{D}^2 \rangle}}{2}. \tag{5.11}$$

Table 5.2: Performances of the $B_s^0 \to \phi\gamma$ taggers in 2011+2012 data, calculated with the effective per-event mistag probability $\omega_{\rm eff}$.

$B_s^0 \to \phi\gamma$ tagger	$\epsilon_{\rm tag}$ (%)	$\omega_{\rm eff}$ (%)	$\epsilon_{\rm power}^{\omega_{\rm eff}}$ (%)
OSComb	33.03	34.80	3.05
SSKaonNNet	63.89	40.82	2.16
OSComb calibrated	33.03	35.40	2.82
SSKaonNNet calibrated	63.89	41.27	1.95
Combination	74.53	37.61	4.58

The tagging power of the calibrated OS and SS taggers are calculated to be 2.82% and 1.95%, respectively, while the combination of the taggers raises this efficiency to 4.58%. It can be compared to the tagging power in $B_s^0 \to J/\psi K^+K^-$ decays, $(3.73 \pm 0.15)\%$ [135], and in $B_s^0 \to D_s^+ D_s^-$ decays, $(5.33 \pm 0.25)\%$ [136]. The quoted performance corresponds to the expected $\epsilon_{\rm tag}$ when using the per-event mistag probabilities, as it is the case in this analysis. Another approach would be to use an average mistag rate, calculated from the sample as:

$$\langle\omega\rangle = \frac{\sum_i^{\rm tagged} \omega_i W_i}{\sum_i^{\rm tagged} W_i}. \tag{5.12}$$

This approach would degrade the performance down to 2.40%. Although this approach is not used in the analysis, it can serve as a cross-check, and has helped to understand the effect of the tagging performance.

5.3 Decay time resolution

The decay time resolution is one of the key elements required to measure the observables $\mathcal{S}$ and $\mathcal{C}$. In Section 5.3.1, the resolution is studied in radiative simulation samples. A dependency with the decay time t is observed, which can be taken into account using the per-event decay time uncertainty σ_t multiplied by a constant scaling factor. In Section 5.3.2, several resolution functions are fitted against the simulation samples. When using a per-event σ_t, a minimum of two Gaussians (with two scaling factors) are required to correctly reproduce the simulated data. The resolution parameters extracted in this section are used in the final fit to the data. Section 5.3.3 explains how the resolution can be controlled using real data, through fake $B^0 \to K^{*0}J/\psi$ candidates and prompt-ϕ candidates.

5.3.1 Resolution as a function of the decay time

The decay time resolution of the radiative $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays is shown in Figure 5.7 as a function of the reconstructed decay time, obtained from the simulated samples. For each bin in t, the distribution of the decay time resolution is fitted to a single Gaussian to obtain a value for the mean and width.

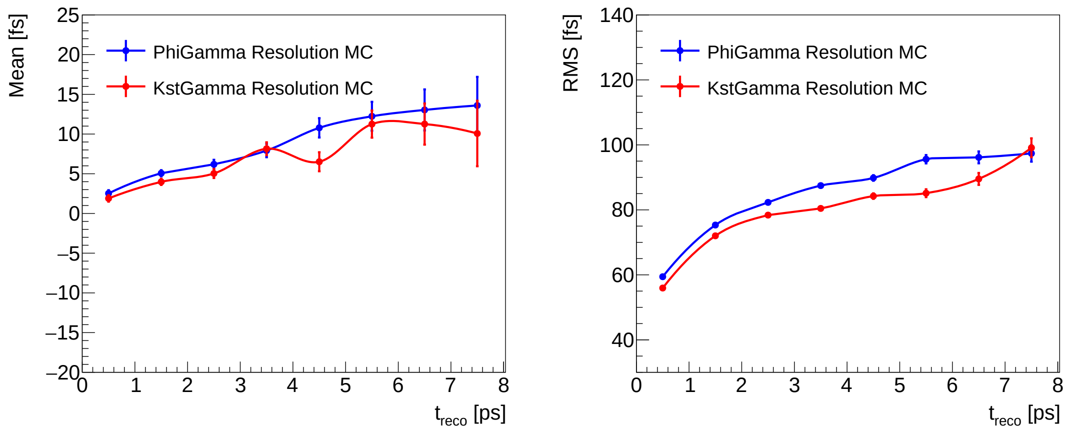
Both the mean and width of the resolution depend on t; the mean ranges linearly from 0 to 10 fs, while the width take values from 50 to 100 fs, with a non-linear dependency at low decay times. The resolution is slighly larger in $B_s^0 \to \phi\gamma$ than in $B^0 \to K^{*0}\gamma$, with a difference of 1 fs in mean and 5 fs in width, but the dependency with t follows the same pattern.

Scaling the decay time uncertainty

A way to take into account the t dependency is to use the per-event decay time uncertainty σ_t, which can be expressed in terms of the uncertainties of the momentum and the decay length:

$$\sigma_t = \sqrt{\left(\frac{t}{p}\right)^2 \sigma_p^2 + \left(\frac{m}{p}\right)^2 \sigma_L^2}, \tag{5.13}$$

where the first term depends directly on the decay time t. However, σ_t is usually underestimated due to reconstruction effects and does not represent



(a) Mean

(b) RMS

Figure 5.7: Decay time resolution (mean and width) of $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays, in slices of the reconstructed decay time, obtained from Run 1 simulated events.

the actual resolution of the event, as can be appreciated in Figure 5.8 from simulated events. Therefore, a scaling factor is needed to correct for it:

$$\sigma_{\text{res.}} = s\,\sigma_t, \tag{5.14}$$

which is extracted from the radiative simulation samples.

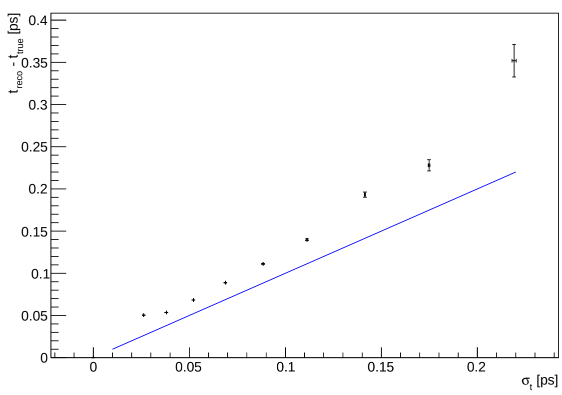


Figure 5.8: MC resolution in slices of the decay-time uncertainty. The blue line corresponds to the $x = y$ case.

5.3.2 Resolution description

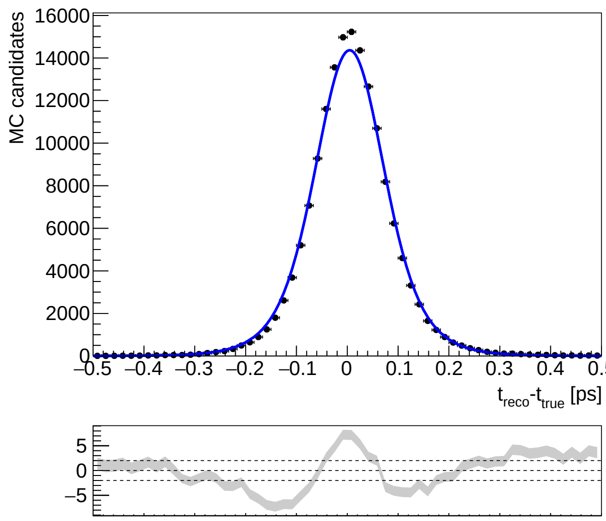
The resolution parameters are obtained from an unbinned maximum likelihood fit to the per-event resolution $\Delta t = t_{\text{reco}} - t_{\text{true}}$ of simulated $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays. Several descriptions of the decay time resolution are explored, with single or double Gaussian functions, and are detailed in the following.

Single Gaussian

The per-event decay time resolution can be described by a single Gaussian function:

$$R_{1\text{G}} \equiv G(\mu, s\sigma_t) = \frac{1}{\sqrt{2\pi}\, s\sigma_t} e^{-(\Delta t-\mu)^2/2(s\sigma_t)^2}, \tag{5.15}$$

where $\Delta t = t_{\text{reco}} - t_{\text{true}}$ is the true resolution, μ is the mean or the resolution bias, and $s\sigma_t$ is the scaled per-event uncertainty. There are two free parameters in the fit: μ and s, which are shown in Figure 5.9 from simulated $B_s^0 \to \phi\gamma$ decays. The $\chi^2 = 10.6$ goodness-of-fit indicates that at least two Gaussians are needed to correctly reproduce the decay time resolution.



$R_{1\text{G}}$	$B_s^0 \to \phi\gamma$
μ (fs)	4.8 ± 0.2
s	1.222 ± 0.002
χ^2	10.6

Figure 5.9: Fit to the per-event decay time resolution in simulated $B_s^0 \to \phi\gamma$ decays, where the resolution is described as a single Gaussian $R_{1\text{G}}$ (Equation 5.15).

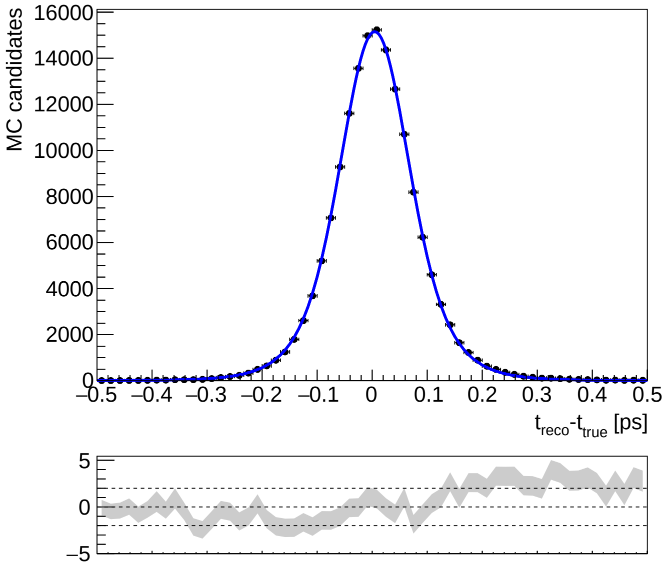
Double Gaussian

This section analyzes the decay time resolution as a sum of two Gaussian functions:

$$\begin{aligned} R_{2\mathrm{G}} \equiv & f G_1(\mu, s_1\sigma_t) + \\ & (1-f) G_2(\mu, s_2\sigma_t), \end{aligned} \tag{5.16}$$

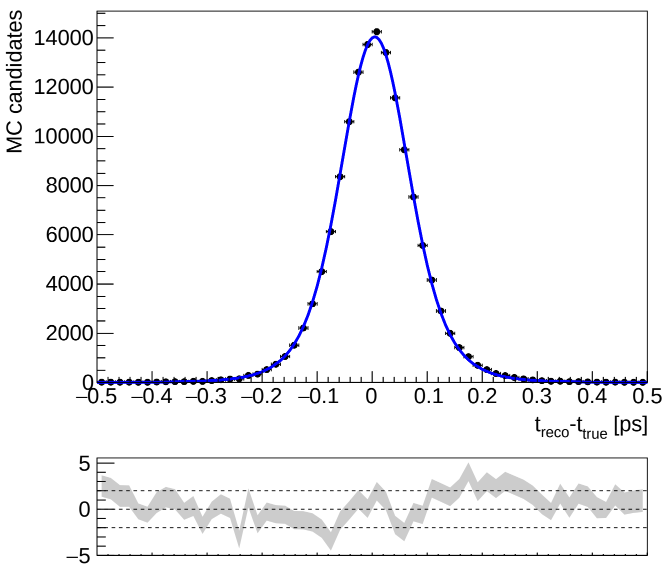
where f is the fraction of the first Gaussian, and s_1 and s_2 are the scaling factors of the first and second Gaussian. The fits to the simulated $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays are presented in Figure 5.10 together with the fitted parameters: f, μ, s_1 and s_2. A significant reduction of the χ^2 test statistic (3.1 for $B_s^0 \to \phi\gamma$, 2.2 for $B^0 \to K^{*0}\gamma$) is observed compared to the single Gaussian case. The resolution parameters of the two radiative decays are compatible within the statistical precision.

The fit to the per-event resolution is repeated in slices of the reconstructed decay time, in order to study the dependency of the resolution parameters with t. The individual fits are shown in Figure 5.11. The fitted parameters, provided in Table 5.3 for simulated $B_s^0 \to \phi\gamma$ decays, show a dependency for the mean $\mu(t)$ whose values range from -0.4 ± 1.7 fs at low decay times to 9.8 ± 0.9 fs at high decay times. The rest of the parameters, however, remain compatible within the statistical precision. The χ^2 of the individual slices stay below 1.3.



R_{2G}	$B_s^0 \to \phi\gamma$
f	0.978 ± 0.002
μ (fs)	4.5 ± 0.2
s_1	1.144 ± 0.003
s_2	2.79 ± 0.08
χ^2	3.1

(a) $B_s^0 \to \phi\gamma$ simulation



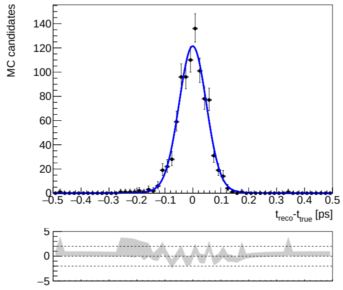
R_{2G}	$B^0 \to K^{*0}\gamma$
f	0.974 ± 0.003
μ (fs)	4.8 ± 0.2
s_1	1.141 ± 0.004
s_2	2.65 ± 0.07
χ^2	2.2

(b) $B^0 \to K^{*0}\gamma$ simulation

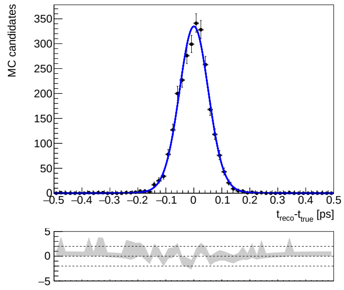
Figure 5.10: Fit to the per-event decay time resolution in simulated (a) $B_s^0 \to \phi\gamma$ and (b) $B^0 \to K^{*0}\gamma$ decays. The resolution is defined as a double Gaussian R_{2G}.

Table 5.3: Fit to the per-event decay time resolution in slices of t, in simulated $B_s^0 \to \phi\gamma$ decays. The resolution is defined as a double Gaussian R_{2G}.

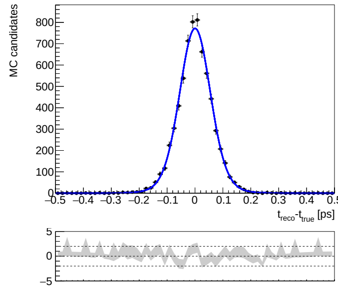
Slice in t (ps)	f	μ	s_1	s_2	χ^2
$[0.3, 0.5]$	0.980 ± 0.015	-0.4 ± 1.7	1.10 ± 0.03	2.91 ± 0.69	1.0
$[0.5, 0.6]$	0.988 ± 0.007	0.3 ± 1.0	1.11 ± 0.02	3.29 ± 0.63	0.8
$[0.6, 0.75]$	0.988 ± 0.008	2.3 ± 0.7	1.11 ± 0.02	2.79 ± 0.35	0.8
$[0.75, 1.0]$	0.978 ± 0.007	2.8 ± 0.5	1.11 ± 0.01	2.55 ± 0.22	0.7
$[1.0, 1.5]$	0.987 ± 0.003	3.8 ± 0.4	1.13 ± 0.01	2.97 ± 0.24	1.1
$[1.5, 2.1]$	0.970 ± 0.007	4.4 ± 0.4	1.12 ± 0.01	2.51 ± 0.15	1.1
$[2.1, 3.1]$	0.970 ± 0.006	5.2 ± 0.4	1.14 ± 0.01	2.61 ± 0.14	1.3
$[3.1, 5.0]$	0.970 ± 0.003	5.2 ± 0.5	1.18 ± 0.01	3.38 ± 0.27	1.3
$[5.0, 7.0]$	0.980 ± 0.008	9.8 ± 0.9	1.25 ± 0.02	3.45 ± 0.49	0.8
$[7.0, 10.0]$	0.893 ± 0.051	9.5 ± 1.7	1.17 ± 0.04	2.27 ± 0.27	0.8
$[0.3, 10.0]$	0.978 ± 0.002	4.5 ± 0.2	1.144 ± 0.003	2.79 ± 0.08	3.1



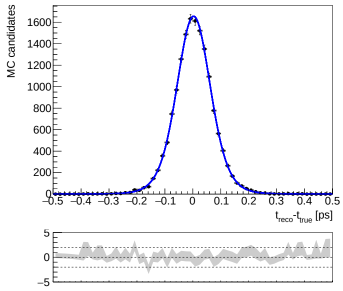
(a) $[0.4, 0.5]$ ps



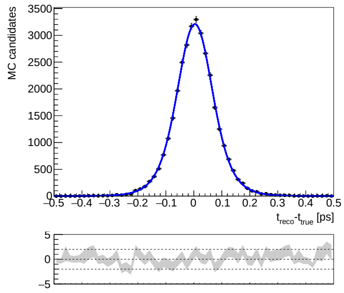
(b) $[0.5, 0.6]$ ps



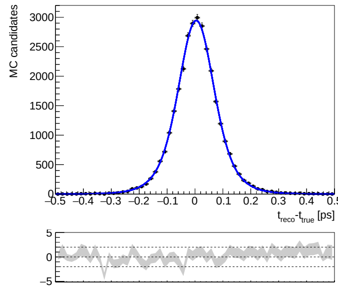
(c) $[0.6, 0.75]$ ps



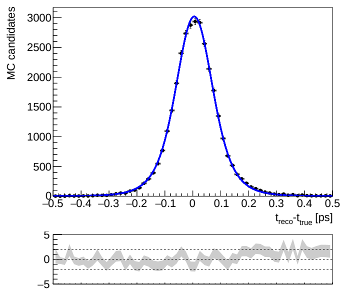
(d) $[0.75, 1.0]$ ps



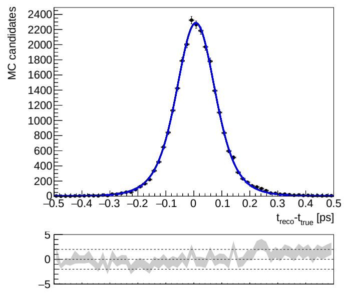
(e) $[1.0, 1.5]$ ps



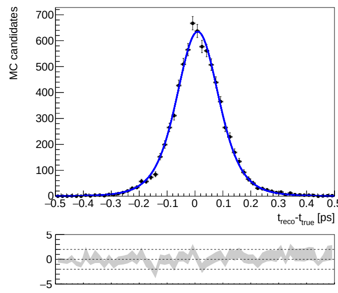
(f) $[1.5, 2.1]$ ps



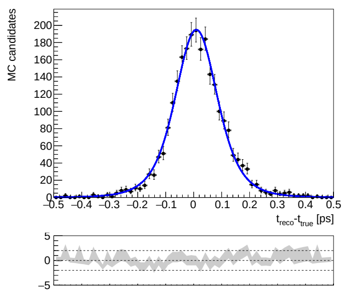
(g) $[2.1, 3.1]$ ps



(h) $[3.1, 5.0]$ ps



(i) $[5.0, 7.0]$ ps



(j) $[7.0, 10.0]$ ps

Figure 5.11: The decay time resolution in slices of the reconstructed decay time, in simulated $B_s^0 \to \phi\gamma$ decays. The resolution is defined as a double Gaussian $R_{2\mathrm{G}}$.

Double Gaussian, scaling the resolution mean

The dependency of the resolution bias with the reconstructed decay time $\mu(t)$, observed in Table 5.3, can be taken into account by scaling the mean with the decay time uncertainty σ_t, in the same way as it is done for the width. The suitability of this approach is tested in simulated $B_s^0 \to \phi\gamma$ decays. Figure 5.12 (a) shows the fit of the per-event resolution using a linear dependency with the uncertainty:

$$\mu(\sigma_t) = s_\mu \sigma_t, \tag{5.17}$$

while Figure 5.12 (b) does the same with a quadratic dependency:

$$\mu(\sigma_t^2) = s_\mu \sigma_t^2. \tag{5.18}$$

The χ^2 of the fits is reduced to 2.0, for a linear dependency, and to 1.5, for a quadratic dependency. The improvement in the residual distribution is subtle but can be appreciated.

However, these configurations double the computational time required to perform the final fit for the measurement. In addition, it is observed that the effect of the μ scaling on the observables $\mathcal{S}$ and $\mathcal{C}$ is very small, of the order of 0.01, so this contribution is added into the systematic uncertainty and the final fit is implemented using a double Gaussian without μ scaling, $R_{2\mathrm{G}}$.

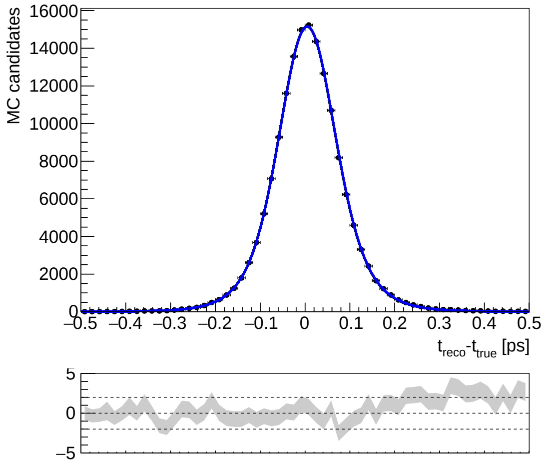
Crystal Ball, without per-event σ_t

The resolution can be described without using the per-event decay time uncertainty σ_t, involving a double-tailed Crystal Ball distribution [137–139] with a mean and width that depend on t:

$$\mathrm{CB}(t) = \begin{cases} A_L\left(B_L - \frac{t-\mu}{\sigma}\right)^{-n_L}, & \text{for } \frac{t-\mu}{\sigma} \leq -\alpha_L \\ \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right), & \text{for } -\alpha_L < \frac{t-\mu}{\sigma} < \alpha_R \\ A_R\left(B_R + \frac{t-\mu}{\sigma}\right)^{-n_R}, & \text{for } \frac{t-\mu}{\sigma} \geq \alpha_R \end{cases} \tag{5.19}$$

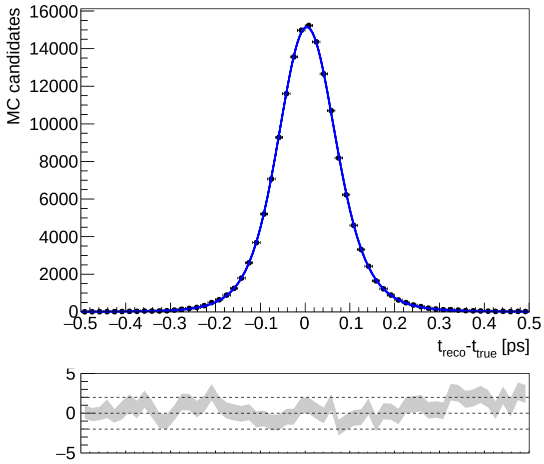
where $\alpha_{L,R} > 0$, $A_{L,R}$ and $B_{L,R}$ are:

$$\begin{aligned} A_i &= \left(\frac{n_i}{\alpha_i}\right)^{n_i} \exp\left(-\frac{\alpha_i^2}{2}\right), \\ B_i &= \frac{n_i - 1}{\alpha_i}, \end{aligned} \tag{5.20}$$



$R_{2\text{G}}, \mu(\sigma_t)$	$B_s^0 \to \phi\gamma$
f	0.978 ± 0.002
s_μ	0.089 ± 0.003
s_1	1.143 ± 0.003
s_2	2.79 ± 0.08
χ^2	2.0

(a) Double Gaussian, with $\mu(\sigma_t) = s_\mu \sigma_t$



$R_{2\text{G}}, \mu(\sigma_t^2)$	$B_s^0 \to \phi\gamma$
f	0.978 ± 0.002
s_μ	1.35 ± 0.04
s_1	1.143 ± 0.003
s_2	2.80 ± 0.08
χ^2	1.5

(b) Double Gaussian, with $\mu(\sigma_t^2) = s_\mu \sigma_t^2$

Figure 5.12: Fit to the per-event decay time resolution in simulated $B_s^0 \to \phi\gamma$ decays, where the resolution is described as a double Gaussian where the mean is scaled by (a) a linear dependency σ_t and (b) a quadratic dependency σ_t^2.

and the parameters $\mu(t)$ and $\sigma(t)$ depend on the decay time:

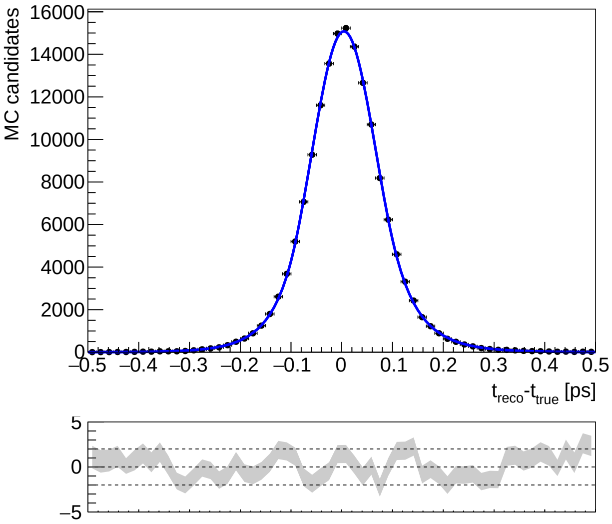
$$\begin{aligned}\mu(t) &= \mu_0 + \mu_s\, t, \\ \sigma(t) &= \sigma_0\, t^{\sigma_p}.\end{aligned} \tag{5.21}$$

Therefore, a total of eigth parameters are needed to characterize the time-dependent resolution: $(\mu_0, \mu_s, \sigma_0, \sigma_p, \alpha_L, \alpha_R, n_L, n_R)$. The $\mu(t)$ and $\sigma(t)$ parameters are the mean and width of the Gaussian distribution defined between $-\alpha_L$ and α_R. It has been checked that the rest of the Crystal Ball parameters do not vary with the decay time. Figure 5.13 shows the resulting parameters of the the Crystal Ball for simulated $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays, in an unbinned maximum likelihood fit to the per-event resolution $\Delta t = t_{\text{reco}} - t_{\text{true}}$. A slightly smaller σ_0 parameter is obtained for $B^0 \to K^{*0}\gamma$ compared to $B_s^0 \to \phi\gamma$, $\Delta\sigma_0 = -1.5 \pm 0.4$, as expected from the average decay time resolution plots (see Figure 5.7). The rest of the parameters are compatible within statistical uncertainties between the two channels.

