Università degli Studi di Cagliari

Ph.D. DEGREE

Physics

Cycle XXXIV

TITLE OF THE Ph.D. THESIS

Precise measurements of CP-violating and mixing parameters of B^0_s and B^0_d systems with the LHCb detector.

Scientific Disciplinary Sector(s)

FIS/04

Ph.D. Student: Piera Muzzetto

Supervisor A.Cardini, F. Dordei

Final exam. Academic Year 2020/2021
Thesis defence: April 2022 Session

Precise measurements of CP-violating and mixing parameters of B_s^0 and B_d^0 systems with the LHCb detector

by

Piera Muzzetto

Department of Physics
Universita' degli Studi di Cagliari
Dottorato di Ricerca in Fisica - Ciclo XXXIV

A dissertation submitted to the University of Cagliari in accordance with the requirements of the degree of Doctor of Philosophy.

Supervisors:
Dr. A. Cardini
Dr. F. Dordei

Coordinator:
Prof. P. Ruggerone

18th March 2022

Abstract

In this dissertation, the analysis strategies and the preliminary results on the ongoing studies of *CP*-violating and mixing parameters of the B_s^0 and B_d^0 systems are presented. The analysed dataset corresponds to the proton-proton interactions collected by the LHCb experiment from 2015 to 2018 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV/c^2, during the so-called Run 2, corresponding to an integrated luminosity of 5.9 fb^{-1}.

In particular, the determination of the *CP*-violating phase ϕ_s, of the decay-width difference $\Delta\Gamma_s$ and of the decay widths Γ_s and Γ_H in the decay channels $B_s^0 \to J/\psi\phi$ and $B_s^0 \to J/\psi\pi^+\pi^-$ are discussed, together with the measurement of the mixing parameter $\Delta\Gamma_d$, exploiting the decay channels $B_d^0 \to J/\psi K^*(892)^0$ and $B_d^0 \to J/\psi K_s^0$. These measurements represent an important test of the Standard Model (SM), the current most complete theoretical description of elementary particles and their interactions, and a sensitive probe of physics phenomena beyond the SM.

The analysis presented in this thesis, using the $B_s^0 \to J/\psi\phi$ decay, allows to obtain a statistical precision on the ϕ_s measurement of 0.022 rad and thus it will be the world's most precise single measurement of this parameter, with the possibility to observe *CP* violation in the interference between a decay with and without mixing.

The preliminary measurement of $\Delta\Gamma_d$ using data collected in 2016 shows a statistical precision of 0.008 ps^{-1} and an achievable precision, with the full Run 2 dataset, of 0.004 ps^{-1}. This final measurement will supersede the precision reached in the previous LHCb analysis [1] and of the current world average value, $\Delta\Gamma_d = 0.0007 \pm 0.0066$ ps^{-1} [2], representing the major contribution to the future world average combination.

Introduction

Since the dawn of time, human beings have questioned the nature of matter and cosmos, with methods of investigation that have gradually improved throughout history, from mythological explanations to contemporary science, during a long journey that has seen the development of increasingly sophisticated experiments and of theories that have been successful and others that have been abandoned. It has been shown that atoms, which in ancient times were thought to be the elementary constituents of matter, are divisible and composed of protons, neutrons and electrons. Similarly, it was discovered that protons and neutrons are particles composed of more elementary constituents. Moreover, thanks to the discovery of new other particles, it was understood that they are organised into leptons, quarks and bosons and further divided into generations, and that they respect specific rules of symmetry, that regulate their nature and their interactions. Today, after this long journey of many years, we have come to a theory called the Standard Model (SM) of particle physics, a quantum field theory that is able to explain, in the most complete way possible, the nature of the elementary constituents of matter, antimatter and their mutual interactions, namely the weak, electromagnetic and strong interactions. The theory has been extensively tested during the last decades and it recently received, in 2012, an important new confirmation with the observation of the Higgs boson, responsible for the spontaneous breaking of the electroweak symmetry that allows elementary particles to acquire mass.