Due to the complexity of the function, the convolution of this resolution with the decay rate cannot be performed analytically. The RooFit framework has two ways of dealing with numerical convolutions: `RooNumConvPdf` and `RooFFTConvPdf`; the former is a brute force numerical convolution, while the latter makes use of the fast Fourier transform (FFT) theorem to speed up the convolution processing time. However, the FFT technique requires that the resolution parameters do not depend on the convolution variable, a condition that is not fulfilled in our case, as shown in Equation 5.21. On the other hand, the pure numerical convolution presents some instabilities that were observed in the minimization steps, in addition to the unmanageable large processing times. For these reasons, this method has not been considered for the measurement.

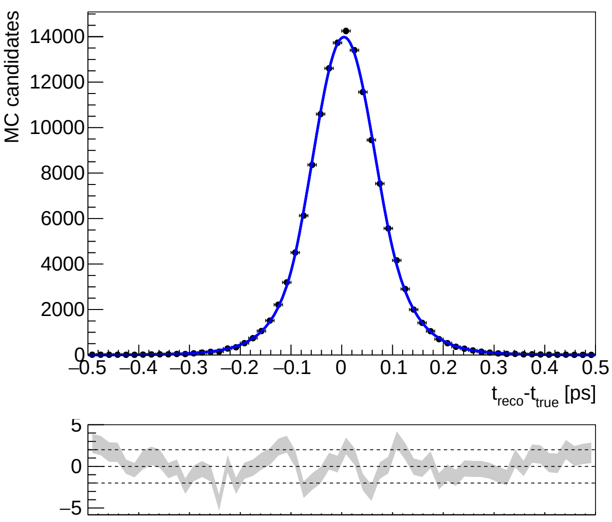
5.3.3 Control of the resolution with real data

To control the potential differences between data and simulation in the description of the decay time resolution, two data-driven studies are performed using (1) $B^0 \to J/\Psi(\to \mu^+\mu^-)K^{*0}$ decays in which the two muons are reconstructed emulating a photon, and (2) ϕ candidates produced in the pp collision vertices.



CB(t)	$B_s^0 \to \phi\gamma$
μ_0 (fs)	0.8 ± 0.4
μ_s	1.81 ± 0.14
σ_0 (fs)	59.5 ± 0.3
σ_p	0.16 ± 0.04
α_L	1.44 ± 0.02
α_R	1.38 ± 0.02
n_L	21.7 ± 3.8
n_R	14.7 ± 1.8
χ^2	1.4

(a) CB(t) in $B_s^0 \to \phi\gamma$ simulation



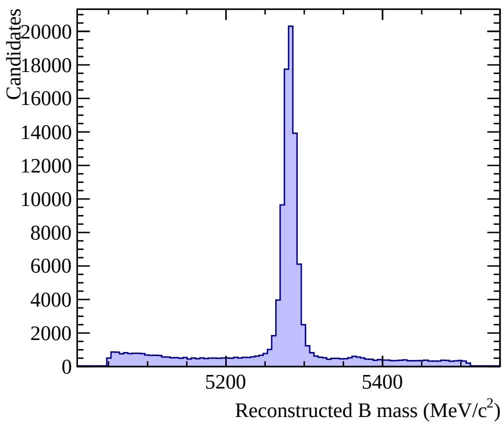
CB(t)	$B^0 \to K^{*0}\gamma$
μ_0 (fs)	0.1 ± 0.4
μ_s	2.21 ± 0.14
σ_0 (fs)	58.0 ± 0.3
σ_p	0.15 ± 0.04
α_L	1.49 ± 0.03
α_R	1.36 ± 0.02
n_L	15.3 ± 2.1
n_R	38.3 ± 11.3
χ^2	1.8

(b) CB(t) in $B^0 \to K^{*0}\gamma$ simulation

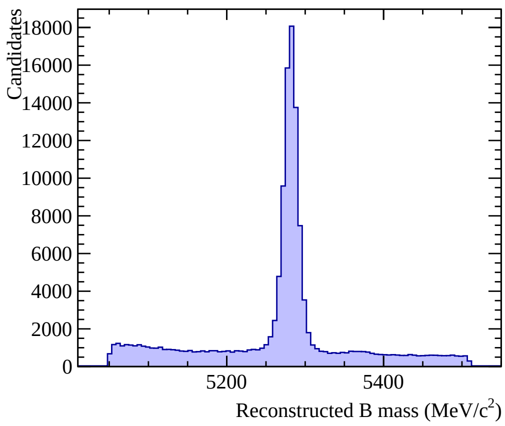
Figure 5.13: Fit to the per-event decay time resolution $\Delta t = t_{\rm reco} - t_{\rm true}$ described as a time-dependent double-tailed Crystal Ball, obtained from the simulation samples of (a) $B_s^0 \to \phi\gamma$ and (b) $B^0 \to K^{*0}\gamma$ decays.

Data-simulation differences in fake $\mathbf{B^0 \to J/\Psi\, K^*}$ decays

The large number of $B^0 \to J/\Psi(\to \mu^+\mu^-)\, K^*(\to K^+\pi^-)$ candidates reconstructed at LHCb, together with a background-clean mass spectrum (Figure 5.14), makes it an excellent control channel to study the decay time resolution. However, its topology is different as compared to radiative decays: there are two muons in the final state $J/\Psi \to \mu^+\mu^-$ instead of a photon, and therefore the vertex is reconstructed from four particles instead of only two. In order to use $B^0 \to J/\Psi\, K^*$ decays to control the decay time resolution in radiative decays, the contribution of the muons to the decay vertex reconstruction is ignored. This behaviour can be implemented in the decay time reconstruction algorithm by setting the vertex covariance matrix of the muons to infinity, so in practice they do not contribute to the vertex fitting. Events reconstructed in this way are called *fake*, since the photon is emulated or faked, while $B^0 \to J/\Psi\, K^*$ events including the muons in the vertex reconstruction are called *full*, since the full covariance matrix contributes to the vertex fitting.



(a) Full-reconstructed

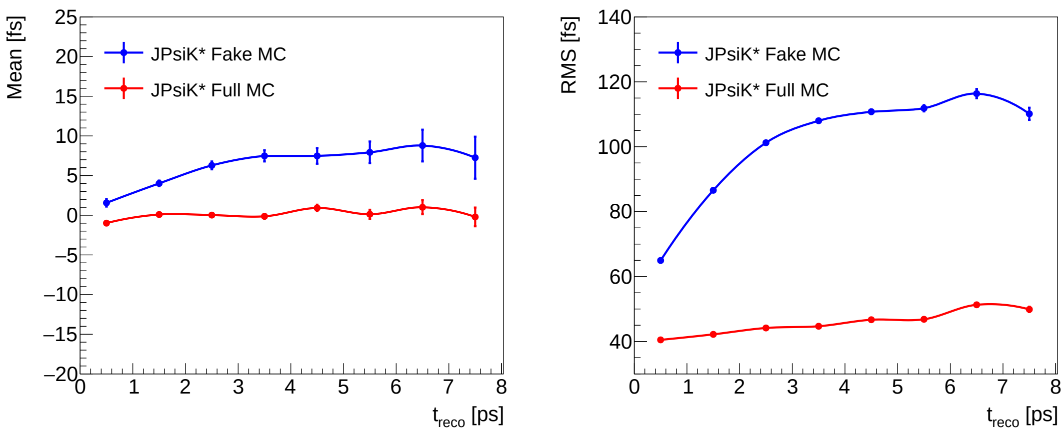


(b) Fake-reconstructed

Figure 5.14: The reconstructed B mass of the *full* and *fake* $B^0 \to J/\Psi K^*$ decays in 2011+2012 data, after the event selection.

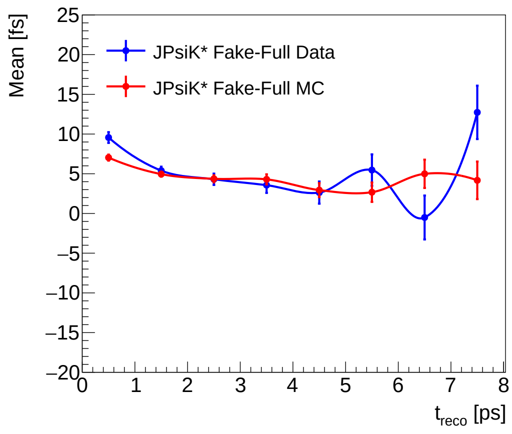
Event selection. For this study, $B^0 \to J/\Psi\, K^*$ candidates are reconstructed with the same stripping as radiative channels, and the offline selection of $B^0 \to K^{*0}\gamma$ (see Section 3.3) is applied except for the photon requirements. The resulting distribution of the reconstructed B mass is shown in Figure 5.14, which is less affected by background contaminations. The final candidates are selected from a mass window $5250 < m_B < 5310$ MeV$/c^2$, since the mass resolution is better than for radiative decays. A total of 79 258 (80 235) full (fake) candidates are selected.

Average resolution in simulation. In simulation samples, the decay time resolution can be obtained by means of the $t_{\text{reco}} - t_{\text{true}}$ distribution, as explained in Section 5.3. Figure 5.15 compares the mean and RMS of the decay time resolution of the fake-reconstructed and the full-reconstructed $B^0 \to J/\Psi K^*$ candidates. As it can be observed from the plots, the decay time resolution is wider in the fake samples (from 60 fs^{-1} to 120 fs^{-1} depending on the decay time) than in the full samples (from 40 fs^{-1} to 50 fs^{-1}). The bias on the resolution, represented by the mean, increases to about 5 fs^{-1} in the fake sample, whereas it remains unbiased in the full sample. Comparing the figures of the fake samples to the average decay time resolution of $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ simulated decays in Figure 5.7, the resolution width is larger and grows more rapidly with the decay time, while the resolution mean is compatible.

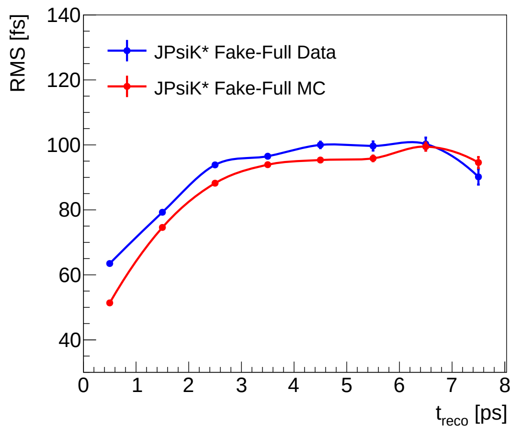


(a) Mean (b) RMS

Figure 5.15: Resolution mean and RMS of the fake- and full-reconstructed $B^0 \to J/\Psi K^*$ decays, from 2011+2012 simulation.



(a) Mean



(b) RMS

Figure 5.16: Mean and RMS of the $t_{\text{fake}} - t_{\text{full}}$ distribution of $B^0 \to J/\Psi K^*$ decays on data and simulation, in the 2011+2012 period.

Fake-Full: comparing data and simulation. The decay time resolution of the data and simulated $B^0 \to J/\Psi K^*$ samples can be compared by using the fake and full samples together. Since the fake and full samples are different reconstructions of the same events, their decay time can be compared event-by-event to construct the variable $t_{\text{fake}} - t_{\text{full}}$, in both data and simulation.

Figure 5.16 compares the mean and RMS of the $t_{\text{fake}} - t_{\text{full}}$ distribution in data and simulation. The mean of such distribution is compatible within statistics, whereas the RMS is about 5 fs^{-1} (10 fs^{-1} for $t < 1$ ps) larger in data than in simulation. The time dependence of the RMS distribution, shown in Figure 5.16 (b), is comparable to the decay time resolution of the radiative channels, previously shown in Figure 5.7.

Data-simulation differences in prompt-ϕ events

An alternative way to control the decay time resolution in data involves fragmentation ϕ mesons (also called prompt-ϕ), produced in the pp collision vertices. These ϕ mesons can be reconstructed together with a random photon in the event to emulate an artificial B_s^0 meson. By reconstructing the decay time of the B_s^0 in data and simulation samples, the possible differences on the decay time resolution can be estimated.

Event selection. In data, prompt $\phi \to K^+K^-$ decays can be selected by using the stripping21 line `StrippingSbarSCorrelationsPhiLine`, which demands two kaons with $\chi^2_{\rm IP} < 49$ coming from the same vertex, along with strong kinematical ($p_K > 5000$ MeV/c) and particle identification requirements. The ϕ meson is combined with a random photon in the event with $p_{\rm T,\gamma} > 3000$ MeV/c to reconstruct the B_s^0 candidate, requiring a minimal compatiblity with a primary vertex $\chi^2_{{\rm IP},B} < 15$. In addition, the kinematical distribution of the random photon is reweighted to match the distribution of radiative channels. To select only real ϕ particles, a background subtraction procedure is applied using the reconstructed ϕ mass as the discriminant variable, as displayed in Figure 5.17.

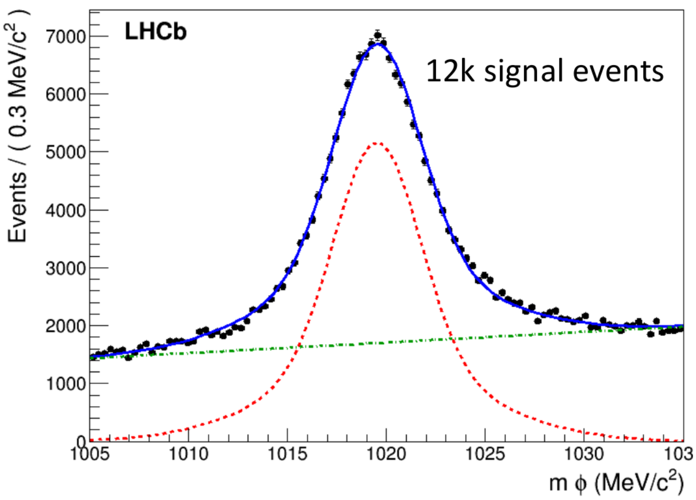


Figure 5.17: Reconstructed mass of prompt-ϕ candidates in Run 1 data. The signal (red) and the combinatorial background (green) are shown.

The same selection is applied to reconstruct prompt-ϕ simulated events, in samples in which the full pp collision is simulated without any additional requirement. These are called *minimum bias* Monte Carlo samples, and are provided by the LHCb collaboration. A total of 2.5k (7.4k) prompt-ϕ candidates are reconstructed in data (minimum bias MC).

Average resolution. Figure 5.18 shows the decay time resolution of $B_s^0 \to \phi\gamma$ decays reconstructed with a prompt-ϕ and a random photon, in data and simulation, compared also to the resolution of signal $B_s^0 \to \phi\gamma$ decays in simulation. Table 5.4 compares the resolution width of the mentioned samples, where all of them are compatible. The data-MC resolution agrees within -1 ± 5 fs, and is compatible with the fake-full $B^0 \to J/\Psi K^*$ study and the radiative decays.

Table 5.4: Decay time resolution for $B_s^0 \to \phi\gamma$ decays reconstructed with a prompt-ϕ and a random photon, in data and simulation, compared to the resolution in signal $B_s^0 \to \phi\gamma$ decays in simulation.

Sample	Decay time resolution
Prompt-ϕ data	74 ± 5 fs
Prompt-ϕ simulation	75 ± 2 fs
$B_s^0 \to \phi\gamma$ simulation	74.8 ± 0.4 fs

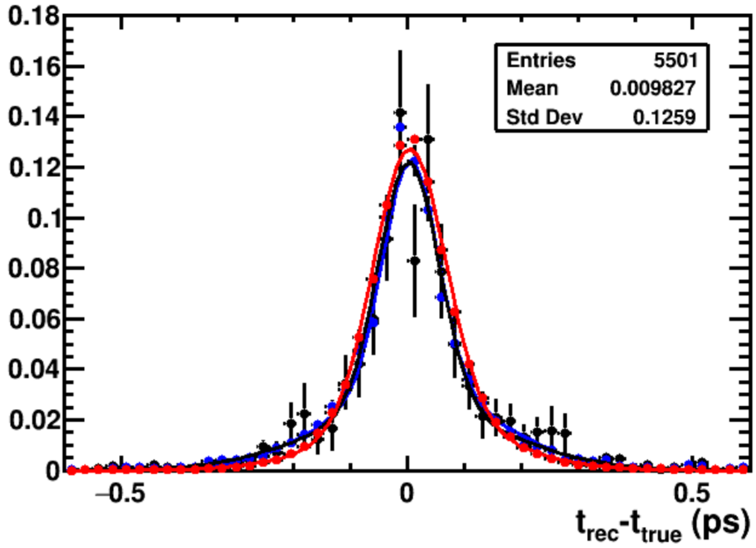


Figure 5.18: Decay time resolution for $B_s^0 \to \phi\gamma$ decays reconstructed with a prompt-ϕ and a random photon. The curves correspond to Run 1 data (black), minimum bias Monte Carlo (blue), and signal $B_s^0 \to \phi\gamma$ decays in simulation (red).

5.4 Fit strategy

A weighted unbinned maximum likelihood fit is performed to the decay time distribution, using the sFit approach (see Section 4.3.2), as explained in the following. The code of the analysis, available in reference [140], is implemented using the RooFit framework [102] with its internal implementation of the MINUIT [103] package to handle the minimization.

5.4.1 Full PDF description

The time-dependent decay rate is formally described by the following PDF:

$$F(t, q, \omega, \sigma_t) = F(t, q|\omega, \sigma_t) \cdot F(\omega) \cdot F(\sigma_t). \tag{5.22}$$

However, $F(\omega)$ and $F(\sigma_t)$ are implicitly taken into account since the mistag probability and the decay time uncertainty are considered as per-event observables. On the other hand, the conditional PDF is described by:

$$\begin{aligned} F(t, q|\omega, \sigma_t) &= \Gamma(t', q|\omega) \otimes R_A(t, t'|\sigma_t) \\ &= \Big\{ e^{-\Gamma_s t'} \Big[\cosh(\Delta\Gamma_s t'/2) - \mathcal{A}^{\Delta} \sinh(\Delta\Gamma_s t'/2) \\ &\qquad + q\,(1-2\omega)\,\mathcal{C}\cos(\Delta m_s t') - q\,(1-2\omega)\,\mathcal{S}\sin(\Delta m_s t') \Big] \Big\} \\ &\quad \otimes \Big\{ A(t_i)\, R(t, t'|\sigma_t) \Big\}. \end{aligned} \tag{5.23}$$

The physical decay rate is convoluted by an acceptance-resolution function $R_A(t, t'|\sigma_t)$ that includes acceptance effects. The resolution $R(t, t'|\sigma_t)$ is described by a double Gaussian with constant mean and a scaled per-event σ_t (see Section 5.3), while the acceptance $A(t_i)$ is binned in the reconstructed decay time. The flavour tagging decision of the event, $q = \{-1, 0, 1\}$ for $\{\bar{B}_s^0, \text{untagged}, B_s^0\}$, affects the cosinus and sinus terms of the decay rate, which are also diluted by the per-event mistag probability $(1 - 2\omega)$.

Conditional observables

Three of the parameters (q, ω and σ_t) are treated as per-event observables in the fit, taking the value that each event provides. A conditional PDF is related to a two-dimensional PDF as:

$$P(x, y) = P(x|y)P(y), \tag{5.24}$$

where $P(x|y)$ is the PDF of x given a constant y. The PDF of y is given implicitly when the observables are taken per-event.

The normalization of a conditional-PDF is different than the normalization of a 2D-PDF, a fact that is relevant for the fit minimization. Given a standard 2D-PDF, $P(x, y)$, the likelihood function is normalized by a common factor for all the n events:

$$\mathcal{L} = \frac{\prod^n P(x_i, y_i)}{\left(\int_x \int_y P(x, y)dxdy\right)^n}. \tag{5.25}$$

However, in a conditional-PDF, $P(x|y)$, each event has a different normalization factor since the y_i per-event observable changes the evaluation of the function:

$$\mathcal{L} = \prod^n \frac{P(x_i, y_i)}{\int_x P(x, y_i)dx}. \tag{5.26}$$

A parameter is treated as a conditional observable in RooFit by activating the option `ConditionalObservables(...)` in the `fitTo` routine. This option tells RooFit to normalize the PDF as in Equation 5.26 for the variables included inside the parenthesis `(...)`. The computational time spent in the unbinned fit minimization grows as a consequence of the inclusion of the per-event observables, since the normalization integrals have to be computed for each event as shown in Equation 5.26.

5.4.2 The acceptance-resolution function

Normalization integrals

The normalization integrals for a full PDF defined with a generic acceptance function,

$$F(t) = \left[\Gamma(t^{'}) \otimes R(t,t^{'})\right] A(t), \tag{5.27}$$

would give:

$$\int F(t)\, dt = \int \left[\int \Gamma(t^{'})\, R(t,t^{'})\, dt^{'}\right] A(t)\, dt, \tag{5.28}$$

an integral that has no analytical solution in the general case. The analytical integration can be implemented by binning the acceptance function, so that the acceptance integral becomes a sum, and it can be reorganized [141]:

$$\begin{aligned} \int F(t)\, dt &= \int \Gamma(t^{'}) \left\{\sum_i A(t_i) \int R(t,t^{'})\, dt\right\} dt^{'} \\ &= \int \Gamma(t^{'}) \left\{\int R_A(t,t^{'})\, dt\right\} dt^{'}. \end{aligned} \tag{5.29}$$

The initial PDF $F(t)$ can be interpreted as a decay rate $\Gamma(t^{'})$ convoluted by an acceptance-resolution function $R_A(t,t^{'})$:

$$F(t) = \Gamma(t^{'}) \otimes R_A(t,t^{'}) = \Gamma(t^{'}) \otimes \left[A(t_i)\, R(t,t^{'})\right], \tag{5.30}$$

where the normalization integrals can have an analytical solution, speeding up the fit computational time by two orders of magnitude and improving the robustness of the minimization procedure.

Binning of the acceptance function

A total of 100 bins are used for the acceptance function binning, with a higher granularity at low decay times where the function changes more rapidly, as follows:

- $n_{\rm bins}^{(1)} = 50$ in the range $[0.3, 2]$ ps, with a bin width $\Delta t = 0.034$ ps.
- $n_{\rm bins}^{(2)} = 30$ in the range $[2, 5]$ ps, with a bin width $\Delta t = 0.1$ ps.
- $n_{\rm bins}^{(3)} = 20$ in the range $[5, 10]$ ps, with a bin width $\Delta t = 0.25$ ps.

5.4.3 Technical implementation

The convolution of the physical rate with the acceptance-resolution is performed by the `RooBDecay` class included in the RooFit framework, which implements a general decay rate function with the four trigonometrical terms of Equation 5.23:

$$
\begin{aligned}
& f_0\, e^{-\Gamma t} \cosh\left(\Delta\Gamma\, t/2\right), \\
& f_1\, e^{-\Gamma t} \sinh\left(\Delta\Gamma\, t/2\right), \\
& f_2\, e^{-\Gamma t} \cos\left(\Delta m\, t\right), \\
& f_3\, e^{-\Gamma t} \sin\left(\Delta m\, t\right),
\end{aligned}
\tag{5.31}
$$

where the coefficients f_i, the parameters Γ, $\Delta\Gamma$ and Δm, and a resolution function are the inputs of the class. The information of the mixing-induced observables is encoded in the coefficients:

$$
\begin{aligned}
& f_0 = 1, \\
& f_1 = -\mathcal{A}^{\Delta}, \\
& f_2 = q\,(1-2\omega)\,\mathcal{C}, \\
& f_3 = -q\,(1-2\omega)\,\mathcal{S}.
\end{aligned}
\tag{5.32}
$$

The resolution function can be any class inheriting from `RooResolutionModel`, where the analytical convolution with the four terms of Equation 5.31 has to be implemented. If there is no analytical convolution defined in the class methods, then it is performed numerically.

The resolution model, `RooEffResModel`, represents the acceptance-resolution $R_A(t, t', \sigma_t)$. This class takes as input a `RooResolutionModel` and an efficiency function; the resolution is a `RooAddModel` that sums two `RooGaussianModel` instances (see Section 5.3), while the efficiency function is a `RooBinnedPdf` representing a binned version of the acceptance function. These classes have been imported from the source code repositories of the $B \to DX$ [142, 143] and $B \to J/\psi X$ [144] analyses, and their implementation is similar as it is explained in references [141, 145]. The `RooEffResModel` implements the analytical convolution of the `RooBDecay` terms (Equation 5.31) with any resolution model multiplied by a constant efficiency [146].

Finally, the fit is performed simultaneously to the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ samples using a `RooSimultaneous` PDF, where $B^0 \to K^{*0}\gamma$ help to constrain the acceptance function in the same way as it was explained in Chapter 4. The $B^0 \to K^{*0}\gamma$ PDF is described with an average resolution since it does not affect the shape of the acceptance function.

5.4.4 Validation of the fit

Validation with full-simulation samples

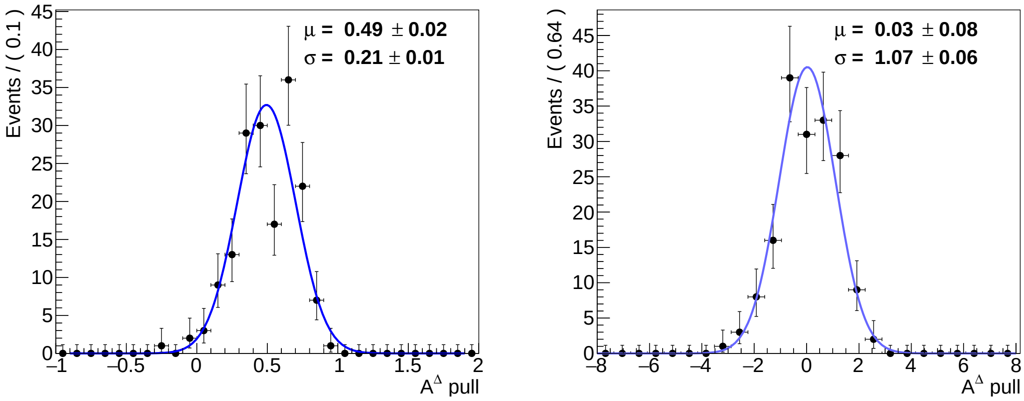
The fit strategy is validated using the full-simulation samples of radiative decays: $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$. A simultaneous fit is performed to these two samples, using the same strategy as in data, changing only the calibration of the flavour taggers. Table 5.5 shows the result of the fit. The three observables are compatible with zero, which are the generated values in this sample.

Table 5.5: Result of the fit to the simulation samples.

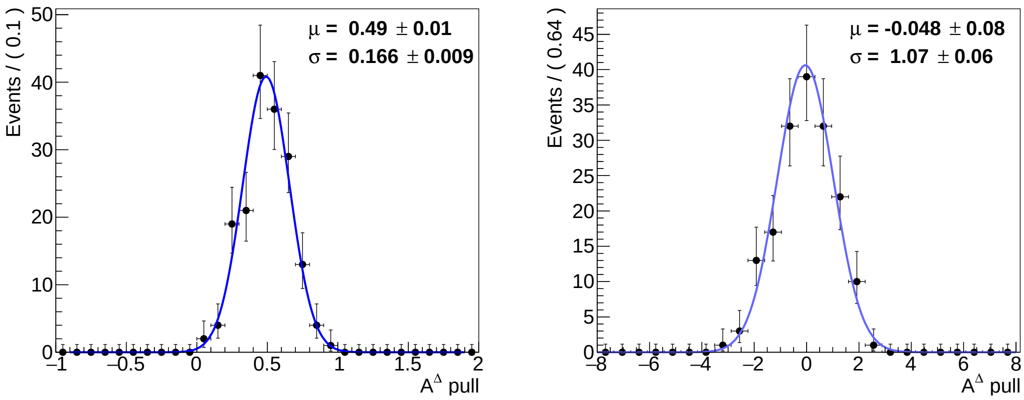
Parameter	Value
$\mathcal{A}^{\Delta}$	0.031 ± 0.047
$\mathcal{S}$	0.022 ± 0.019
$\mathcal{C}$	0.008 ± 0.019
b (acceptance)	0.739 ± 0.003
c (acceptance)	0.049 ± 0.004

Validation with toy Monte Carlo experiments

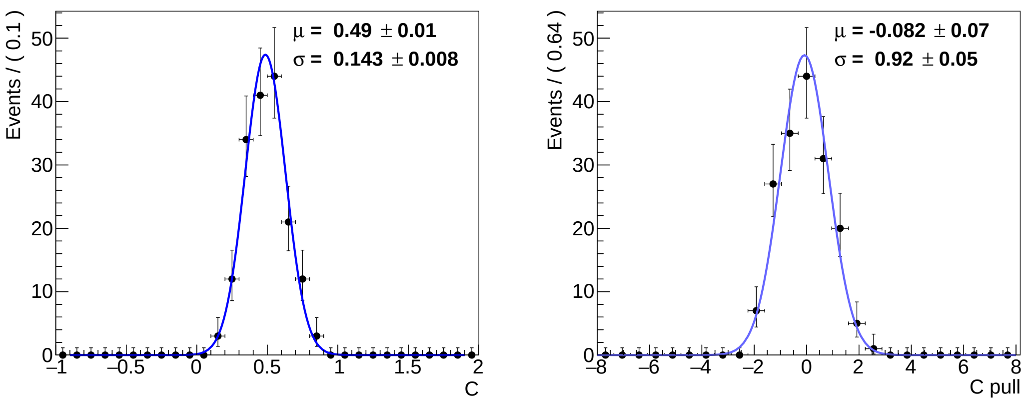
The fit strategy is tested with toy Monte Carlo experiments. Each experiment is generated with $N_{B_s^0\to\phi\gamma} = 5300$ events and the decay time acceptance used in the nominal fit, in the scenario $\mathcal{A}^{\Delta} = \mathcal{S} = \mathcal{C} = 0.5$, where only the $B_s^0 \to \phi\gamma$ distribution is generated and fitted. The rest of the observables (mistag probability, decay time resolution) are taken randomly from the data sample. Figure 5.19 shows the dispersion of the fitted values and the pull distribution for each of the three observables. All the dispersion distributions are compatible with the generated value, 0.5, while the pull distributions present no biases (μ compatible with zero) and a correctly evaluated uncertainty (σ compatible with one), within their statistical precision.



(a) $\mathcal{A}^{\Delta}$



(b) $\mathcal{S}$



(c) $\mathcal{C}$

Figure 5.19: Simulated experiments generated with $\mathcal{A}^{\Delta} = \mathcal{S} = \mathcal{C} = 0.5$, performed with the fit strategy of the tagged analysis. The dispersion of the observables $\mathcal{A}^{\Delta}$ (a), $\mathcal{S}$ (b), and $\mathcal{C}$ (c), is shown at the left-hand side of the figure, while the pull distributions are shown at the right-hand side. A Gaussian fit is superimposed.

5.5 Sensitivity studies

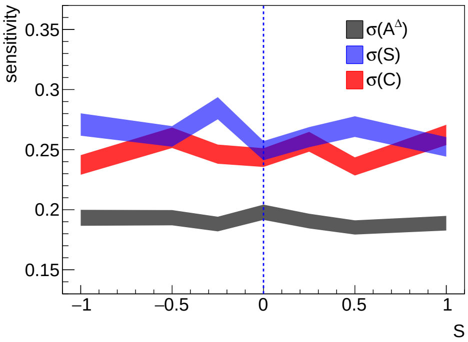
This section studies the effect of the different physical and reconstruction quantities to the uncertainties and central values of the observables $\{\mathcal{A}^{\Delta}, \mathcal{C}, \mathcal{S}\}$, with the help of toy Monte Carlo techniques.

To perform the toys, the total number of events was set to the expected Run 1 yield for $B_s^0 \to \phi\gamma$, $N_{\text{Run 1}} = 5300$, the mistag rate was taken to be $\omega \simeq 0.41$, and the decay time resolution was modelled by a Gaussian function with constant width $\sigma_{\text{res.}} \simeq 70$ fs for all the events. These values correspond to the expected average mistag fraction $\langle\omega\rangle$ and an effective resolution width for the whole sample. The generated values for the parameters were chosen to be $\mathcal{A}^{\Delta} = \mathcal{S} = \mathcal{C} = 0.5$, to have enough sensitivity to the different effects. Note that for other central values the sensitivity effects can be rather different.

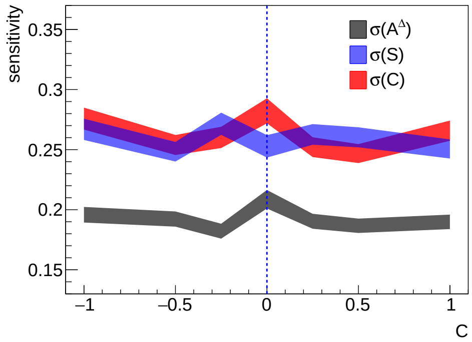
5.5.1 Effect introduced by $\mathcal{A}^{\Delta}$, $\mathcal{S}$ and $\mathcal{C}$

In this section, the dependency of the uncertainties $\{\sigma_{\mathcal{A}^{\Delta}}, \sigma_{\mathcal{C}}, \sigma_{\mathcal{S}}\}$ with the central value of the observables $\{\mathcal{A}^{\Delta}, \mathcal{C}, \mathcal{S}\}$ is studied, motivated by the strong dependence between $\mathcal{A}^{\Delta}$ and its uncertainty found in the untagged analysis, as detailed in Section 4.4.

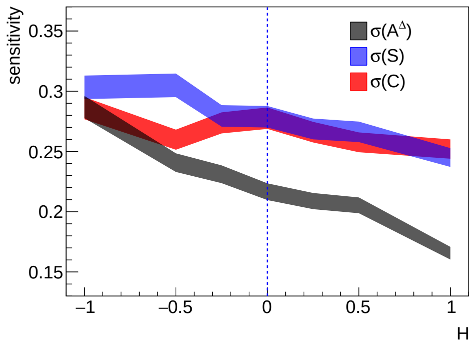
Several hundred pseudoexperiments are generated, with the three observables ranging from -1 to 1. The rest of the inputs (ω, σ_t) are identical to the values of the previous section. The results are shown in Figure 5.20. A clear dependency is found between the three uncertainties and the $\mathcal{A}^{\Delta}$ value, although it is more pronounced in the case of $\sigma_{\mathcal{A}^{\Delta}}$. On the other hand, the three uncertainties are not affected by the central values of $\mathcal{S}$ and $\mathcal{C}$.



(a) Sensitivities vs $\mathcal{S}_{\text{gen}}$



(b) Sensitivities vs $\mathcal{C}_{\text{gen}}$



(c) Sensitivities vs $\mathcal{A}^{\Delta}_{\text{gen}}$

Figure 5.20: The uncertainty of the observables $\{\sigma_{\mathcal{A}^{\Delta}}, \sigma_{\mathcal{C}}, \sigma_{\mathcal{S}}\}$ as a function of their central values $\{\mathcal{A}^{\Delta}, \mathcal{C}, \mathcal{S}\}$, in the $\mathcal{A}^{\Delta} = \mathcal{S} = \mathcal{C} = 0.5$ scenario.

5.5.2 Effect introduced by the mistag rate and the decay time resolution

The dilution factor $\mathcal{D} = (1-2\omega)$ introduced by the flavour tagging algorithms in the probability density function (Equation 5.23) degrades the sensitivity to the $\mathcal{S}$ and $\mathcal{C}$ parameters, by reducing the amplitude of the $\cos(\Delta m_s t)$ and $\sin(\Delta m_s t)$ terms. The decay time resolution also degrades the $\mathcal{S}$ and $\mathcal{C}$ sensitivity; the cosinus and sinus are highly oscillating terms, thus the migration of events affects the amplitude of the oscillation.

The dilution in the oscillation amplitude introduced after each reconstruction effect (mistag rate, acceptance and resolution) can be appreciated in Figure 5.21, for a tagged B particle ($q = 1$) in the case $\mathcal{S} = \mathcal{C} \simeq 0.5$. The sinus and cosinus terms of the physical decay rate (Equation 5.23) produce a highly oscillating distribution, represented as the red curve. The mistag rate introduces a dilution of the amplitude, decreasing the sensitivity to the observables. After the acceptance effects produced by the selection,

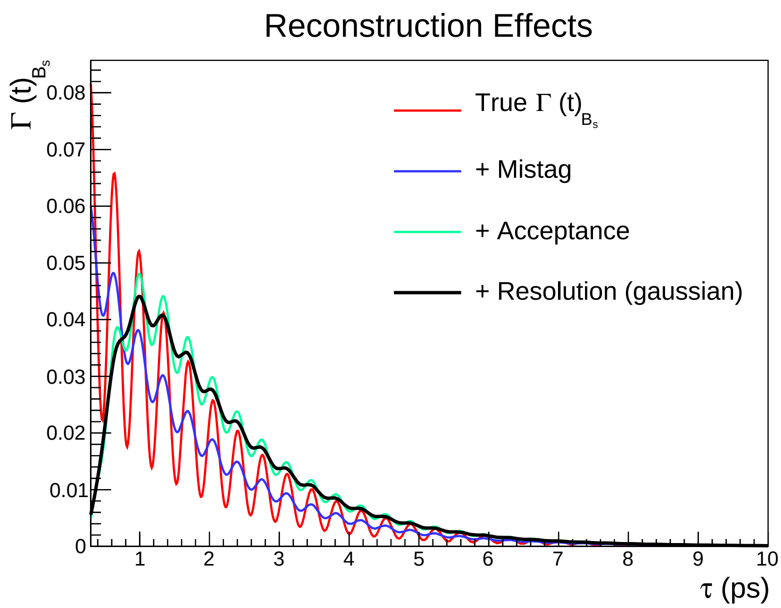


Figure 5.21: Generated decay time distribution after each reconstruction effect, for the $B_s^0 \to \phi\gamma$ decay with an initial B_s^0 particle, for the case $\mathcal{S} = \mathcal{C} \simeq 0.5$. The red pattern is the physical decay rate (from Equation 5.23). The mistag, acceptance and resolution effects are added one-by-one, and are represented by the blue, green and black curves, respectively.

the resolution (black curve) reduces further the amplitude. Thus, both the mistag rate and the resolution degrade the sensitivity of $\mathcal{S}$ and $\mathcal{C}$, while the acceptance affects mostly the $\mathcal{A}^{\Delta}$ parameter.

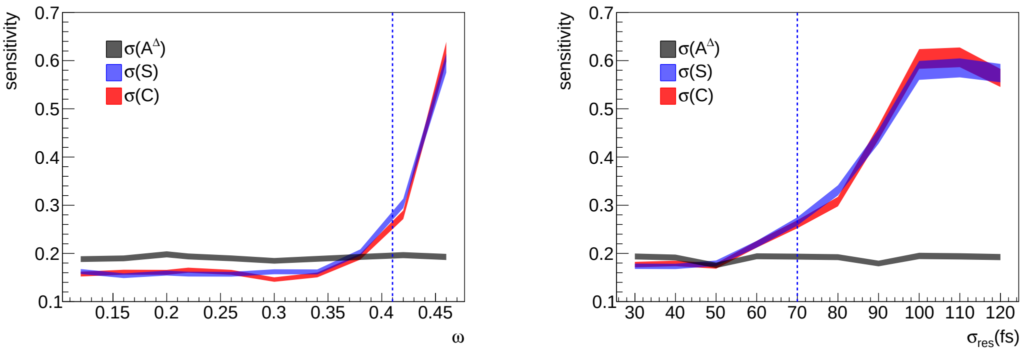
Effect in the uncertainties

In this section, a set of toy Monte Carlo experiments are produced to study the effect of the mistag rate and decay time resolution on the statistical uncertainties $\sigma_{\mathcal{C}}$ and $\sigma_{\mathcal{S}}$. The pseudoexperiments are generated fixing the central values to $\sigma_{\mathcal{A}^{\Delta}} = \sigma_{\mathcal{S}} = \sigma_{\mathcal{C}} = 0.5$, the mistag rate to $\omega = 0.41$, and the decay time resolution width to $\sigma_t = 70$ fs, described by a single Gaussian.

The pseudoexperiments are repeated modifying the value of ω or σ_t in both generation and fitting PDFs, maintaining fixed the other variable. The results are shown in Figure 5.22. As expected, the $\mathcal{S}$ and $\mathcal{C}$ uncertainties improve with a lower mistag fraction and a narrower resolution, while $\mathcal{A}^{\Delta}$ remains insensitive to these two factors. For the case in which $\omega = 0.41$ and $\sigma_t = 70$ fs, the uncertainties of the three observables are found to be comparable:

$$\sigma_{\mathcal{A}^{\Delta}} \simeq \sigma_{\mathcal{S}} \simeq \sigma_{\mathcal{C}} \simeq 0.2. \tag{5.33}$$

Note that this uncertainty corresponds to the signal contribution, and does not include the contribution from the background subtraction procedure.



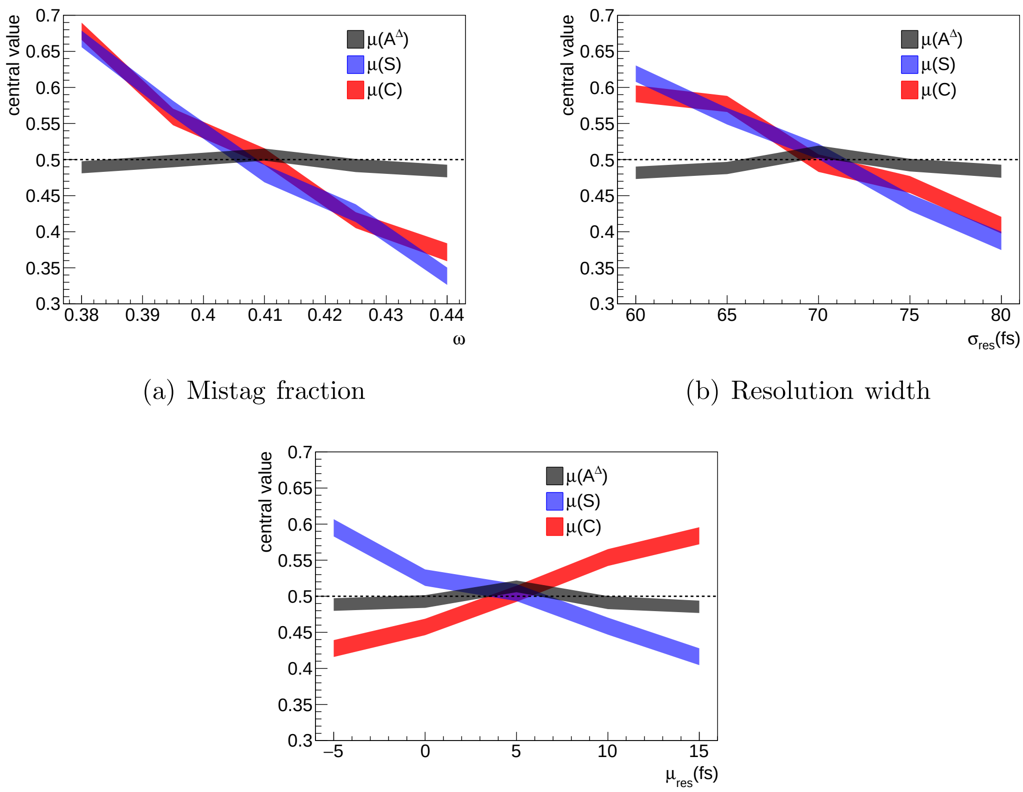
(a) Mistag fraction

(b) Resolution width

Figure 5.22: The uncertainties of the observables as a function of the (a) mistag fraction and (b) decay time resolution width, for the $\mathcal{A}^{\Delta} = \mathcal{S} = \mathcal{C} = 0.5$ scenario. The dotted blue line corresponds to the value used to produce the other plot.

Effect in the central values

Another set of toy Monte Carlo experiments is conducted to assess the possible biasing effects when making a mistake on the determination of the mistag fraction ω or the resolution parameters μ_t and σ_t. A scanning is performed in the three variables, but this time only the PDF of generation is modified, fixing the PDF of fitting to $\omega = 0.41$, $\mu_t = 5$ fs and $\sigma_t = 70$ fs. The results are shown in Figure 5.23. The $\sigma_{\mathcal{C}}$ and $\sigma_{\mathcal{S}}$ uncertainties are sensitive to the three variables (ω, μ_t and σ_t), while $\sigma_{\mathcal{A}^\Delta}$ remains insensitive to all of them. In the $\mathcal{A}^\Delta = \mathcal{S} = \mathcal{C} = 0.5$ scenario, a mistake on 1 fs in the determination of the resolution mean or width modifies the value of the parameters by 0.01 (in absolute value), while a mistake of $\Delta\omega = 0.01$ modifies by 0.05 the observables.



(a) Mistag fraction

(b) Resolution width

(c) Resolution mean

Figure 5.23: The bias in the observables produced by (a) the mistag fraction ω, (b) the decay time resolution width σ_t and (c) the decay time resolution mean μ_t. The unbiased case is marked by a dotted line for the $\mathcal{A}^\Delta = \mathcal{S} = \mathcal{C} = 0.5$ scenario.

5.6 Systematic uncertainties

The summary of the relevant systematic contributions is presented in Table 5.6, where the systematics of the three observables, $\mathcal{A}^{\Delta}$, $\mathcal{S}$ and $\mathcal{C}$, are quoted. The following paragraphs detail the calculation of each contribution in depth.

Table 5.6: Summary of the systematic uncertainties in the tagged analysis of $B_s^0 \to \phi\gamma$ decays with Run 1 data.

	Systematic source	$\sigma(\mathcal{A}^{\Delta})$	$\sigma(\mathcal{S})$	$\sigma(\mathcal{C})$
External measurements	$(\Gamma_s, \Delta\Gamma_s)$	0.067	0	0
	Γ_d	0.040	0	0
	Δm_s	0	0.008	0.013
Decay time	Acceptance: MC limited statistics	0.067	0	0
	Acceptance: modelling	0.028	0	0
	Resolution: MC limited statistics	0	0.001	0.002
	Resolution: modelling	0	0.034	0.026
Mistag probability	OS tagger calibration	0	0.010	0.010
	SS tagger calibration	0	0.049	0.031
Background subtraction	Mass modelling: signal	< 0.01	< 0.01	< 0.01
	Mass modelling: combinatorial	0.034	< 0.01	< 0.01
	Mass modelling: partial	0.093	0.010	0.020
	Peaking backgrounds	0.036	0.018	0.030
	Mass-time correlation	0.110	< 0.01	< 0.01
Total	$\sigma_{\text{ext.}}$	0.078	0.008	0.013
	$\sigma_{\text{syst.}}$	0.170	0.064	0.055

5.6.1 External measurements

Among the input parameters from the HFLAV world-average measurements, listed in Table 4.4, $\Delta m_s = 17.757 \pm 0.021$ ps^{-1} [39] is the only one affecting the measurement of $\mathcal{C}$ and $\mathcal{S}$, since it appears inside the trigonometrical terms $\cos(\Delta m_s\, t)$ and $\sin(\Delta m_s\, t)$ of the time-dependent decay rate.

The systematic uncertainty introduced by Δm_s is obtained from a Gaussian-constrained maximum likelihood fit, as it was done in the untagged $\mathcal{A}^{\Delta}$ measurement for the systematic contribution of external measurements (Chapter 4). The $\sigma_{\mathcal{S}}$ and $\sigma_{\mathcal{C}}$ uncertainties with and without the Gaussian-constrained fit are subtracted in quadrature $\sqrt{\sigma^2_{\text{with}} - \sigma^2_{\text{without}}}$, obtaining:

$$\begin{aligned} \sigma_{\text{ext.}}(\mathcal{S}) &= 0.008, \\ \sigma_{\text{ext.}}(\mathcal{C}) &= 0.013, \end{aligned} \tag{5.34}$$

which corresponds to the total external measurements contribution.

5.6.2 Flavour tagging calibration

The flavour tagging algorithms introduce a systematic effect through the uncertainty of the parameters used to calibrate the mistag probability. To estimate the effect of the calibration parameters p_0 and p_1, they are inserted as Gaussian-constraints in the fit. A contribution of $\sigma_{\mathcal{C}} = \sigma_{\mathcal{S}} = 0.01$ is obtained for the OS tagger, while $\sigma_{\mathcal{S}} = 0.049$ and $\sigma_{\mathcal{C}} = 0.031$ are found for the SS tagger. These numbers are obtained considering that all the events of the sample are either OS- or SS-tagged, so they represent an overestimation of the real systematic effect, which would be situated between the two.

5.6.3 Decay time acceptance

The definition of the decay time acceptance in the tagged analysis is different than the untagged acceptance, where it was not binned. The effect of the binning has been tested by increasing the number of bins one order of magnitude. The difference, 0.02 for $\mathcal{A}^{\Delta}$, is the additional systematic contribution. The observables $\mathcal{S}$ and $\mathcal{C}$ remain immune to the acceptance effects, hence they are not affected by the limited statistics of the Monte Carlo simulation.

5.6.4 Decay time resolution

The systematic contribution of the decay time resolution has been splitted into several sources:

- The limited statistics of the full-simulation could induce an uncertainty in the resolution parameters extracted from simulation, in particular the width scaling of the double Gaussian. It has been estimated from a Gaussian-constrained fit, obtaining a residual difference of < 0.01.
- As it has been discussed in Section 5.3, the mean of the double Gaussian depends on the decay time, $\mu(t)$, although in the nominal fit it is considered to be constant. The fit is repeated using the full decay-time dependency, obtaining a difference of 0.02 and 0.01 in $\mathcal{S}$ and $\mathcal{C}$, respectively.
- The description of the resolution may be different in the 2011 and 2012 periods. The fit is repeated separating the samples of the two years, using a different width scaling. The increase in uncertainty, of 0.028 and 0.024 for $\mathcal{S}$ and $\mathcal{C}$, is taken as the systematic uncertainty.

The modelling of the resolution, as quoted in the summary Table 5.6, includes these last two sources.

5.6.5 Background subtraction

The systematic uncertainties of the background subtraction are obtained in the same way as in the untagged analysis, Section 4.5. This contribution, which is one of the main systematic sources in the measurement of the untagged observable $\mathcal{A}^{\Delta}$, becomes a residual source in the measurement of the tagged observables $\mathcal{S}$ and $\mathcal{C}$. This is a consequence of the background affecting the whole decay time distribution, but not the rapid oscillations. The peakings are the main source of background systematics for the tagged observables, representing 0.018 and 0.030 for $\mathcal{S}$ and $\mathcal{C}$, respectively. The partial are a secondary source, representing only about 0.01 and 0.02, respectively. The rest of the backgrounds, and the mass-time correlation, do not affect the value of these observables.

5.7 Results

The results are presented in Table 5.7, where the three observables are quoted together with the parameters of the acceptance for the control channel.

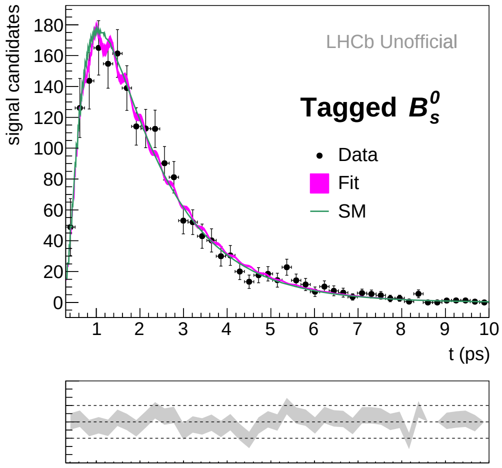
Table 5.7: Results of the tagged analysis, in a simultaneous fit to $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$. The acceptance values correspond to $A_{K^{*0}\gamma}$.

Parameter	Value
$\mathcal{A}^{\Delta}$	$-0.551\pm^{0.330}_{0.359}$
$\mathcal{S}$	0.441 ± 0.252
$\mathcal{C}$	0.122 ± 0.238
b (ps)	0.816 ± 0.015
c (ps^{-1})	0.070 ± 0.011

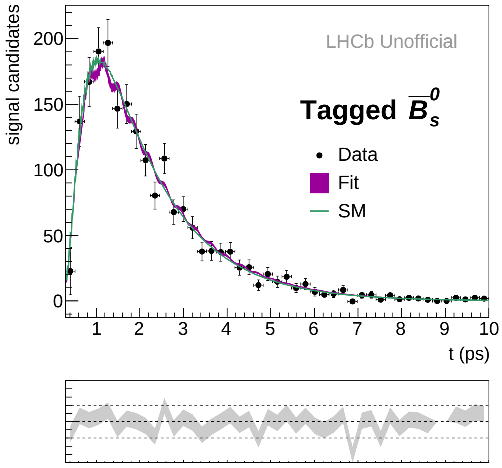
Figure 5.24 shows the fit projected into each one of the categories: particle, antiparticle and untagged for $B_s^0 \to \phi\gamma$, and the control channel which corresponds to $B^0 \to K^{*0}\gamma$. The pull distributions below the plots do not show any particular pattern. Considering the systematic uncertainties, the results are:

$$\begin{aligned}
\mathcal{A}^{\Delta}_{\phi\gamma} &= -0.551\,^{+0.330}_{-0.359}(\text{stat.}) \pm 0.170(\text{syst.}) \pm 0.067(\text{ext.}), \\
\mathcal{S}_{\phi\gamma} &= 0.441 \pm 0.252(\text{stat.}) \pm 0.064(\text{syst.}) \pm 0.008(\text{ext.}), \\
\mathcal{C}_{\phi\gamma} &= 0.122 \pm 0.238(\text{stat.}) \pm 0.055(\text{syst.}) \pm 0.013(\text{ext.}),
\end{aligned} \tag{5.35}$$

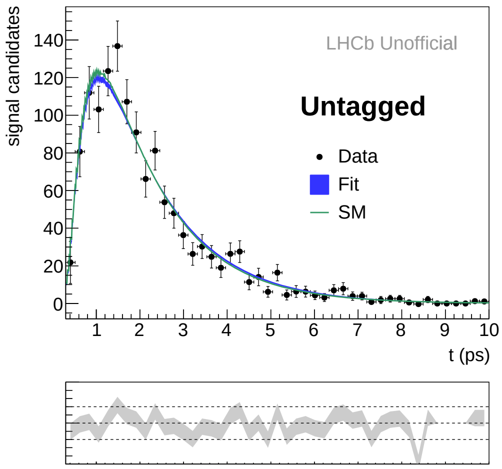
where the statistical contribution of the background subtraction is calculated to be 0.15 in the $\mathcal{S}$ and $\mathcal{C}$ observables, and the control channel do not modify their uncertainty. At this point, the measurement is completely dominated by the statistical uncertainty.



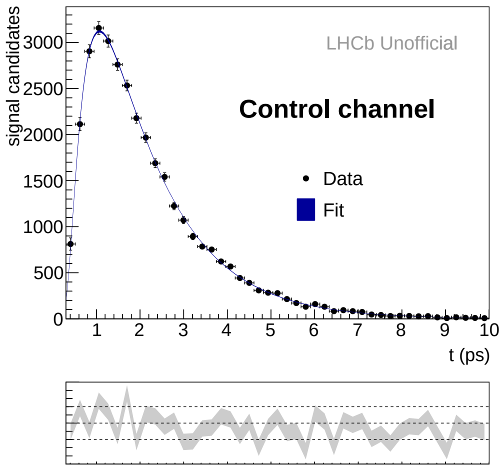
(a) Tagged B_s^0



(b) Tagged $\bar{B}_s^0$



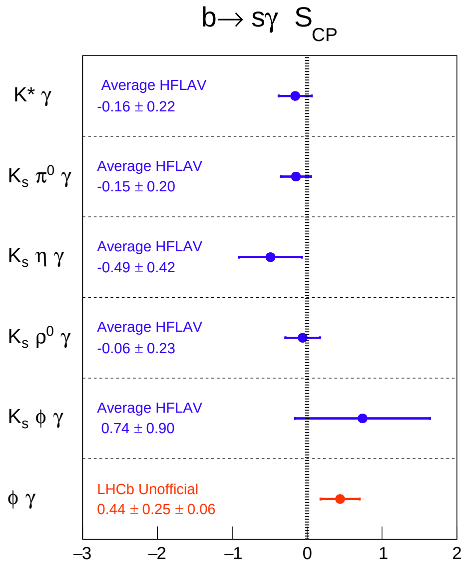
(c) Untagged



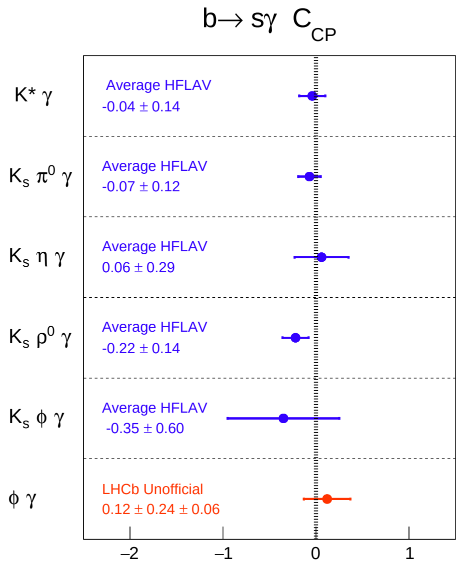
(d) Control channel

Figure 5.24: Simultaneous fit to the data samples of $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays, with the tagged analysis. The plots are separated by categories: (a) tagged $B_s^0 \to \phi\gamma$ decays, (b) tagged $\bar{B}_s^0 \to \phi\gamma$ decays, (c) untagged $B_s^0 \to \phi\gamma$ decays, and (d) $B^0 \to K^{*0}\gamma$ decays. The best estimation and the SM prediction curves are superimposed over the data points.

The list of all the $S_{\rm CP}$ and $C_{\rm CP}$ measurements sensitive to the $b \to s\gamma$ transition was shown in Section 1.5.2, when discussing the experimental status. Figure 5.25 extends the list with the current results, with the difference that this is the first measurement from a B_s^0 meson. The uncertainties of the current measurement are competitive to the measurements performed by the B-factories.



(a) $S_{\rm CP}$ measurements



(b) $C_{\rm CP}$ measurements

Figure 5.25: Summary of the $S_{\rm CP}$ and $C_{\rm CP}$ measurements in $b \to s\gamma$ transitions, adding the current measurement. Decays from B^0 mesons are shown in blue.

5.8 Constraints on new physics

The implications of the current measurements on new physics constraints can be presented in a common framework together with other radiative decay measurements. The effective weak Hamiltonian theory, introduced in Section 5.8.1, is a theoretical tool to study the phenomenology of flavour processes. Radiative decays are sensitive to the C_7 and C_7' Wilson coefficients. The constraints on these coefficients are shown in Section 5.8.2, for the current measurement alone and for all the radiative decay measurements together.

5.8.1 Effective weak Hamiltonian

The phenomenology of rare B decays involving flavour-changing hadron transitions can be studied using the effective weak Hamiltonian formalism, where the scattering matrix S can be expanded in terms of local operators $\mathcal{O}$:

$$\langle F|\, iS\, |I\rangle = 4\frac{G_F}{\sqrt{2}} \sum_i C_i(\mu) \langle F|\, \mathcal{O}_i(\mu)\, |I\rangle + \dots \tag{5.36}$$

In the case of $b \to s\gamma$ transitions, the Hamiltonian can be written as [147–149]:

$$H_{\text{eff}} = -\frac{4\, G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_{i=1}^{8} \Big(C_i(\mu)\mathcal{O}_i(\mu) + C_i^{'}(\mu)\mathcal{O}_i^{'}(\mu) \Big) , \tag{5.37}$$

where G_F is the Fermi constant, V_{ij} are the CKM matrix elements, the operators $\mathcal{O}_i(\mu)$ represent the effective structure of the interaction, and the Wilson coefficients $C_i(\mu)$ are the effective coupling constants for a given energy scale μ. In the $b \to s$ transition, the physical process occurs at energies of the order of the b-quark pole mass, $m_b = 4.78 \pm 0.06$ GeV [20], so the scale is usually chosen to be $\mu_b = 4.8$ GeV. $\mathcal{O}_{1-6}$ are the four-quark operators, while $\mathcal{O}_7$ and $\mathcal{O}_8$ are the electromagnetic and chromomagnetic dipole operators. New physics effects are expected to show more clearly in the latter two, since the other operators receive SM contributions at tree level.

The formalism arises from an operator product expansion (OPE) [150] in the local operators, renormalized at a given scale μ which sets the threshold between the short- and long-distance regimes [151–153]. The operators

$\mathcal{O}_i(\mu)$ contain the physics contributions from scales below μ, corresponding to the non-perturbative part, while the Wilson coefficients $C_i(\mu)$ receive the contribution from scales above μ, which can be calculated with perturbative techniques. Consequently, $C(\mu)$ is affected by processes mediated by the weak bosons $W^{\pm}$ and Z^0, the t-quark, and new heavy particles, as well as short-distance QCD effects. Physics beyond the SM could change the value of the Wilson coefficients due to the contribution of new heavy particles, but also introduce new operator structures not covered by the SM.

The electromagnetic operators have the following structure:

$$\begin{aligned} \mathcal{O}_7 &= \frac{e}{16\pi^2} m_b \left(\bar{s}\sigma_{\mu\nu} P_R b\right) F^{\mu\nu}, \\ \mathcal{O}_7' &= \frac{e}{16\pi^2} m_b \left(\bar{s}\sigma_{\mu\nu} P_L b\right) F^{\mu\nu}, \end{aligned} \tag{5.38}$$

where $P_{R,L} = (1 \pm \gamma_5)/2$ are the chiral projection operators. Primed operators have the chirality flipped, with vanishing coefficients for $\mathcal{O}_{1-6}$ and suppressed coefficients for $\mathcal{O}_{7-8}$. The primed Wilson coefficient C_7' is supressed by the ratio of the light over the heavy quark masses:

$$C_7^{\prime\text{SM}} = \frac{m_s}{m_b} C_7^{\text{SM}} \simeq 0.02\, C_7^{\text{SM}}. \tag{5.39}$$

The $b \to s\gamma$ transition is sensitive to the $\mathcal{O}_7$ ($\mathcal{O}_7'$) electromagnetic operators, related to the $\bar{s}_L \sigma_{\mu\nu} b_R$ ($\bar{s}_R \sigma_{\mu\nu} b_L$) structures. Due to angular momentum conservation, the emitted photon will be left-handed (right-handed). Since the W boson couples exclusively to left-handed quarks, a chirality flip is needed in one of the quarks of the mentioned transition, introduced by the mass operators: $b_R \to b_L$ will happen with a strength m_b (and $s_R \to s_L$ with a strength m_s). Therefore, the fraction of helicity amplitudes is given by

$$\frac{A_R}{A_L} = \frac{m_s}{m_b},$$

and the photon emitted in the $b \to s\gamma$ transition is predominantly left-handed (right-handed in the $\bar{b} \to \bar{s}\gamma$ transition and the $B_s^0 \to \phi\gamma$ decay).

5.8.2 Constraints on the C_7-C_7' Wilson coefficients

The studies presented in this section are performed with `flavio` [154], a Python package for flavour physics phenomenology, which in turn uses the `wilson` [155] package to calculate the constraints on the Wilson coefficients. The study adopts a similar strategy as in reference [156].

The constraints of the current measurement on the Wilson coefficient $C_7'^{\rm NP}$ are shown in Figure 5.26, where the Standard Model contribution has been subtracted, $C_7'^{\rm NP} = C_7' - C_7'^{\rm SM}$. The observables $\mathcal{A}^{\Delta}_{\phi\gamma}$ and $\mathcal{S}_{\phi\gamma}$ constrain the real and imaginary parts of the coefficients, respectively:

$$\begin{aligned} \mathcal{A}^{\Delta}_{\phi\gamma} &\simeq \frac{2\,\mathrm{Re}\left(e^{-i\phi_s}\, C_7 C_7'\right)}{|C_7|^2 + |C_7'|^2}, \\ \mathcal{S}_{\phi\gamma} &\simeq \frac{2\,\mathrm{Im}\left(e^{-i\phi_s}\, C_7 C_7'\right)}{|C_7|^2 + |C_7'|^2}, \end{aligned} \tag{5.40}$$

complementing each other in the complex plane. The plot shows the 1σ contours for the individual constraints, and the 1σ and 2σ contours for the combination. The 2D combination constraint is compatible with the SM within 1.8σ, with the best point being $\left\{\mathrm{Re}\, C_7'^{\rm NP}, \mathrm{Im}\, C_7'^{\rm NP}\right\} = \{0.14, -0.10\}$.

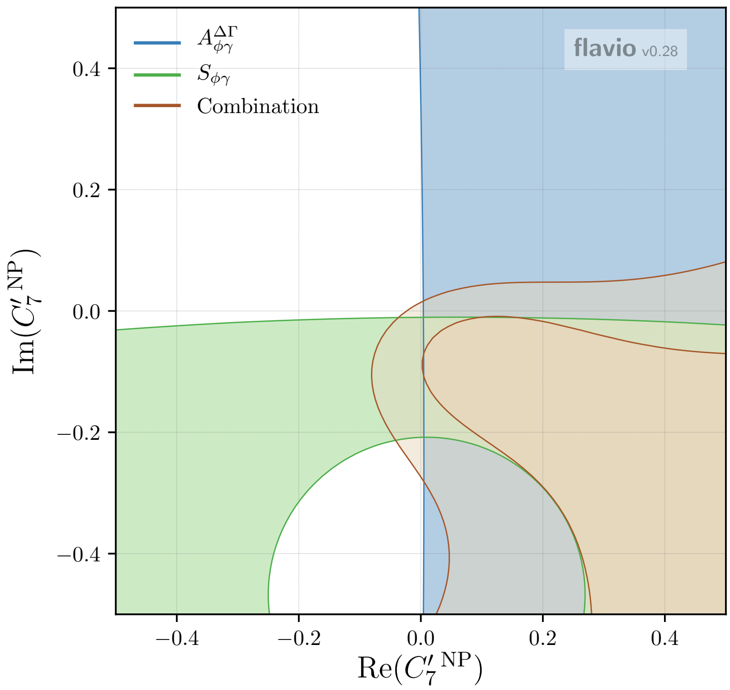


Figure 5.26: Constraints on the $C_7'^{\rm NP}$ complex plane, from the measurement of $\mathcal{A}^{\Delta}_{\phi\gamma}$ and $\mathcal{S}_{\phi\gamma}$. The 1σ and 2σ contours are shown for the combination, in brown.

The constraints from all the radiative decay measurements, including the current analysis, are shown in Figure 5.27. The plot includes the following measurements:

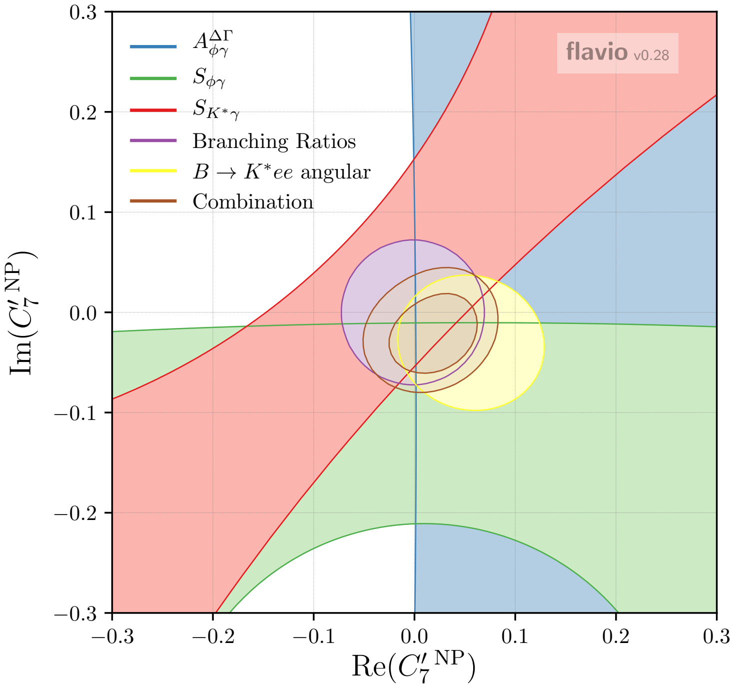


Figure 5.27: Constraints on the $C_7'^{\,\mathrm{NP}}$ complex plane, from all the radiative decay measurements.

- $S_{K^{*0}\gamma}$ makes reference to the analysis of the time-dependent CP asymmetry in $B^0 \to K^{*0}(\to K_s^0\pi^0)\gamma$ decays. The red band correspond to the combination of the Belle [53] and BaBar [52] measurements, which is sensitive to:
$$S_{K^{*0}\gamma} \simeq \frac{2\,\mathrm{Im}\left(e^{-i\phi_d}\, C_7 C_7'\right)}{|C_7|^2 + \left|C_7'\right|^2}. \tag{5.41}$$
Since the mixing angle of B^0 decays is sizeable, $\phi_d \simeq 43°$, the resulting constraint band is diagonally shaped.

- The branching ratios constraint, in purple, includes the inclusive $\mathcal{B}(B \to X_s\gamma)_{E_\gamma>1.6\text{ GeV}}$, and the exclusive $\mathcal{B}(B^+ \to K^{*+}\gamma)$, $\mathcal{B}(B_s^0 \to \phi\gamma)$ and $\mathcal{B}(B^0 \to K^{*0}\gamma)$ branching ratios. They are sensitive to the total amplitude:
$$\mathcal{B}(B \to X_s\gamma) \propto |C_7|^2 + |C_7'|^2, \tag{5.42}$$
hence the projected constraint band in the C_7' complex plane is circularly shaped.

- The angular analysis of $B^0 \to K^{*0}e^+e^-$ decays at $q^2 = 0$, performed by LHCb [46], can access to the transverse asymmetries sensitive to the real and imaginary parts:

$$A_{\mathrm{T}}^{(2)}(q^2 \to 0) = \frac{2\,\mathrm{Re}\left(C_7 C_7'^*\right)}{|C_7|^2 + |C_7'|^2},$$
$$A_{\mathrm{T}}^{\mathrm{Im}}(q^2 \to 0) = \frac{2\,\mathrm{Im}\left(C_7 C_7'^*\right)}{|C_7|^2 + |C_7'|^2}. \tag{5.43}$$

The 2D combination constraint is compatible with the SM within 0.8σ. The measurement of $\mathcal{A}^{\Delta}_{\phi\gamma}$ and $\mathcal{S}_{\phi\gamma}$ displaces the best point towards the NP region, with the values being:

$$\left\{\mathrm{Re}\,C_7'^{\,\mathrm{NP}}, \mathrm{Im}\,C_7'^{\,\mathrm{NP}}\right\} = \{0.011, -0.004\} \quad (\text{without } \mathcal{A}^{\Delta}_{\phi\gamma} \text{ and } \mathcal{S}_{\phi\gamma}),$$
$$\left\{\mathrm{Re}\,C_7'^{\,\mathrm{NP}}, \mathrm{Im}\,C_7'^{\,\mathrm{NP}}\right\} = \{0.021, -0.024\} \quad (\text{with } \mathcal{A}^{\Delta}_{\phi\gamma} \text{ and } \mathcal{S}_{\phi\gamma}). \tag{5.44}$$

For completeness, the constraints on the real parts of C_7^{NP} and $C_7'^{\,\mathrm{NP}}$ is shown in Figure 5.28, for all the radiative decay measurements. The real part of C_7^{NP} is basically constrained by the branching ratio measurements, in agreement with the SM prediction.

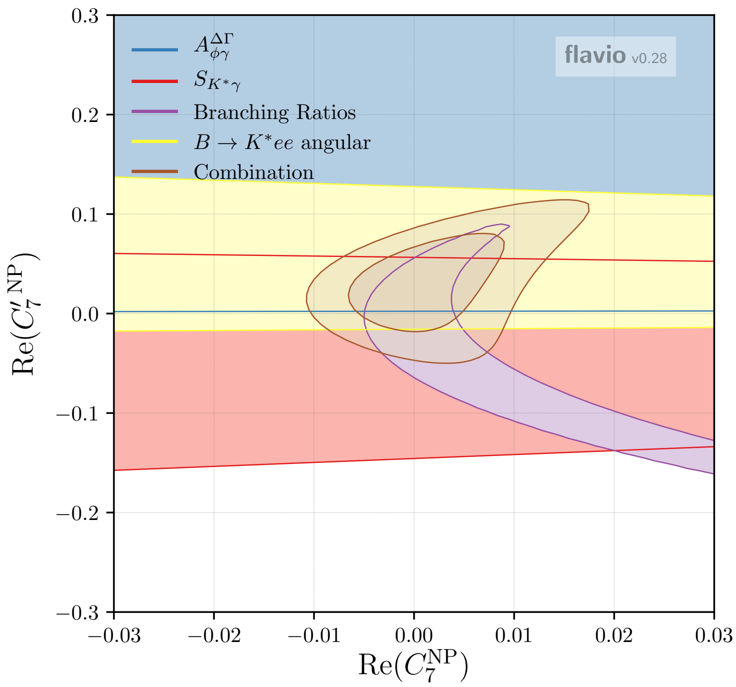


Figure 5.28: Constraints on the real parts $\mathrm{Re}(C_7^{\mathrm{NP}})$ and $\mathrm{Re}(C_7'^{\,\mathrm{NP}})$, from all the radiative decay measurements.

6 Conclusion

The mixing-induced and CP-violating observables $\mathcal{A}^{\Delta}_{\phi\gamma}$, $\mathcal{S}_{\phi\gamma}$ and $\mathcal{C}_{\phi\gamma}$ are measured for the first time, from a fit to the time-dependent decay rate of $B^0_s \to \phi\gamma$ decays. The measurement is performed with an integrated luminosity of 3 fb^{-1} recorded by the LHCb in pp collisions, at centre-of-mass energies of $\sqrt{s} = 7$ TeV in 2011 and $\sqrt{s} = 8$ TeV in 2012. Two analyses are presented, with (tagged) and without (untagged) identifying the flavour of the initial meson. The selection of candidates is common for the two analyses, and is optimized to maximize the signal significance, obtaining a total of $5\,266 \pm 104$ reconstructed $B^0_s \to \phi\gamma$ candidates. The flavour-specific $B^0 \to K^{*0}\gamma$ decay, with a yield about six times larger, is used as a control channel.

The untagged analysis performed for this thesis led to the first measurement of $\mathcal{A}^{\Delta}_{\phi\gamma} = -0.98^{+0.46}_{-0.52}{}^{+0.23}_{-0.20}$, reported in [59], corresponding to the first measurement of an observable sensitive to the photon polarization in a B^0_s decay. Compared to the published result, the reconstruction and selection of candidates is further improved, increasing the total yield by about 20%. The obtained result is:

$$\mathcal{A}^{\Delta}_{\phi\gamma} = -0.555\,^{+0.331}_{-0.359}(\text{stat.}) \pm 0.170(\text{syst.}) \pm 0.078(\text{ext.}),$$

where the first uncertainty is statistical, the second is systematic, and the third is the contribution from external measurements. The value is compatible with the Standard Model prediction, $\mathcal{A}^{\Delta,\,\text{SM}}_{\phi\gamma} = 0.045 \pm 0.029$, within 1.5 standard deviations.

The observables $\mathcal{S}_{\phi\gamma}$ and $\mathcal{C}_{\phi\gamma}$ are measured in an analysis where the flavour tagging information is used for the first time in a radiative decay at LHCb, with an effective tagging performance of $4.58\pm0.13\%$. The observable $\mathcal{S}_{\phi\gamma}$ is sensitive to the photon polarization, and $\mathcal{C}_{\phi\gamma}$ is equivalent to the direct CP violation. The obtained results are:

$$\mathcal{A}^{\Delta}_{\phi\gamma} = -0.551\,^{+0.330}_{-0.359}(\text{stat.}) \pm 0.170(\text{syst.}) \pm 0.078(\text{ext.}),$$
$$\mathcal{S}_{\phi\gamma} = 0.441 \pm 0.252(\text{stat.}) \pm 0.064(\text{syst.}) \pm 0.008(\text{ext.}),$$
$$\mathcal{C}_{\phi\gamma} = 0.122 \pm 0.238(\text{stat.}) \pm 0.055(\text{syst.}) \pm 0.013(\text{ext.}),$$

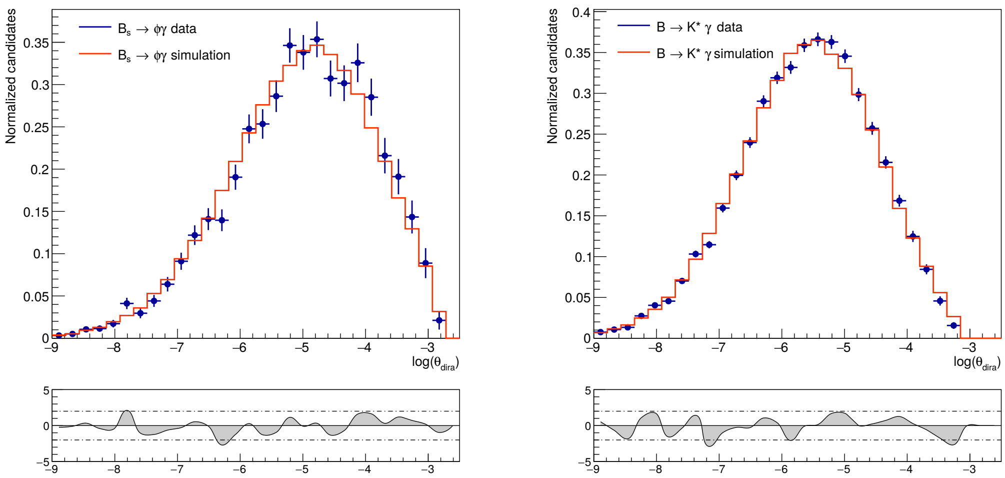
where the uncertainties are statistical, systematic, and from the external measurements, respectively, with the measurement being dominated by the statistical uncertainty. The uncertainty is competitive with the S_{CP} and C_{CP} measurements of the B-factories. The value of $\mathcal{A}^{\Delta}_{\phi\gamma}$ is compatible with the untagged result, and the values of $\mathcal{S}_{\phi\gamma}$ and $\mathcal{C}_{\phi\gamma}$ are compatible with the vanishing Standard Model prediction within 1.7 and 0.5 standard deviations, respectively. The implications of the measurement are analyzed within the effective weak Hamiltonian context, calculating the new physics constraints for the $b \to s\gamma$ transition. The current measurement pushes the experimental value towards the new physics region, while still being compatible with the SM within 1.8σ.

$\mathcal{A}$

Appendix

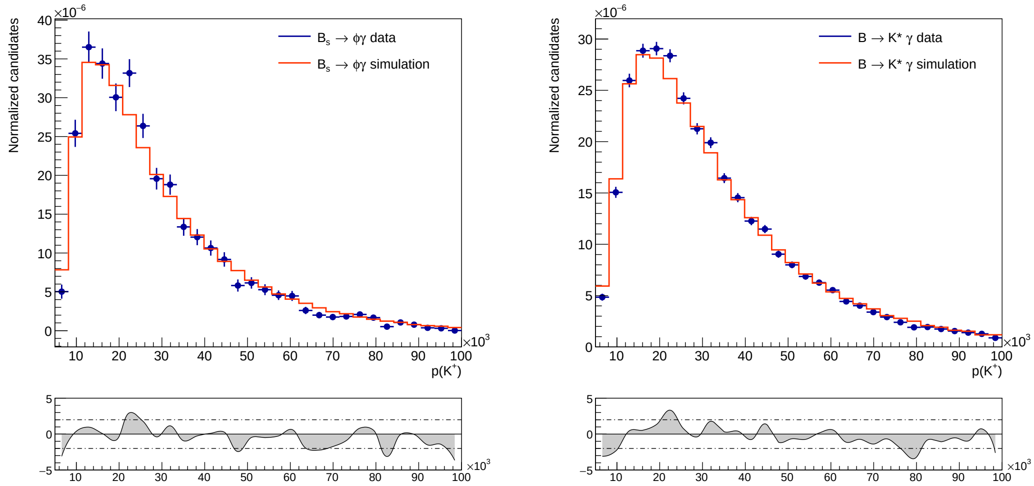
A.1 Data/MC comparison

Comparison between the data and simulation samples, for several variable distributions in $B_s^0 \to \phi\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) decays. The background subtraction procedure of Section 3.4 has been applied to data.

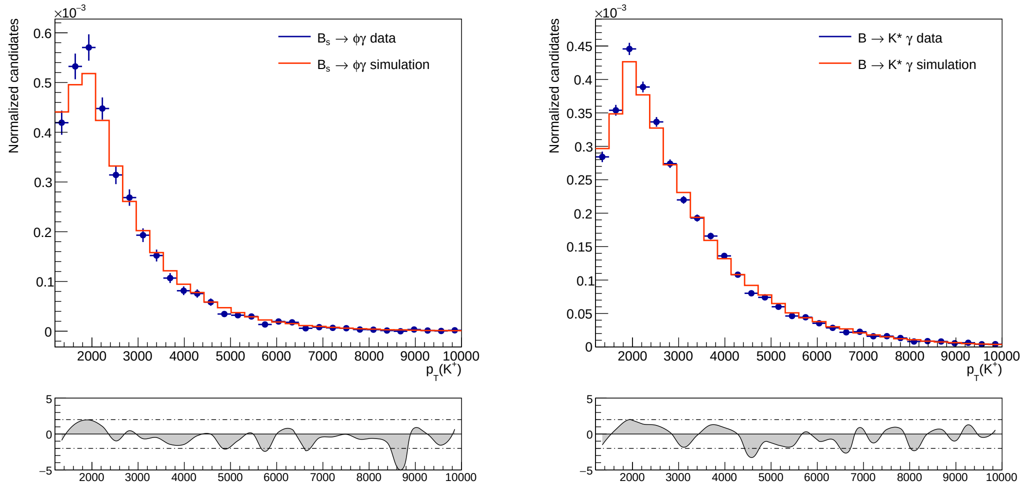


(a) $\log\theta_{\rm dira}$

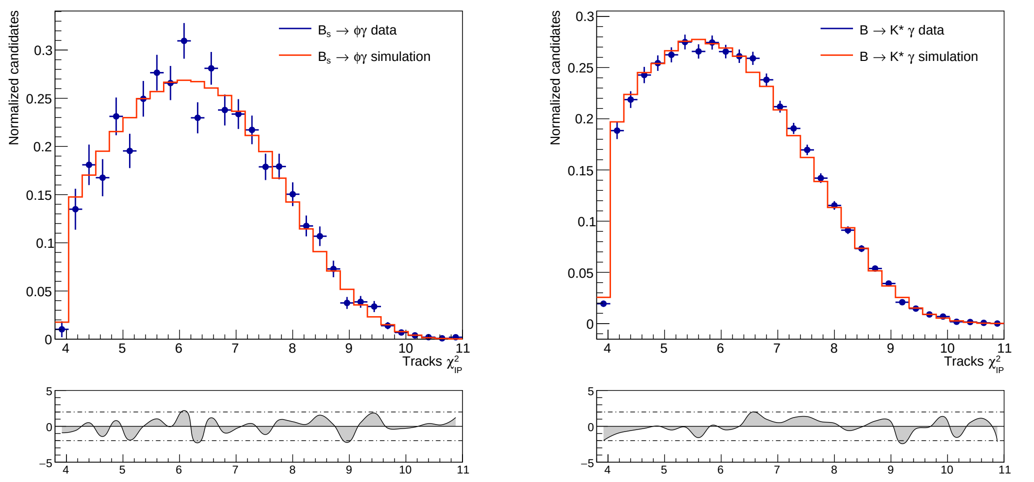
Figure A.1: (I) Comparison between data (blue) and simulation (red), for $B_s^0 \to \phi\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) decays. The pull distribution is shown below.



(a) Momentum h^+

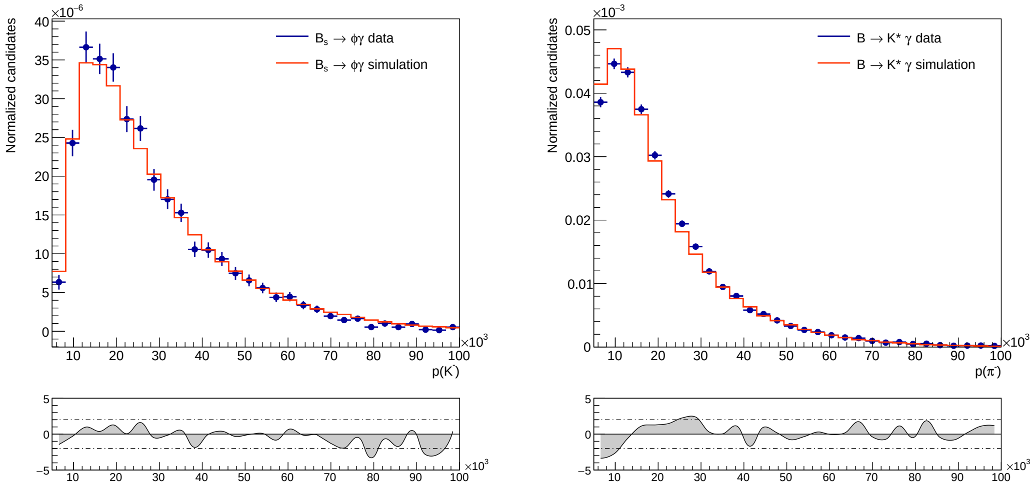


(b) Transverse momentum h^+

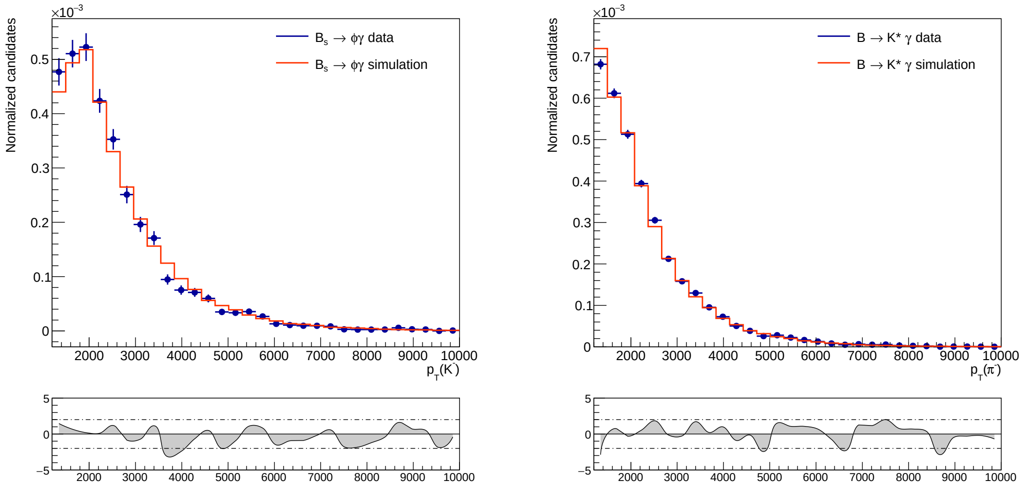


(c) Tracks min $\log \chi^2_{\mathrm{IP}}$

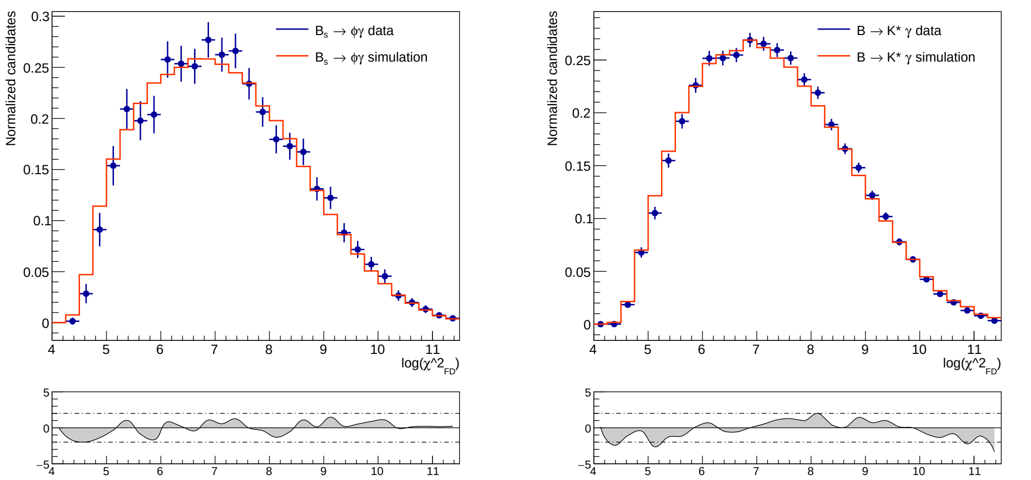
Figure A.2: (II) Comparison between data (blue) and simulation (red), for $B_s^0 \to \phi\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) decays. The pull distribution is shown below.



(a) Momentum h^-

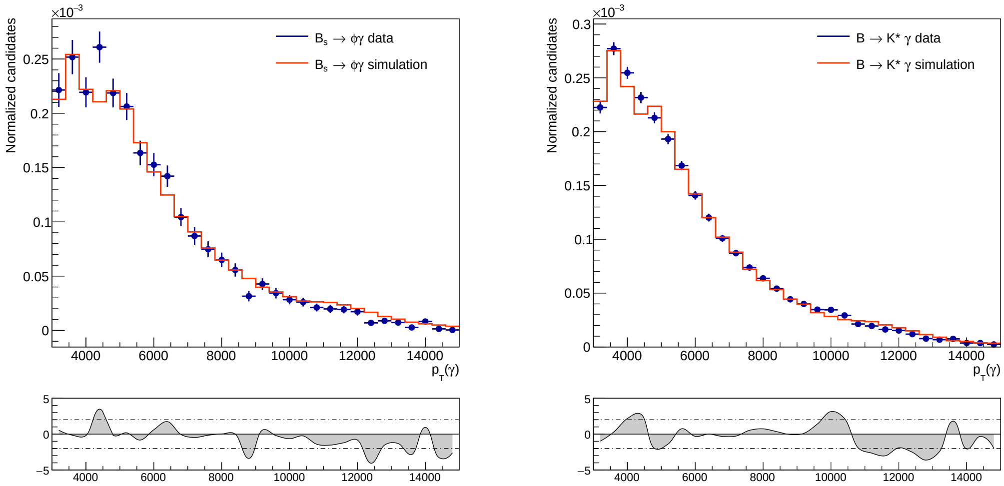


(b) Transverse momentum h^-

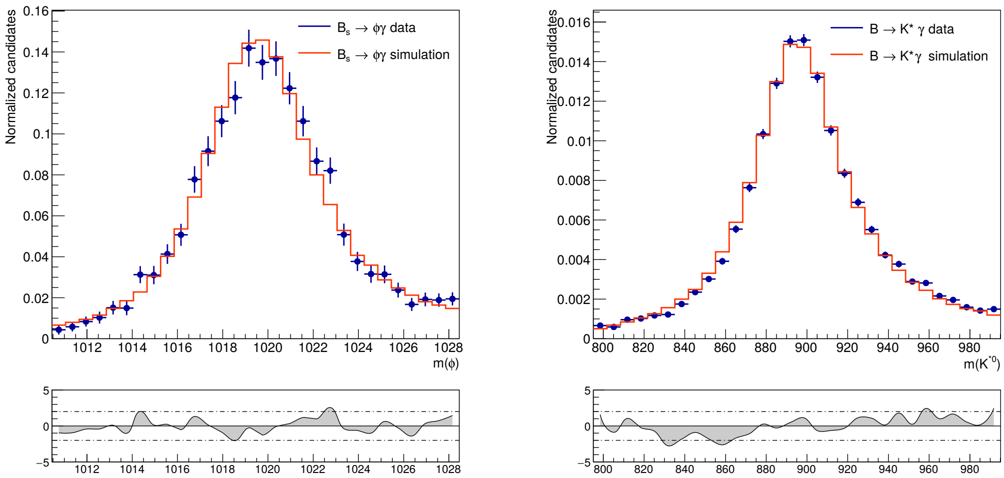


(c) $\log \chi^2_{\mathrm{FD}}$ of the B meson

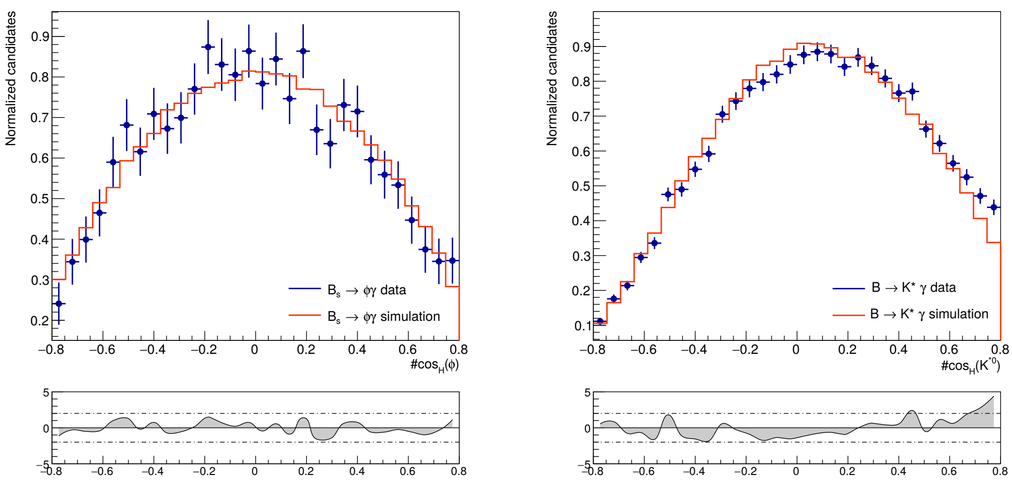
Figure A.3: (III) Comparison between data (blue) and simulation (red), for $B_s^0 \to \phi\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) decays. The pull distribution is shown below.



(a) Photon p_T



(b) ϕ/K^{*0} mass



(c) Helicity angle $\cos\theta_H$

Figure A.4: (IV) Comparison between data (blue) and simulation (red), for $B_s^0 \to \phi\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) decays. The pull distribution is shown below.

Resumen

El Modelo Estándar (SM) de la física de partículas describe las partículas elementales y sus interacciones. A pesar de que ha sido capaz de explicar la mayor parte de los fenómenos físicos con extraordinaria precisión, sabemos que se trata de una teoría incompleta, pues existen fenómenos que no logra incorporar como la gravedad, la materia y energía oscuras, o las oscilaciones de neutrinos. Por ello, buena parte de los experimentos de física de partículas se centran en buscar anomalías respecto a las predicciones del SM, con el objetivo de encontrar las piezas que faltan y construir modelos de nueva física.

La búsqueda de nueva física se puede realizar mediante el estudio de la desintegración $B_s^0 \to \phi\gamma$. Esta desintegración se describe a nivel de quarks por la transición $b \to s\gamma$, que sucede a través de un proceso mediado por la interacción débil a nivel de bucle, como muestra la Figura R.1 (izquierda). Según el SM, el fotón es emitido con helicidad levógira en la gran mayoría de las ocasiones, una magnitud que depende de la dirección del espín y

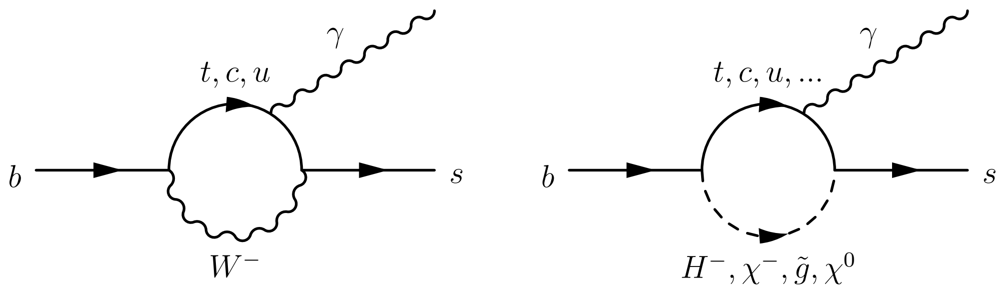


Figure R.1: Diagrama de la transición $b \to s\gamma$, mediado por partículas del Modelo Estándar (izquierda) y por nueva física (derecha).

del momento de la partícula. Sin embargo, existen modelos de nueva física que predicen una fracción mayor de fotones dextrógiros, debido al efecto producido por nuevas partículas en el bucle de la desintegración, ilustrado en la Figura R.1 (derecha). Como no es posible medir la helicidad del fotón directamente, su medida ha de ser llevada a cabo de manera indirecta, por ejemplo a través de la fracción de desintegración dependiente del tiempo en la desintegración $B_s^0 \to \phi\gamma$:

$$\begin{aligned}\Gamma_{B_s^0\to\phi\gamma}(t) \propto e^{-\Gamma_s t}[&\cosh(\Delta\Gamma_s t/2) - \mathcal{A}^\Delta \sinh(\Delta\Gamma_s t/2) \\ &+ \mathcal{C}\cos(\Delta m_s t) - \mathcal{S}\sin(\Delta m_s t)], \\ \Gamma_{\bar{B}_s^0\to\phi\gamma}(t) \propto e^{-\Gamma_s t}[&\cosh(\Delta\Gamma_s t/2) - \mathcal{A}^\Delta \sinh(\Delta\Gamma_s t/2) \\ &- \mathcal{C}\cos(\Delta m_s t) + \mathcal{S}\sin(\Delta m_s t)].\end{aligned} \tag{A.1}$$

Los observables $\mathcal{A}^\Delta$ y $\mathcal{S}$ dependen de la fracción de helicidades del fotón de la siguiente forma:

$$\begin{aligned}\mathcal{A}^\Delta &\simeq \frac{2r}{1+r^2}\cos(\phi_s - \phi_L - \phi_R)\cos(\delta_L - \delta_R), \\ S &\simeq \frac{2r}{1+r^2}\sin(\phi_s - \phi_L - \phi_R)\cos(\delta_L - \delta_R),\end{aligned}$$

donde $r = |A_L|/|A_R|$ es la fracción de helicidades, ϕ_s es el ángulo de mezcla débil del mesón B_s^0, y δ es la fase fuerte. Por otra parte, el observable $\mathcal{C}$ representa la violación de CP directa. Estos observables aparecen debido a la interferencia entre las oscilaciones partícula-antipartícula del mesón neutro B_s^0 y la desintegración al estado final $\phi\gamma$, el cual es accesible a través de ambos estados iniciales B_s^0 y $\bar{B}_s^0$ al ser un autoestado de CP. La predicción del SM para estos observables es muy cercana a cero, por tanto cualquier desviación significativa de este valor sería un indicio de nueva física.

Para realizar esta medida se han utilizado los datos registrados por LHCb, un experimento diseñado para llevar a cabo medidas de alta precisión relacionadas con la física del sabor, como medidas de la violación de la simetría CP o procesos que involucren quarks pesados como el quark belleza b. LHCb es uno de los cuatro experimentos principales del Gran Colisionador de Hadrones (LHC), operado por la Organización Europea para la Investigación Nuclear (CERN) en la frontera franco-suiza, donde se aceleran protones a una energía nominal de centro de masas de 14 TeV. La Figura R.2 muestra un esquema de perfil del detector LHCb, el cual está formado por diferentes subsistemas:

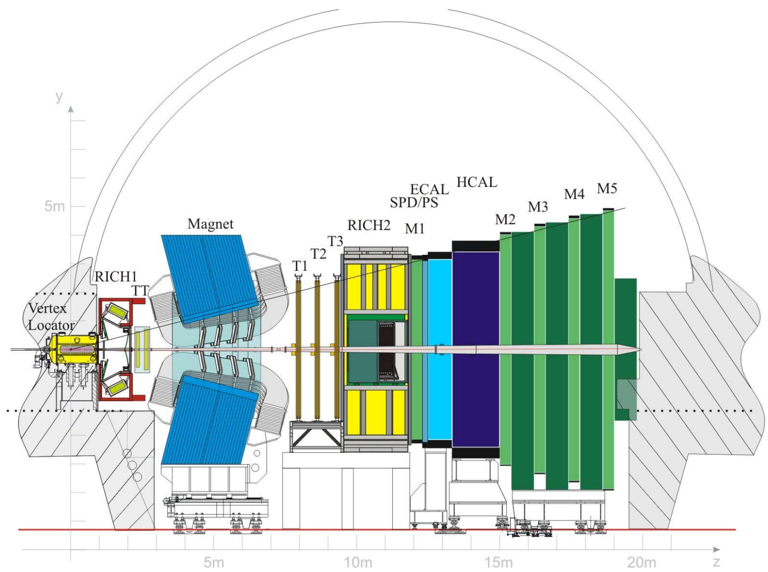


Figure R.2: Vista de perfil del detector LHCb, el cual mide 21 metros de largo, 10 metros de alto, y pesa 5 600 toneladas.

- El sistema de *tracking* sirve para rastrear las partículas que se producen en la colisión. Está formado por el Vertex Locator (VeLo), que es el detector más interno situado cerca de donde se producen las colisiones, y las estaciones de rastreo antes (TT) y después (T1-T3) del imán. La desviación que el imán induce sobre las partículas permite determinar sus cargas y momentos. La mayor parte del sistema de tracking consiste en detectores de tiras de silicio.

- Los detectores RICH1 y RICH2 son detectores de radiación Cherenkov, los cuales permiten identificar las partículas cargadas que lo atraviesan a través de la firma que dejan a su paso.

- El sistema calorimétrico está formado por los calorímetros electromagnético (ECAL) y hadrónico (HCAL), los cuales permiten identificar y medir la energía de las partículas electromagnéticas (como fotones y electrones) y de los hadrones (como piones y kaones), respectivamente. El ECAL se encuentra precedido de una placa de plomo y de dos subdetectores (SPD y PS) que ayudan a la identificación de las partículas.

- Las cámaras de muones (M1-M5) rastrean y miden los momentos de las partículas homónimas. Éstas se sitúan al final del detector, ya que los muones interaccionan con la materia en menor medida que otras partículas.

La identificación y reconstrucción de los fotones se lleva a cabo gracias al calorímetro electromagnético del LHCb. Se utilizan diversas técnicas para calibrar la respuesta de las celdas calorimétricas, así como varios algoritmos que se encargan de identificar y distinguir entre las diferentes partículas neutras, como los fotones γ y los piones π^0. Para conseguir esto último, los algoritmos se basan en el llamado perfil energético de los fotones, un modelo que predice la respuesta media de las celdas ante fotones individuales. Este trabajo presenta una mejora de este perfil, mostrado en la Figura R.3. Los perfiles se han extraido de los datos de simulaciones Monte Carlo donde se simula la respuesta del detector LHCb, utilizándose varias muestras de simulaciones con energías iniciales del fotón de entre 1 y 168 GeV. Se han obtenido seis perfiles diferentes, separándolos para las diferentes regiones del calorímetro (interna, media y externa), y para fotones estándar y convertidos (se dice que un fotón es convertido si se ha transformado en un par electrón-positrón antes de llegar al calorímetro). Estos perfiles se utilizan actualmente en el software del LHCb.

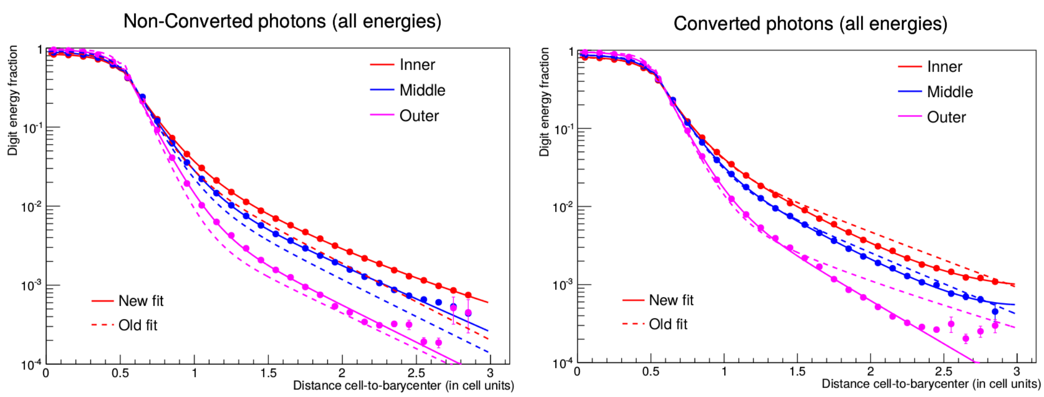


Figure R.3: Perfil energético de fotones en escala logarítmica, para fotones estándar (izquierda) y convertidos (derecha). El nuevo ajuste, representado por la línea sólida, reemplaza al ajuste anterior, representado por la línea discontinua.

En esta tesis se presenta la medida de los observables $\mathcal{A}^{\Delta}_{\phi\gamma}$, $\mathcal{S}_{\phi\gamma}$ y $\mathcal{C}_{\phi\gamma}$ en desintegraciones $B^0_s \to \phi\gamma$, analizando la fracción de desintegración dependiente del tiempo. La medida se ha realizado utilizando la muestra de datos de 3 fb^{-1} registrados por LHCb durante los años 2011 y 2012, con energías en centro de masas de $\sqrt{s} = 7$ TeV y $\sqrt{s} = 8$ TeV, respectivamente. Para lograr la medida de $\mathcal{A}^{\Delta}_{\phi\gamma}$ no es necesario determinar el sabor del mesón inicial (medida sin etiquetado de sabor). Sin embargo, para la medida de $\mathcal{S}_{\phi\gamma}$ y $\mathcal{C}_{\phi\gamma}$ es necesario el etiquetado de sabor, ya que los términos que dependen de estos parámetros en la fracción de desintegración (Ecuación A.1) poseen diferente signo para B^0_s y $\bar{B}^0_s$, y de otro modo cancelarían en la suma. En este trabajo se han realizado los dos tipos de análisis, separándolos en sendos capítulos al requerir diferentes estrategias. Los análisis se han desarrollado en todo momento a ciegas, sin conocer el valor de los observables, desvelando éstos únicamente antes de la publicación de la presente tesis.

La reconstrucción y selección de candidatos se ha realizado de manera común para los dos análisis. Este proceso cuenta con dos partes: (1) una selección eficiente de sucesos, que mantenga la señal al mismo tiempo que deseche sucesos de fondo, y (2) un método de sustracción de fondo, de manera que solo los sucesos señal se utilicen para la medida. La desintegración $B^0 \to K^{*0}(\to K^{\pm}\pi^{\mp})\gamma$ es utilizada para controlar los efectos de reconstrucción en el tiempo propio que pudieran afectar a la desintegración de señal $B^0_s \to \phi(\to K^+K^-)\gamma$. La Figura R.4 muestra un esquema de ambas desintegraciones, consistentes en un mesón inicial que recorre de media varios centímetros, un mesón intermedio que se desintegra mediante interacción fuerte y por tanto no vuela, y dos hadrones y un fotón en el estado final. La selección se basa en encontrar dos hadrones que formen un vértice desplazado de la colisión inicial, junto con un fotón de alta energía, y consta de tres etapas diferenciadas: *trigger*, *stripping* y *offline*. La etapa de trigger se lleva a cabo al mismo tiempo que se recogen los datos del detector LHCb, con el objetivo de seleccionar los sucesos potencialmente interesantes. La etapa de stripping es una reconstrucción que pretende descartar aún más sucesos, con el fin de disponer de una cantidad de datos manejable en cualquier análisis. Finalmente, la selección offline la desarrolla el analista para adecuar la muestra a los objetivos del análisis.

En este trabajo se han escogido las líneas de trigger y stripping, y se ha optimizado una selección offline basada en cortes, maximizando la significancia de señal $\mathcal{S} = S/\sqrt{S+B}$, donde S y B representan el número de sucesos de señal y de fondo, respectivamente. Para ello, se han utilizado las variables que más discriminan entre señal y fondo:

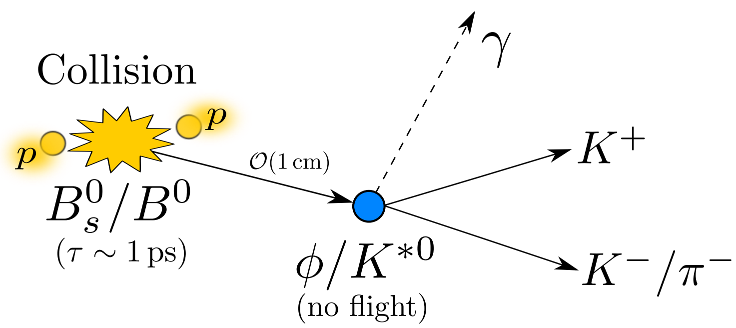


Figure R.4: Topología de las desintegraciones $B_s^0 \to \phi\gamma$ y $B^0 \to K^{*0}\gamma$.

- χ^2_{IP}. El parámetro de impacto, IP, es la distancia mínima entre la trayectoria de una partícula y el vértice de colisión. Un valor elevado garantiza que las partículas se han producido en un vértice desplazado como consecuencia de una desintegración. La variable $\chi^2_{\mathrm{IP}} = \mathrm{IP}^2/\sigma^2_{\mathrm{IP}}$ tiene en cuenta la incertidumbre, y por tanto es una variable discriminatoria más adecuada, siendo la principal en este análisis.

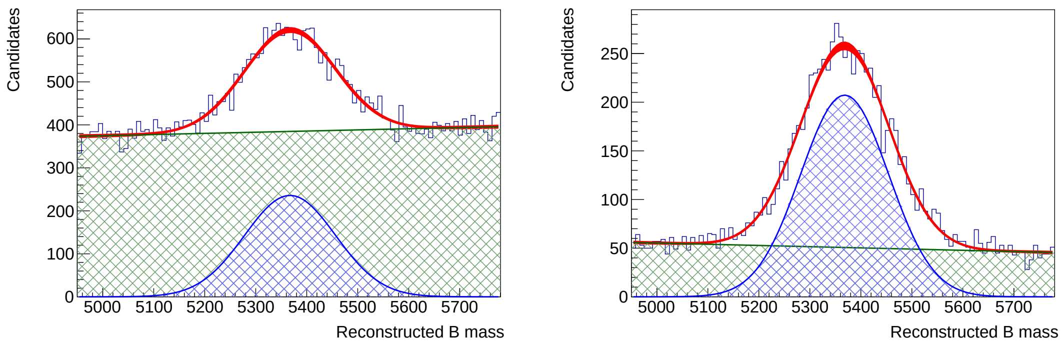
- γ/π^0. Es la variable que permite distinguir entre fotones y piones neutros, producida a partir de múltiples variables calorimétricas.

- $\cos\theta_H$. El ángulo de helicidad del mesón intermedio. Se trata de una variable útil puesto que la señal y los diferentes tipos de fondo producen distribuciones diferentes en esta variable.

- $p_T(B)$. Para finalizar, se realiza un corte algo más severo en el momento transversal del mesón B_s^0, una variable reconstruida a partir de los momentos de sus productos de desintegración.

La Figure R.5 muestra la distribución de masa del mesón inicial, $m(B_s^0)$, antes (izquierda) y después (derecha) de aplicar la selección offline. Una vez seleccionados los candidatos, se procede a sustraer el fondo por completo mediante la técnica sPlot. La masa reconstruida del mesón B permite distinguir entre las diferentes componentes (señal, fondo combinatorial, y fondos físicos), cuya distribución se ha extraído parcialmente de simulaciones Monte Carlo, y se asigna un peso (sWeight) a cada candidato. En total, se han obtenido $5\,266 \pm 104$ sucesos señal reconstruidos para $B_s^0 \to \phi\gamma$, y aproximadamente seis veces dicha cifra para el canal de control $B^0 \to K^{*0}\gamma$.

El primer análisis presentado en esta tesis corresponde a la medida del



(a) Antes de la selección offline (b) Después de la selección offline

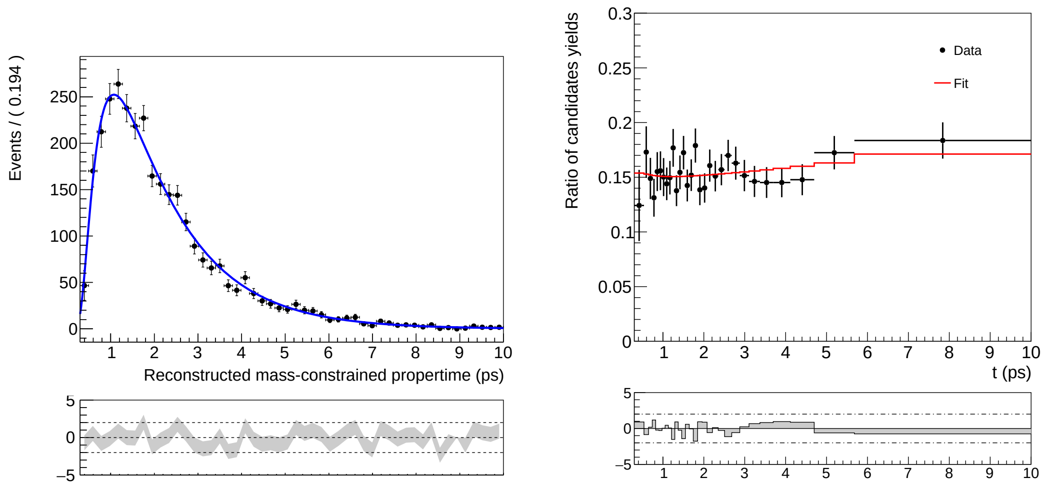
Figure R.5: Distribución de la masa reconstruida del mesón inicial $m(B_s^0)$ en la desintegración $B_s^0 \to \phi\gamma$, antes y después de aplicar la selección offline.

observable $\mathcal{A}^{\Delta}$, mediante un análisis sin etiquetado de sabor. En esta medida resulta de capital importancia la determinación de la aceptancia del tiempo de desintegración, una función que describe las ineficiencias en la distribución del tiempo propio producidas por la reconstrucción y la selección de sucesos. La aceptancia es obtenida de datos, utilizando el canal de control $B^0 \to K^{*0}\gamma$, mientras que las diferencias en aceptancia entre los canales de señal y control son obtenidas de simulaciones Monte Carlo donde se simula la respuesta del detector LHCb. Se han desarrollado dos estrategias de ajuste diferentes. La primera, un ajuste *unbinned* de máxima verosimilitud, donde los canales $B_s^0 \to \phi\gamma$ y $B^0 \to K^{*0}\gamma$ son ajustados simultáneamente. La segunda, un ajuste a la fracción de candidatos de $B_s^0 \to \phi\gamma$ sobre $B^0 \to K^{*0}\gamma$. Para ésta última, el autor de esta tesis ha desarrollado un código [124] que permite ajustes genéricos donde los datos son agrupados (*binned data*), agrupando los candidatos en bins de igual tamaño, e integrando la función densidad de probabilidad (PDF) en cada bin, lo cual resulta necesario en distribuciones donde existan regiones con poca población de sucesos. Ambas estrategias de análisis han sido verificadas con la generación de pseudoexperimentos mediante técnicas de simulación Monte Carlo.

Utilizando la estrategia unbinned, el resultado obtenido ha sido:

$$\mathcal{A}^{\Delta}_{\phi\gamma} = -0.555\,^{+0.331}_{-0.359}(\text{stat.}) \pm 0.170(\text{syst.}) \pm 0.078(\text{ext.}),$$

donde la primera incertidumbre es estadística, la segunda es sistemática, y la tercera corresponde a medidas externas. Ésta última hace referencia a los parámetros que han sido tomados de medidas de otros análisis. Los sistemáticos principales provienen de la sustracción del fondo, los cuales podrían reducirse en futuros análisis, aunque actualmente la medida está dominada por la incertidumbre estadística. El ajuste se representa en la Figura R.6 Utilizando la segunda estrategia de análisis, el resultado obtenido ha sido $\mathcal{A}^{\Delta} = -0.473\,^{+0.321}_{-0.343}$, donde el ajuste se encuentra representado en la Figura R.6 (derecha). Los valores obtenidos por ambos métodos son compatibles entre sí, lo cual se ha contrastado mediante pseudoexperimentos. Este trabajo ha dado lugar a la primera medida de este observable, $\mathcal{A}^{\Delta}_{\phi\gamma} = -0.98^{+0.46}_{-0.52}\,^{+0.23}_{-0.20}$, publicado en [59], tratándose de la primera medida de un observable sensible a la fracción de helicidades del fotón en una desintegración B^0_s. En esta tesis se ha mejorado la reconstrucción y selección de candidatos con respecto a la referencia mencionada, incrementando en un 20% el número de sucesos señal.



(a) Ajuste unbinned

(b) Ajuste a la fracción de candidatos

Figure R.6: Ajustes al tiempo de desintegración de $B^0_s \to \phi\gamma$ mediante (a) el método unbinned y (b) el método de la fracción de candidatos $B^0_s \to \phi\gamma / B^0 \to K^{*0}\gamma$.

El segundo análisis de esta tesis corresponde a la medida de los tres observables $\mathcal{A}^{\Delta}$, $\mathcal{S}$ y $\mathcal{C}$, utilizando técnicas de etiquetado de sabor del mesón inicial. Para ello, se ha utilizado la información proporcionada por los algoritmos de etiquetado de sabor de LHCb, que pueden clasificarse en dos grupos:

algoritmos OS (*other-side*) y algoritmos SS (*same-side*). Los algoritmos OS aprovechan el hecho de que tras la colisión pp los quarks b se producen en pares, y por tanto intentan clasificar el sabor del otro hadrón del suceso. Los algoritmos SS se basan en la búsqueda del pión (kaón) que se produce junto a la hadronización del B_s^0 (B^0). En este análisis se ha calibrado la salida de los algoritmos para su correcta utilización en muestras $B_s^0 \to \phi\gamma$, y se han combinado en una única decisión que es la que finalmente se ha usado para el ajuste a los datos. La Tabla R.1 muestra el rendimiento de cada algoritmo y de la combinación, estimándose una eficiencia de etiquetado efectiva de $\epsilon_{\rm eff} = 4.58 \pm 0.13\%$.

Table R.1: Rendimiento de los algoritmos de etiquetado en $B_s^0 \to \phi\gamma$, para el período 2011+2012. $\epsilon_{\rm tag}$ es la eficiencia de etiquetado, ω la fracción de etiquetados incorrectos, y $\epsilon_{\rm power} = \epsilon_{\rm tag}\,(1-2\omega)^2$ la eficiencia efectiva.

Algoritmo de etiquetado	$\epsilon_{\rm tag}$ (%)	$\omega_{\rm eff}$ (%)	$\epsilon_{\rm power}^{\omega_{\rm eff}}$ (%)
OSComb calibrado	33.03	35.40	2.82
SSKaonNNet calibrado	63.89	41.27	1.95
Combinación	74.53	37.61	4.58

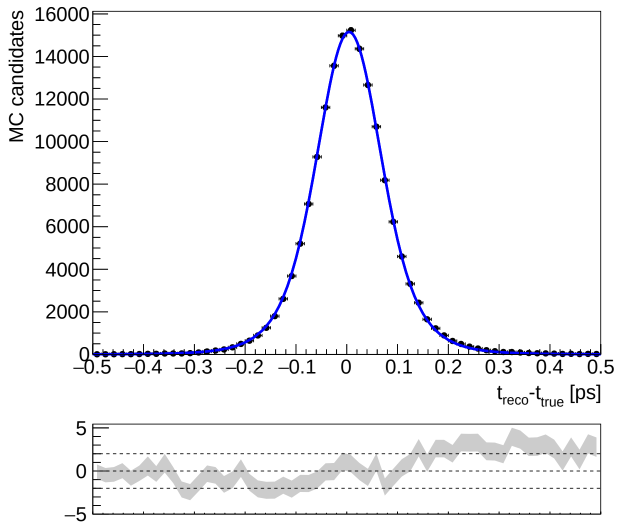


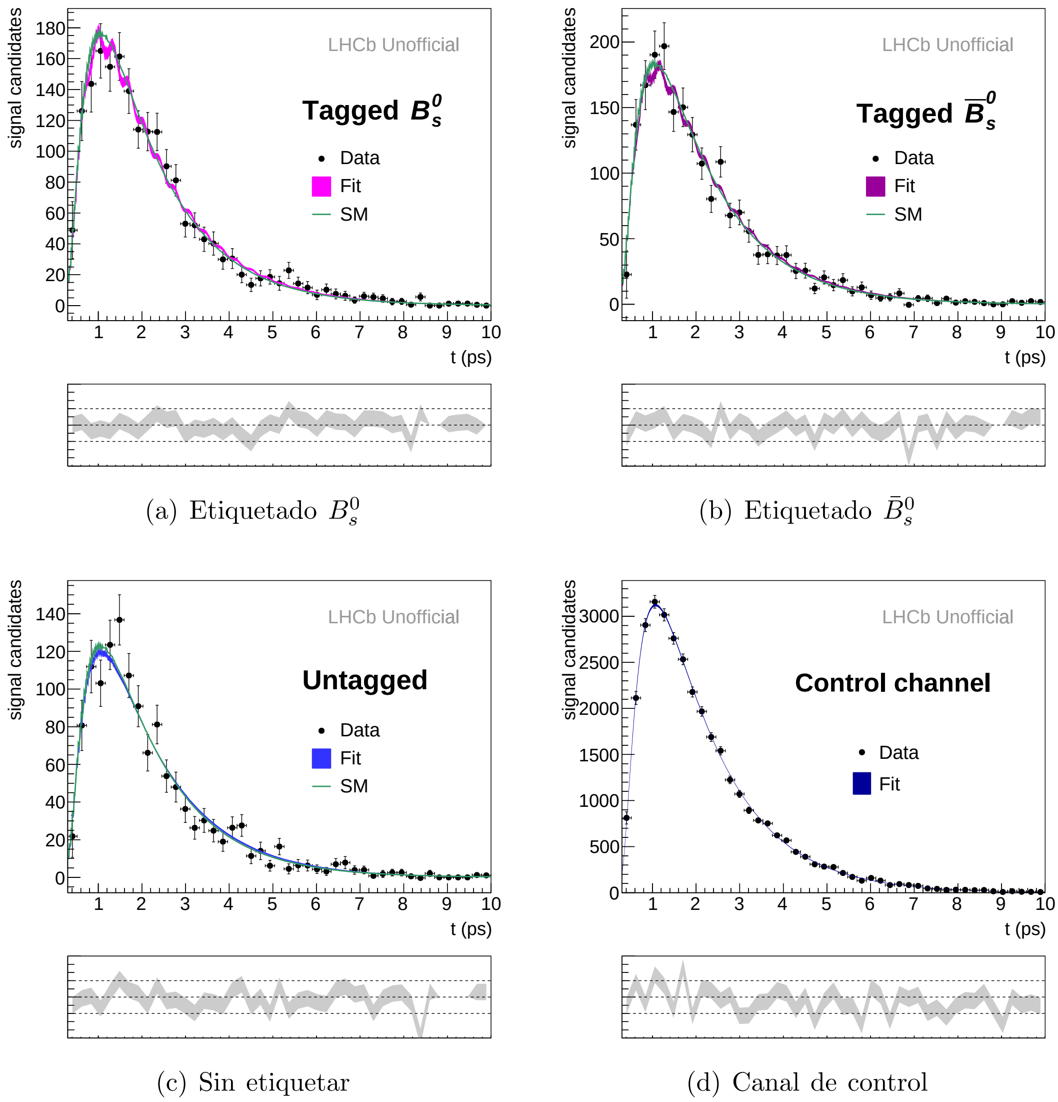
Figure R.7: Ajuste a la resolución del tiempo de desintegración en sucesos de $B_s^0 \to \phi\gamma$ simulados, con la simulación completa del detector LHCb.

Otro de los aspectos fundamentales para el éxito de la medida es la caracterización de la resolución del tiempo de desintegración. Su importancia radica en el hecho de que los términos que acompañan a los coeficientes $\mathcal{S}$ y $\mathcal{C}$ en la fracción de desintegración (Ecuación A.1) oscilan con una frecuencia $\Delta m_s = 17.757 \pm 0.021$ ps^{-1} [39], cuyo período es tan solo cuatro veces mayor que la resolución experimental en desintegraciones radiativas. La resolución se ha caracterizado haciendo uso de la incertidumbre del tiempo de desintegración proporcionada por la muestra de datos, la cual debe ser escalada por un factor de corrección que se extrae de las muestras de simulación de LHCb, donde se tiene acceso a la resolución verdadera puesto que se conoce el tiempo verdadero de desintegración. La Figura R.7 muestra el ajuste a la resolución de Monte Carlo, cuya función densidad de probabilidad es la suma de dos gausianas, cada una con un factor de corrección diferente. El mejor ajuste consiste en una proporción de la primera gausiana de $97.8 \pm 0.2\%$, con factores de corrección $s_1 = 1.144 \pm 0.003$ y $s_2 = 2.79 \pm 0.08$, y una media o sesgo de $\mu = 4.5 \pm 0.2$ fs.

La estrategia de ajuste se ha desarrollado con un método unbinned de máxima verosimilitud, donde se introduce el etiquetado de sabor y la descripción de la resolución del tiempo de desintegración. La convolución de la fracción de desintegración con la resolución se calcula analíticamente, integrando el código desarrollado en las referencias [142, 144], y utilizando una función densidad de probabilidad que engloba tanto la aceptancia como la resolución. Finalmente, la estrategia de análisis ha sido verificada mediante la generación de pseudoexperimentos simulados, obteniendo resultados consistentes. El resultado la medida es el siguiente:

$$\mathcal{A}^{\Delta}_{\phi\gamma} = -0.551\,^{+0.330}_{-0.359}(\text{stat.}) \pm 0.170(\text{syst.}) \pm 0.067(\text{ext.}),$$
$$\mathcal{S}_{\phi\gamma} = 0.441 \pm 0.252(\text{stat.}) \pm 0.064(\text{syst.}) \pm 0.008(\text{ext.}),$$
$$\mathcal{C}_{\phi\gamma} = 0.122 \pm 0.238(\text{stat.}) \pm 0.055(\text{syst.}) \pm 0.013(\text{ext.}),$$

donde la primera incertidumbre es estadística, la segunda es sistemática, y la tercera corresponde a medidas externas. En los tres casos, la incertidumbre se encuentra dominada por la contribución estadística. Las gráficas del ajuste se muestran en la Figura R.8, junto con la curva esperada por el Modelo Estándar. Los resultados son compatibles con la predicción del Modelo Estándar, desviándose de la misma en 1.5, 1.7 y 0.5 desviaciones estándar,



(a) Etiquetado B_s^0

(b) Etiquetado $\bar{B}_s^0$

(c) Sin etiquetar

(d) Canal de control

Figure R.8: Resultado del ajuste con etiquetado de sabor. Las figuras superiores indican las muestras etiquetadas como B_s^0 y $\bar{B}_s^0$, respectivamente, mientras que las figuras inferiores indican la muestra sin etiquetar de $B_s^0 \to \phi\gamma$ y la muestra del canal de control $B^0 \to K^{*0}\gamma$. Todas ellas se han ajustado simultáneamente.

respectivamente. Este trabajo corresponde a la primera medida de los observables $\mathcal{S}_{\phi\gamma}$ y $\mathcal{C}_{\phi\gamma}$, y a la primera aplicación de los algoritmos de etiquetado de sabor en un análisis radiativo en LHCb.

Las implicaciones de los resultados han sido analizadas desde una perspectiva fenomenológica, en el contexto del hamiltoniano débil efectivo. Los parámetros $\mathcal{A}^{\Delta}$ y $\mathcal{S}$ se pueden expresar en función del coeficiente de Wilson C_7' que acompaña al operador electromagnético $\mathcal{O}_7'$, de la siguiente forma:

$$\mathcal{A}^{\Delta}_{\phi\gamma} \simeq \frac{2\,\mathrm{Re}\left(e^{-i\phi_s}\,C_7C_7'\right)}{|C_7|^2+|C_7'|^2},$$
$$\mathcal{S}_{\phi\gamma} \simeq \frac{2\,\mathrm{Im}\left(e^{-i\phi_s}\,C_7C_7'\right)}{|C_7|^2+|C_7'|^2},$$

siendo sensibles a la parte real e imaginaria, respectivamente. Las restricciones que impone la medida en el coeficiente de Wilson C_7' se muestran en la Figura R.9 (izquierda), dando como resultado una compatibilidad con el Modelo Estándar de en torno a 1.8σ. En la Figura R.9 (derecha) se muestra el mismo gráfico incluyendo el resto de las medidas radiativas sensibles a la transición $b \to s\gamma$, donde la banda de restricción es más compatible con el SM, situándose en torno a 0.8σ.

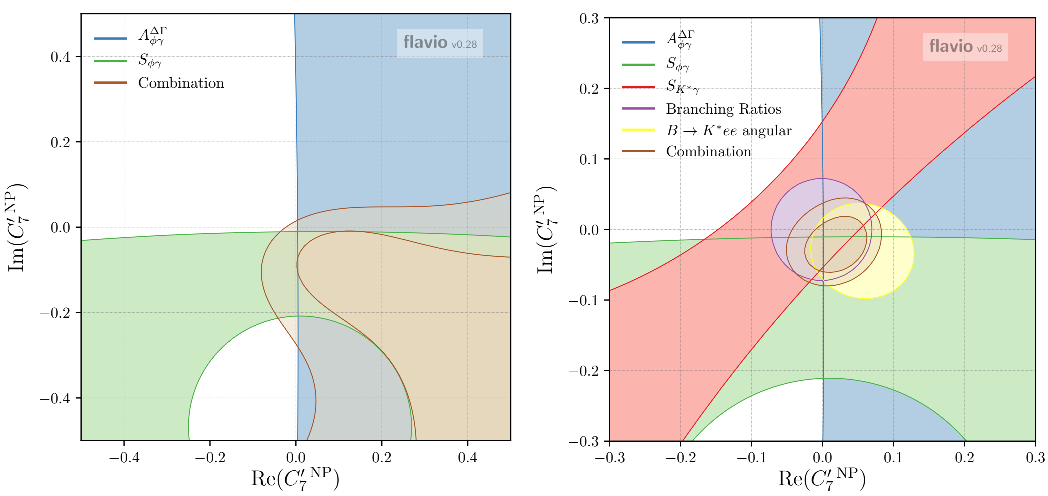


Figure R.9: Restricciones en el plano complejo $C_7'^{\,\mathrm{NP}}$, producidas por la medida de $\mathcal{A}^{\Delta}_{\phi\gamma}$ y $\mathcal{S}_{\phi\gamma}$ (izquierda) y por todos los observables sensibles a la transición $b \to s\gamma$ (derecha).

En conclusión, este trabajo presenta la primera medida de los observables $\mathcal{A}^{\Delta}_{\phi\gamma}$, $\mathcal{S}_{\phi\gamma}$ y $\mathcal{C}_{\phi\gamma}$ en desintegraciones $B^0_s \to \phi\gamma$. Para ello, se ha analizado la fracción de desintegración dependiente del tiempo de dicho canal, utilizando los datos recogidos por el detector LHCb durante los años 2011 y 2012. Estos observables, inducidos por las oscilaciones del mesón neutro B_s^0, podrían presentar efectos de nueva física. El trabajo de esta tesis ha dado lugar a la primera medida de $\mathcal{A}^{\Delta}_{\phi\gamma}$ [59], tratándose del primer observable sensible a la polarización del fotón en desintegraciones B_s^0, mientras que la medida de $\mathcal{S}_{\phi\gamma}$ y $\mathcal{C}_{\phi\gamma}$ se encuentra pendiente de publicación. Los valores son compatibles con la predicción del Modelo Estándar en torno a 1.5, 1.7 y 0.5 desviaciones estándar para $\mathcal{A}^{\Delta}_{\phi\gamma}$, $\mathcal{S}_{\phi\gamma}$ y $\mathcal{C}_{\phi\gamma}$, respectivamente.

Bibliography

[1] Wikimedia Commons. *File:Standard Model of Elementary Particles.svg*. 2006. URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (Cited on page 2).

[2] ATLAS Collaboration, G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett.* B716 (2012), pp. 1–29. arXiv: 1207.7214 [hep-ex] (Cited on page 2).

[3] CMS Collaboration, S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* B716 (2012), pp. 30–61. arXiv: 1207.7235 [hep-ex] (Cited on page 2).

[4] A. Esposito et al. "Four-Quark Hadrons: an Updated Review", *Int. J. Mod. Phys.* A30 (2015), p. 1530002. arXiv: 1411.5997 [hep-ph] (Cited on page 3).

[5] R. F. Lebed, R. E. Mitchell, and E. S. Swanson. "Heavy-quark QCD exotica", *Progress in Particle and Nuclear Physics* 93.Supplement C (2017), pp. 143–194. DOI: 10.1016/j.ppnp.2016.11.003 (Cited on page 3).

[6] Belle Collaboration, S. K. Choi et al. "Observation of a narrow charmonium-like state in exclusive $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}J/\psi$ decays", *Phys. Rev. Lett.* 91 (2003), p. 262001. arXiv: hep-ex/0309032 [hep-ex] (Cited on page 3).

[7] LHCb Collaboration, R. Aaij et al. “Observation of the resonant character of the $Z(4430)^-$ state”, *Phys. Rev. Lett.* 112.22 (2014), p. 222002. arXiv: 1404.1903 [hep-ex] (Cited on pages 3, 26).

[8] LHCb Collaboration, R. Aaij et al. “Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays”, *Phys. Rev. Lett.* 115 (2015), p. 072001. arXiv: 1507.03414 [hep-ex] (Cited on pages 3, 26).

[9] S. L. Glashow. “Partial Symmetries of Weak Interactions”, *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2 (Cited on page 3).

[10] A. Salam and J. C. Ward. “Electromagnetic and weak interactions”, *Phys. Lett.* 13 (1964), pp. 168–171. DOI: 10.1016/0031-9163(64)90711-5 (Cited on page 3).

[11] S. Weinberg. “A Model of Leptons”, *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264 (Cited on page 3).

[12] S. L. Glashow, J. Iliopoulos, and L. Maiani. “Weak Interactions with Lepton-Hadron Symmetry”, *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285 (Cited on page 3).

[13] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* 13 (1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508 (Cited on page 4).

[14] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* 13 (1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321 (Cited on page 4).

[15] N. Cabibbo. “Unitary Symmetry and Leptonic Decays”, *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531 (Cited on page 4).

[16] M. Kobayashi and T. Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”, *Progress of Theoretical Physics* 49.2 (1973), p. 652. DOI: 10.1143/PTP.49.652 (Cited on page 4).

[17] B. Pontecorvo. “Inverse beta processes and nonconservation of lepton charge”, *Sov. Phys. JETP* 7 (1958). [Zh. Eksp. Teor. Fiz. 34, 247 (1957)], pp. 172–173 (Cited on page 5).

[18] Z. Maki, M. Nakagawa, and S. Sakata. "Remarks on the unified model of elementary particles", *Prog. Theor. Phys.* 28 (1962), pp. 870–880. DOI: 10.1143/PTP.28.870 (Cited on page 5).

[19] L. Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix", *Phys. Rev. Lett.* 51 (21 Nov. 1983), pp. 1945–1947. DOI: 10.1103/PhysRevLett.51.1945 (Cited on page 5).

[20] Particle Data Group Collaboration, C. Patrignani et al. "Review of Particle Physics", *Chin. Phys.* C40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001 (Cited on pages 5, 8, 12, 14, 47, 61, 73, 161).

[21] CKMfitter Group. *Global CKM fit as of Summer 2016*. URL: http://ckmfitter.in2p3.fr/www/html/ckm_results.html (Cited on page 6).

[22] K. Lande et al. "Observation of long-lived neutral V particles", *Phys. Rev.* 103 (1956), pp. 1901–1904. DOI: 10.1103/PhysRev.103.1901 (Cited on page 7).

[23] F. Muller et al. "Regeneration and mass difference of neutral K mesons", *Phys. Rev. Lett.* 4 (1960), pp. 418–421. DOI: 10.1103/PhysRevLett.4.418 (Cited on page 7).

[24] ARGUS Collaboration. "Observation of $B^0 - \bar{B}^0$ mixing", *Physics Letters B* 192.1 (1987), pp. 245–252. DOI: 10.1016/0370-2693(87)91177-4 (Cited on page 7).

[25] UA1 Collaboration. "Search for $B^0 - \bar{B}^0$ oscillations at the CERN proton-antiproton collider", *Physics Letters B* 186.2 (1987), pp. 247–254. DOI: 10.1016/0370-2693(87)90288-7 (Cited on page 7).

[26] CDF Collaboration. "Observation of $B_s^0 - \bar{B}_s^0$ oscillations", *Phys. Rev. Lett.* 97 (24 Dec. 2006), p. 242003. DOI: 10.1103/PhysRevLett.97.242003 (Cited on page 8).

[27] BaBar Collaboration, B. Aubert et al. "Evidence for $D^0 - \overline{D}^0$ mixing", *Phys. Rev. Lett.* 98 (2007), p. 211802. arXiv: hep-ex/0703020 [hep-ex] (Cited on page 8).

[28] Belle Collaboration, M. Staric et al. "Evidence for D^0 - $\bar{D}^0$ mixing", *Phys. Rev. Lett.* 98 (2007). [,65(2007)], p. 211803. arXiv: hep-ex/0703036 [hep-ex] (Cited on page 8).

[29] LHCb Collaboration, R. Aaij et al. “Observation of $D^0 - \overline{D}^0$ oscillations”, *Phys. Rev. Lett.* 110.10 (2013), p. 101802. arXiv: 1211.1230 [hep-ex] (Cited on page 8).

[30] M. S. Sozzi. “Discrete symmetries and CP violation: from experiment to theory”, Oxford Graduate Texts. New York, NY: Oxford Univ. Press, 2008. URL: https://cds.cern.ch/record/1087897 (Cited on pages 9, 11).

[31] G. C. Branco, L. Lavoura, and J. P. Silva. “CP violation”, Internat. Ser. Mono. Phys. Oxford: Clarendon Press, 1999. URL: https://cds.cern.ch/record/396964 (Cited on page 9).

[32] E. Noether. “Invariant Variation Problems”, *Gott. Nachr.* 1918 (1918). [Transp. Theory Statist. Phys.1,186(1971)], pp. 235–257. arXiv: physics/0503066 [physics] (Cited on page 11).

[33] O. W. Greenberg. “Why is CPT fundamental?”, *Found. Phys.* 36 (2006), pp. 1535–1553. arXiv: hep-ph/0309309 [hep-ph] (Cited on page 12).

[34] M. E. Peskin and D. V. Schroeder. “An introduction to quantum field theory”, Boulder, CO: Westview, 1995. URL: https://cds.cern.ch/record/257493 (Cited on page 12).

[35] G. Luders. “Proof of the TCP theorem”, *Annals Phys.* 2 (1957). [Annals Phys. 281, 1004 (2000)], pp. 1–15. DOI: 10.1016/0003-4916(57)90032-5 (Cited on page 12).

[36] C. S. Wu et al. “Experimental test of parity conservation in beta decay”, *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413 (Cited on page 12).

[37] J. H. Christenson et al. “Evidence for the 2π decay of the K_2^0 meson”, *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138 (Cited on page 12).

[38] A. Bettini. “Introduction to elementary particle physics”, Cambridge: Cambridge Univ. Press, 2008. URL: https://cds.cern.ch/record/1111396 (Cited on page 12).

[39] HFLAV Group, "Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016", *Eur. Phys. J.* C77.12 (2017), p. 895. arXiv: 1612.07233 [hep-ex] (Cited on pages 13, 19, 91, 103, 156, 182).

[40] LHCb Collaboration, R. Aaij et al. "Measurement of the ratio of branching fractions $BR(B^0 \to K^{*0}\gamma)/BR(B_s^0 \to \phi\gamma)$ and the direct CP asymmetry in $B^0 \to K^{*0}\gamma$", *Nucl. Phys.* B867 (2013), pp. 1–18. arXiv: 1209.0313 [hep-ex] (Cited on page 14).

[41] D. Atwood, M. Gronau, and A. Soni. "Mixing induced CP asymmetries in radiative B decays in and beyond the standard model", *Phys. Rev. Lett.* 79 (1997), pp. 185–188. arXiv: hep-ph/9704272 [hep-ph] (Cited on pages 14, 18).

[42] F. Muheim, Y. Xie, and R. Zwicky. "Exploiting the width difference in $B_s \to \phi\gamma$", *Phys. Lett.* B664 (2008), pp. 174–179. arXiv: 0802.0876 [hep-ph] (Cited on pages 16, 19).

[43] LHCb Collaboration, R. Aaij et al. "Observation of photon polarization in the $b \to s\gamma$ transition", *Phys. Rev. Lett.* 112.16 (2014), p. 161801. arXiv: 1402.6852 [hep-ex] (Cited on page 18).

[44] M. Gronau et al. "Measuring the photon polarization in $B \to K\pi\pi\gamma$", *Phys. Rev. Lett.* 88 (2002), p. 051802. arXiv: hep-ph/0107254 [hep-ph] (Cited on page 18).

[45] M. Gronau and D. Pirjol. "Reexamining the photon polarization in $B \to K\pi\pi\gamma$", *Phys. Rev.* D96.1 (2017), p. 013002. arXiv: 1704.05280 [hep-ph] (Cited on page 18).

[46] LHCb Collaboration, R. Aaij et al. "Angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay in the low-q^2 region", *JHEP* 04 (2015), p. 064. arXiv: 1501.03038 [hep-ex] (Cited on pages 18, 164).

[47] Y. Grossman and D. Pirjol. "Extracting and using photon polarization information in radiative B decays", *JHEP* 06 (2000), p. 029. arXiv: hep-ph/0005069 [hep-ph] (Cited on page 18).

[48] T. Mannel and S. Recksiegel. "Flavor changing neutral current decays of heavy baryons: the case $\Lambda_b \to \Lambda\gamma$", *J. Phys.* G24 (1998), pp. 979–990. arXiv: hep-ph/9701399 [hep-ph] (Cited on page 18).

[49] G. Hiller and A. Kagan. “Probing for new physics in polarized Λ_b decays at the Z”, *Phys. Rev.* D65 (2002), p. 074038. arXiv: hep-ph/0108074 [hep-ph] (Cited on page 18).

[50] F. Legger and T. Schietinger. “Photon helicity in $\Lambda_b \to pK\gamma$ decays”, *Phys. Lett.* B645 (2007). [Erratum: Phys. Lett.B647,527(2007)], pp. 204–212. arXiv: hep-ph/0605245 [hep-ph] (Cited on page 18).

[51] D. Atwood et al. “Mixing-induced CP violation in $B \to P_1 P_2 \gamma$ in search of clean new physics signals”, *Phys. Rev.* D71 (2005), p. 076003. arXiv: hep-ph/0410036 [hep-ph] (Cited on page 18).

[52] BaBar Collaboration, B. Aubert et al. “Measurement of time-dependent CP asymmetry in $B^0 \to K^0_S \pi^0 \gamma$ decays”, *Phys. Rev.* D78 (2008), p. 071102. arXiv: 0807.3103 [hep-ex] (Cited on pages 18, 164).

[53] Belle Collaboration, Y. Ushiroda et al. “Time-dependent CP asymmetries in $B^0 \to K^0_S \pi^0 \gamma$ transitions”, *Phys. Rev.* D74 (2006), p. 111104. arXiv: hep-ex/0608017 [hep-ex] (Cited on pages 18, 164).

[54] BaBar Collaboration, B. Aubert et al. “Branching fractions and CP-violating asymmetries in radiative B decays to $\eta K \gamma$”, *Phys. Rev.* D79 (2009), p. 011102. arXiv: 0805.1317 [hep-ex] (Cited on page 18).

[55] BaBar Collaboration, P. del Amo Sanchez et al. “Time-dependent analysis of $B^0 \to K^0_S \pi^- \pi^+ \gamma$ decays and studies of the $K^+\pi^-\pi^+$ system in $B^+ \to K^+\pi^-\pi^+\gamma$ decays”, *Phys. Rev.* D93.5 (2016), p. 052013. arXiv: 1512.03579 [hep-ex] (Cited on page 18).

[56] Belle Collaboration, J. Li et al. “Time-dependent CP asymmetries in $B^0 \to K^0_S \rho^0 \gamma$ decays”, *Phys. Rev. Lett.* 101 (2008), p. 251601. arXiv: 0806.1980 [hep-ex] (Cited on page 18).

[57] Belle Collaboration, H. Sahoo et al. “First observation of radiative $B^0 \to \phi K^0 \gamma$ decays and measurements of their time-dependent CP violation”, *Phys. Rev.* D84 (2011), p. 071101. arXiv: 1104.5590 [hep-ex] (Cited on page 19).

[58] Belle Collaboration, Y. Ushiroda et al. “Time-dependent CP-violating asymmetry in $B^0 \to \rho^0 \gamma$ decays”, *Phys. Rev. Lett.* 100 (2008), p. 021602. arXiv: 0709.2769 [hep-ex] (Cited on page 19).

[59] LHCb Collaboration, R. Aaij et al. "First experimental study of photon polarization in radiative B_s^0 decays", *Phys. Rev. Lett.* 118.2 (2017). [Addendum: *Phys. Rev. Lett.* 118.10, p.109901 (2017)], p. 021801. arXiv: 1609.02032 [hep-ex] (Cited on pages 19, 61, 85, 167, 180, 185).

[60] LHCb Collaboration, C. Sánchez Mayordomo. "Radiative b-hadron decays at LHCb", *Proceedings, 38th International Conference on High Energy Physics (ICHEP 2016)*. (Chicago, USA). PoS, 2016, p. 585 (Cited on page 19).

[61] L. Evans and P. Bryant. "LHC Machine", *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001 (Cited on page 21).

[62] C. Lefèvre. "The CERN accelerator complex. Complexe des accélérateurs du CERN". Dec. 2008. URL: https://cds.cern.ch/record/1260465 (Cited on page 22).

[63] CERN. *The Large Hadron Collider*. URL: https://home.cern/topics/large-hadron-collider (Cited on page 23).

[64] A. Riotto and M. Trodden. "Recent progress in baryogenesis", *Ann. Rev. Nucl. Part. Sci.* 49 (1999), pp. 35–75. arXiv: hep-ph/9901362 [hep-ph] (Cited on page 24).

[65] CERN Bulletin. *LHC report: playing with angles*. URL: https://cds.cern.ch/journal/CERNBulletin/2016/38/News%20Articles/2216373 (Cited on page 24).

[66] LHCb Collaboration. *LHCb Operations Plots Webpage*. URL: https://lbgroups.cern.ch/online/OperationsPlots/index.htm (Cited on page 25).

[67] LHCb Collaboration, R. Aaij et al. "Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0}\mu^+\mu^-$", *Phys. Rev. Lett.* 111 (2013), p. 191801. arXiv: 1308.1707 [hep-ex] (Cited on page 25).

[68] LHCb Collaboration, R. Aaij et al. "Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity", *JHEP* 02 (2016), p. 104. arXiv: 1512.04442 [hep-ex] (Cited on page 25).

[69] LHCb Collaboration, R. Aaij et al. “Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays”, *JHEP* 08 (2017), p. 055. arXiv: 1705.05802 [hep-ex] (Cited on page 25).

[70] LHCb Collaboration, R. Aaij et al. “Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$”, *JHEP* 09 (2015), p. 179. arXiv: 1506.08777 [hep-ex] (Cited on page 25).

[71] LHCb Collaboration, R. Aaij et al. “Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$”, *Phys. Rev. Lett.* 115.11 (2015). [Erratum: Phys. Rev. Lett.115,no.15,159901(2015)], p. 111803. arXiv: 1506.08614 [hep-ex] (Cited on page 25).

[72] LHCb Collaboration, R. Aaij et al. “Measurement of the ratio of the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ-lepton decays”, (2017). arXiv: 1708.08856 [hep-ex] (Cited on page 25).

[73] LHCb Collaboration, R. Aaij et al. “Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$”, *Phys. Rev. Lett.* 120.12 (2018), p. 121801. arXiv: 1711.05623 [hep-ex] (Cited on page 25).

[74] LHCb Collaboration, R. Aaij et al. “First evidence for the decay $B_s^0 \to \mu^+\mu^-$”, *Phys. Rev. Lett.* 110.2 (2013), p. 021801. arXiv: 1211.2674 [hep-ex] (Cited on page 25).

[75] LHCb, CMS Collaboration, V. Khachatryan et al. “Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data”, *Nature* 522 (2015), pp. 68–72. arXiv: 1411.4413 [hep-ex] (Cited on page 25).

[76] LHCb Collaboration, R. Aaij et al. “Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays”, *Phys. Rev. Lett.* 118.19 (2017), p. 191801. arXiv: 1703.05747 [hep-ex] (Cited on page 25).

[77] LHCb Collaboration, R. Aaij et al. “Determination of the X(3872) meson quantum numbers”, *Phys. Rev. Lett.* 110 (2013), p. 222001. arXiv: 1302.6269 [hep-ex] (Cited on page 26).

[78] LHCb Collaboration, R. Aaij et al. “First observation of CP violation in the decays of B_s^0 mesons”, *Phys. Rev. Lett.* 110.22 (2013), p. 221601. arXiv: 1304.6173 [hep-ex] (Cited on page 26).

[79] LHCb Collaboration, R. Aaij et al. “Evidence for CP violation in time-integrated $D^0 \to h^- h^+$ decay rates”, *Phys. Rev. Lett.* 108 (2012), p. 111602. arXiv: 1112.0938 [hep-ex] (Cited on page 26).

[80] LHCb Collaboration, R. Aaij et al. “Measurement of matter-antimatter differences in beauty baryon decays”, *Nature Phys.* 13 (2017), pp. 391–396. arXiv: 1609.05216 [hep-ex] (Cited on page 26).

[81] LHCb Collaboration, R. Aaij et al. “Measurement of the CP-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\phi$”, *Phys. Rev. Lett.* 108 (2012), p. 101803. arXiv: 1112.3183 [hep-ex] (Cited on page 26).

[82] LHCb Collaboration, R. Aaij et al. “LHCb Detector Performance”, *Int. J. Mod. Phys.* A30.07 (2015), p. 1530022. arXiv: 1412.6352 [hep-ex] (Cited on pages 26, 45, 46, 48).

[83] LHCb Collaboration. “The LHCb Detector at the LHC”, *Journal of Instrumentation* 3.08 (2008), S08005. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08005 (Cited on page 27).

[84] R. Aaij et al. “Performance of the LHCb Vertex Locator”, *JINST* 9 (2014), p. 09007. arXiv: 1405.7808 [physics.ins-det] (Cited on page 29).

[85] LHCb Collaboration. *The LHCb Detector*. URL: http://lhcb-public.web.cern.ch/lhcb-public/en/detector/Detector-en.html (Cited on page 35).

[86] M. Brice. “The LHCb electromagnetic calorimeter”. May 2005. URL: https://cds.cern.ch/record/835712 (Cited on page 36).

[87] LHCb Collaboration. *Trigger Schemes*. URL: http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html (Cited on page 39).

[88] LHCb Collaboration, R. Aaij et al. “Absolute luminosity measurements with the LHCb detector at the LHC”, *JINST* 7 (2012), P01010. arXiv: 1110.2866 [hep-ex] (Cited on page 40).

[89] G. Corti et al. "Software for the LHCb experiment", *IEEE Trans. Nucl. Sci.* 53 (2006), pp. 1323–1328. DOI: 10.1109/TNS.2006.872627 (Cited on page 41).

[90] G. Barrand et al. "GAUDI - A software architecture and framework for building HEP data processing applications", *Comput. Phys. Commun.* 140 (2001), pp. 45–55. DOI: 10.1016/S0010-4655(01)00254-5 (Cited on page 41).

[91] LHCb Collaboration, M. Clemencic et al. "The LHCb simulation application, Gauss: Design, evolution and experience", *J. Phys. Conf. Ser.* 331 (2011), p. 032023. DOI: 10.1088/1742-6596/331/3/032023 (Cited on page 42).

[92] LHCb Collaboration. *The GAUSS project*. URL: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/ (Cited on page 42).

[93] T. Sjöstrand et al. "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* 191 (2015), pp. 159–177. arXiv: 1410.3012 [hep-ph] (Cited on page 42).

[94] D. J. Lange. "The EvtGen particle decay simulation package", *Nucl. Instrum. Meth.* A462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4 (Cited on page 42).

[95] E. Barberio and Z. Was. "PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0", *Comput. Phys. Commun.* 79 (1994), pp. 291–308. DOI: 10.1016/0010-4655(94)90074-4 (Cited on page 42).

[96] GEANT4 Collaboration, S. Agostinelli et al. "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8 (Cited on page 42).

[97] LHCb Collaboration. *The BOOLE project*. URL: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/ (Cited on page 42).

[98] LHCb Collaboration. *The MOORE project*. URL: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/ (Cited on page 42).

[99] LHCb Collaboration. *The BRUNEL project*. URL: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/ (Cited on pages 42, 50).

[100] LHCb Collaboration. *The DAVINCI project*. URL: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/ (Cited on page 42).

[101] R. Brun and F. Rademakers. "ROOT: An object oriented data analysis framework", *Nucl. Instrum. Meth.* A389 (1997), pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X (Cited on pages 42, 95).

[102] W. Verkerke and D. P. Kirkby. "The RooFit toolkit for data modeling", *eConf* C0303241 (2003), MOLT007. arXiv: physics/0306116 [physics] (Cited on pages 43, 92, 144).

[103] F. James and M. Roos. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations", *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9 (Cited on pages 43, 92, 95, 144).

[104] M. Pivk and F. R. Le Diberder. "SPlot: A Statistical tool to unfold data distributions", *Nucl. Instrum. Meth.* A555 (2005), pp. 356–369. arXiv: physics/0402083 [physics.data-an] (Cited on pages 43, 71).

[105] O. Deschamps et al. "Photon and neutral pion reconstruction", LHCb-2003-091 (Sept. 2003). URL: http://cds.cern.ch/record/691634 (Cited on pages 45, 48).

[106] V. Breton, N. Brun, and P. Perret. "A clustering algorithm for the LHCb electromagnetic calorimeter using a cellular automaton", LHCb-2001-123 (Sept. 2001). URL: https://cds.cern.ch/record/681262 (Cited on page 46).

[107] M. Calvo et al. "Photon polarization in $B^0_s \to \phi\gamma$", LHCb-ANA-2014-102 (Nov. 2014). URL: https://cds.cern.ch/record/1971169 (Cited on pages 47, 73, 83).

[108] M. Calvo et al. "A tool for γ/π^0 separation at high energies", LHCb-PUB-2015-016 (Aug. 2015). URL: http://cds.cern.ch/record/2042173 (Cited on pages 49, 68).

[109] F. Soomro and V. Belyaev. "HLT2 exclusive selections for $B_s \to \phi\gamma$ and $B_d \to K^*\gamma$", LHCb-PUB-2010-007 (Feb. 2010). URL: http://cds.cern.ch/record/1244785 (Cited on page 58).

[110] L. Anderlini et al. "The PIDCalib package", LHCb-PUB-2016-021 (July 2016). URL: https://cds.cern.ch/record/2202412 (Cited on page 63).

[111] L. Shchutska, A. Golutvin, and I. Belyaev. "Study of radiative penguin decays $B \to K^{*}\gamma$ and $B_s \to \phi\gamma$ at LHCb", LHCb-2007-030 (May 2007). URL: http://cds.cern.ch/record/1033945 (Cited on page 68).

[112] ARGUS Collaboration, H. Albrecht et al. "Search for hadronic $b \to u$ decays", *Phys. Lett.* B241 (1990), pp. 278–282. DOI: 10.1016/0370-2693(90)91293-K (Cited on page 76).

[113] E. Ben Haim and L. Henry. "Correcting $_s\mathcal{P}lots$ in the presence of fixed yields ", LHCb-INT-2016-014. CERN-LHCb-INT-2016-014 (Mar. 2016). URL: https://cds.cern.ch/record/2139626 (Cited on page 78).

[114] P. Ruiz Valls. "Photon polarization in $B_s^0 \to \phi\gamma$ decays at the LHCb experiment". PhD Thesis, presented 13 Dec 2016. URL: http://cds.cern.ch/record/2239984 (Cited on page 82).

[115] Z. Xu. "Searches for the Anomalous Photon Polarisation in Radiative B Decays at LHCb". PhD Thesis, presented 18 Nov 2016. URL: http://cds.cern.ch/record/2235201 (Cited on page 82).

[116] James Phillips. *ZunZun: a site for curve fitting 2D and 3D data.* Source code: https://bitbucket.org/zunzuncode/zunzunsite3/overview. URL: http://zunzun.com/ (Cited on page 88).

[117] LHCb Collaboration. *Parameters of Sim08 Monte Carlo productions.* URL: https://twiki.cern.ch/twiki/bin/view/LHCb/SettingsSim08 (Cited on page 89).

[118] G. D. Cowan. "Statistical data analysis", Oxford: Oxford Univ. Press, 1998. URL: https://cds.cern.ch/record/358560 (Cited on page 92).

[119] R. J. Barlow. "Statistics: a guide to the use of statistical methods in the physical sciences", Manchester physics series. Chichester: Wiley, 1989. URL: https://cds.cern.ch/record/213033 (Cited on page 92).

[120] Y. Xie. "sFit: a method for background subtraction in maximum likelihood fit", (2009). arXiv: 0905.0724 [physics.data-an] (Cited on page 92).

[121] Y. Xie. "Some physics and technical issues in analysis of the decay $B_s \to J/\psi K^+K^-$", LHCb-INT-2012-017 (July 2012). URL: https://cds.cern.ch/record/1461783 (Cited on page 93).

[122] P. Clarke et al. "First ambiguity-free measurement of ϕ_s and $\Delta\Gamma_s$ in $B_s^0 \to J/\psi K^+K^-$ decays", (Oct. 2011). LHCb-ANA-2011-083. URL: https://cds.cern.ch/record/1392936 (Cited on page 93).

[123] LHCb Collaboration, R. Aaij et al. "Determination of the Sign of the Decay Width Difference in the B_s^0 System", *Phys. Rev. Lett.* 108 (24 June 2012), p. 241801. DOI: 10.1103/PhysRevLett.108.241801 (Cited on page 93).

[124] C. Sánchez Mayordomo. *FF2.0: fit framework for binned measurements*. URL: https://gitlab.cern.ch/casanche/FF2.0 (Cited on pages 95, 104, 179).

[125] D. W. Scott. "Multivariate density estimation: theory, practice, and visualization; 2nd ed." Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley, 2015. URL: https://cds.cern.ch/record/2009863 (Cited on page 96).

[126] L. Denby and C. Mallows. "Variations on the Histogram", *J. Comput. Graph. Stat.* 18.1 (2009), pp. 21–31. DOI: 10.1198/jcgs.2009.0002 (Cited on page 96).

[127] R. Aaij et al. "Opposite-side flavour tagging of B mesons at the LHCb experiment", *The European Physical Journal C* 72.6 (June 7, 2012), p. 2022. DOI: 10.1140/epjc/s10052-012-2022-1 (Cited on pages 117, 118).

[128] LHCb Collaboration, R. Aaij et al. "*B* flavour tagging using charm decays at the LHCb experiment", *JINST* 10.10 (2015), P10005. arXiv: 1507.07892 [hep-ex] (Cited on page 117).

[129] LHCb Collaboration, R. Aaij et al. "A new algorithm for identifying the flavour of B_s^0 mesons at LHCb", *JINST* 11.05 (2016), P05010. arXiv: 1602.07252 [hep-ex] (Cited on pages 117, 118).

[130] R. Aaij and etal. “New algorithms for identifying the flavour of B^0 mesons using pions and protons”, *The European Physical Journal C* 77.4 (Apr. 12, 2017), p. 238. DOI: 10.1140/epjc/s10052-017-4731-y (Cited on page 117).

[131] LHCb Collaboration. *Scheme of tagging algorithms.* URL: https://twiki.cern.ch/twiki/bin/view/LHCb/FlavourTaggingConferencePlots (Cited on pages 117, 124).

[132] LHCb Collaboration. *Flavour Tagging Performances and Calibrations for Reco14 data.* URL: https://twiki.cern.ch/twiki/bin/viewauth/LHCbPhysics/FlavourTagging4Reco14 (Cited on pages 120, 121).

[133] Jack Wimberley. *EspressoPerformanceMonitor.* URL: https://gitlab.cern.ch/lhcb-ft/EspressoPerformanceMonitor (Cited on page 121).

[134] M. Calvo, M. Grabalosa, and M. Musy. “Combination of same-side with opposite-side flavour tagging”, LHCb-PUB-2009-027 (Feb. 2010). URL: https://cds.cern.ch/record/1223083 (Cited on page 125).

[135] LHCb Collaboration, R. Aaij et al. “Precision measurement of CP violation in $B_s^0 \to J/\psi K^+K^-$ decays”, *Phys. Rev. Lett.* 114.4 (2015), p. 041801. arXiv: 1411.3104 [hep-ex] (Cited on pages 125, 127).

[136] LHCb Collaboration, R. Aaij et al. “Measurement of the CP-violating phase ϕ_s in $\bar{B}_s^0 \to D_s^+ D_s^-$ decays”, *Phys. Rev. Lett.* 113.21 (2014), p. 211801. arXiv: 1409.4619 [hep-ex] (Cited on pages 125, 127).

[137] M. Oreglia. “A Study of the Reactions $\psi' \to \gamma\gamma\psi$”. PhD thesis. SLAC, 1980. URL: http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=slac-r-236.html (Cited on page 135).

[138] J. E. Gaiser. “Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ'”. PhD thesis. SLAC, 1982. URL: http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=slac-r-255.html (Cited on page 135).

[139] T. Skwarnicki. “A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances”. PhD thesis. Cracow, INP, 1986. URL: http://www-library.desy.de/cgi-bin/showprep.pl?DESY-F31-86-02 (Cited on page 135).

[140] C. Sánchez Mayordomo and C. Remón Alepuz. *rad-bs2phigamma-tagged: Time-dependent tagged analysis for $B_s^0 \to \phi\gamma$*. URL: https://gitlab.cern.ch/LHCb-RD/rad-bs2phigamma-tagged (Cited on page 144).

[141] M. T. Schiller. "RooFit Tutorial on advanced topics", *50th Analysis and Software Week*. CERN, 2013. URL: https://indico.cern.ch/event/201535/contributions/384770/ (Cited on pages 146, 147).

[142] LHCb Collaboration. *B2DX Fitters*. URL: https://gitlab.cern.ch/lhcb/Urania/tree/master/PhysFit/B2DXFitters (Cited on pages 147, 182).

[143] M. T. Schiller. "Overview of the B2DXFitters package", *LHCb B2OC TD workshop*. Padova, 2015. URL: https://agenda.infn.it/contributionDisplay.py?contribId=3&sessionId=1&confId=9718 (Cited on page 147).

[144] LHCb Collaboration. *P2VV Fitters*. URL: https://gitlab.cern.ch/lhcb/Urania/tree/master/PhysFit/P2VV (Cited on pages 147, 182).

[145] G. Tellarini and S. Vecchi. "Experience with the B2DXFitters package in $B_s^0 \to D_s^+ K^-$ analysis", *LHCb B2OC TD workshop*. Padova, 2015. URL: https://agenda.infn.it/contributionDisplay.py?contribId=6&sessionId=8&confId=9718 (Cited on page 147).

[146] T. M. Karbach, G. Raven, and M. Schiller. "Decay time integrals in neutral meson mixing and their efficient evaluation", LHCb-INT-2013-043 (July 2013). URL: https://cds.cern.ch/record/1563169 (Cited on page 147).

[147] S. Descotes-Genon et al. "Exploring New Physics in the C7-C7' plane", *JHEP* 06 (2011), p. 099. arXiv: 1104.3342 [hep-ph] (Cited on page 161).

[148] K. G. Chetyrkin, M. Misiak, and M. Munz. "Weak radiative B meson decay beyond leading logarithms", *Phys. Lett.* B400 (1997). [Erratum: Phys. Lett.B425,414(1998)], pp. 206–219. arXiv: hep-ph/9612313 [hep-ph] (Cited on page 161).

[149] M. Gorbahn and U. Haisch. "Effective Hamiltonian for non-leptonic $|\Delta F| = 1$ decays at NNLO in QCD", *Nucl. Phys.* B713 (2005), pp. 291–332. arXiv: hep-ph/0411071 [hep-ph] (Cited on page 161).

[150] K. G. Wilson. “Nonlagrangian models of current algebra”, *Phys. Rev.* 179 (1969), pp. 1499–1512. DOI: 10.1103/PhysRev.179.1499 (Cited on page 161).

[151] M. Antonelli et al. “Flavor Physics in the Quark Sector”, *Phys. Rept.* 494 (2010), pp. 197–414. arXiv: 0907.5386 [hep-ph] (Cited on page 161).

[152] A. J. Buras. “Weak Hamiltonian, CP Violation and Rare Decays”, tech. rep. München: München Univ. Fachber. Phys., June 1998. URL: https://cds.cern.ch/record/358151 (Cited on page 161).

[153] G. Buchalla, A. J. Buras, and M. E. Lautenbacher. “Weak decays beyond leading logarithms”, *Rev. Mod. Phys.* 68 (1996), pp. 1125–1144. arXiv: hep-ph/9512380 [hep-ph] (Cited on page 161).

[154] D. Straub et al. *flav-io/flavio v0.28*. Apr. 2018. URL: https://doi.org/10.5281/zenodo.1218732 (Cited on page 162).

[155] J. Aebischer, J. Kumar, and D. M. Straub. “Wilson: a Python package for the running and matching of Wilson coefficients above and below the electroweak scale”, (2018). arXiv: 1804.05033 [hep-ph] (Cited on page 162).

[156] A. Paul and D. M. Straub. “Constraints on new physics from radiative B decays”, *Journal of High Energy Physics* 2017.4 (Apr. 5, 2017), p. 27. DOI: 10.1007/JHEP04(2017)027 (Cited on page 162).