However, several open problems remain unanswered in the context of the SM. First of all, from a theoretical point of view, the SM is currently characterised by several free parameters, and it is not able to include and describe the gravitational interaction in a consistent and fundamental way. Secondly, many experimental pieces of evidence, most of them coming from cosmological observations, cannot be explained in the context of the SM. For instance, it has been observed that only the 4% of the energy density of the universe is described by the SM, while the remaining one is composed of the so-called dark matter and dark energy, whose nature is not yet understood. Another cosmological observation is the too large discrepancy between the amount of matter and antimatter present in the universe, and the matter-antimatter unbalance deducible from experimental observations of SM parameters. This problem is strictly linked to the so-called flavour sector of the SM, which describes the different particle flavours and the transitions between them, whose physics is governed by the Cabibbo-Kobayashi-Maskawa

(CKM) matrix by means of the weak interaction, where the violation of the *CP* symmetry of the SM comes into play.

Since the so far measured SM *CP* violation is not sufficient to explain the observed discrepancy, theoretical models describing New Physics (NP) beyond the SM have been developed in the last decades, with the introduction of new heavy particles or new interactions, in which the magnitude of the *CP* violation is enhanced. Since then, the search for new physics effects has become one of the most pursued ways in the field of high energy physics. The heavy degrees of freedom, predicted by the NP models, can also be searched indirectly by studying SM high order (loop) processes, where the heavy particles may contribute significantly, modifying branching fractions, decay-time distributions, angular distributions and finally *CP*-violating parameters, with respect to the SM predictions.

The LHCb experiment is one of the four main experiments exploiting the protons colliding in the Large Hadron Collider (LHC) in order to answer to the aforementioned currently open problems. Its research is mainly focused on the context of flavour physics, searching for indirect NP effects. The LHCb experiment, thanks to its geometry and performance, is an excellent experiment where it is possible to study, with very high precision, the physics of hadrons containing a beauty (b) quark.

Indeed, decays of b-hadrons offer a very rich set of observables within the flavour sector, since the b quark can decay into all the other quark flavours (except the top quark, t) and the corresponding b-hadrons can be easily studied experimentally, being characterised by a relatively long lifetime. Among the most studied b-hadron systems there are the neutral B_s^0 and B_d^0 mesons, which oscillate to their corresponding anti-particles and back, during their lifetime. This mixing process, where two exponential decays with different lifetimes contribute to the decay rate, is extremely sensitive to potential contributions from NP phenomena. In particular, the measurement of the *CP*-violating phase ϕ_s, which is one of the key measurements performed by the LHCb experiment, can receive significant corrections in different NP models. It origins from the interference between the amplitude for the direct decay of a B_s^0 meson and the amplitude for decay after the mixing. Its SM theoretical prediction is very precise and it is taken as the indirect determination of the CKM parameter $-2\beta_s$ via a global fit to experimental data within the SM, namely $-2\beta_s = -0.0370^{+0.0007}_{-0.0008}$ rad [3]. Consequently, any observed deviation from the theoretical prediction would be a clear signature of the presence of NP phenomena. Of great importance are also the measurements of the difference between the inverse of the two lifetimes, defined as $\Delta\Gamma_s$ and $\Delta\Gamma_d$ for the B_s^0 and B_d^0 mesons, respectively. Of special interest for its elusiveness is the $\Delta\Gamma_d$ parameter. It has a very small SM value [4] and it is only weakly constrained by measurements [2]

$$\Delta\Gamma_d^{SM} = (2.6 \pm 0.4) \times 10^{-3} \qquad \frac{\Delta\Gamma_d}{\Gamma_d}^{exp} = 0.001 \pm 0.010, \tag{1}$$

where Γ_d is the average lifetime of the B_d^0 meson. Therefore, any improvement in its determination

would provide an important effort in studying the validity of the SM theory.

Currently, the LHCb experiment provides the most precise measurement of ϕ_s, and the two most important decay channels exploited are the $B_s^0 \to J/\psi\phi(1020)$ and the $B_s^0 \to J/\psi\pi^+\pi^-$ decays [5], [6]. In this thesis, a new measurement of ϕ_s with the above-mentioned decays are presented, performed with data recorded by the LHCb detector from 2015 to 2018. In addition, a preliminary measurement of the $\Delta\Gamma_d$ mixing parameter, using 2016 data, is also provided.

The structure of the thesis is the following. In Chapter 1 a theoretical introduction to the topic is provided. Then, after a description of the detector and its performance given in Chapter 2, the general features of a decay-rate analysis are presented in Chapter 3. Chapter 4 and Chapters 5 present the ongoing measurements of the ϕ_s parameter, respectively in the $B_s^0 \to J/\psi\phi(1020)$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decay channels, where a detailed description of the selection of the decays, together with the study of the most important experimental effects already measured, is provided. Chapter 6 is dedicated to the explanation of the $\Delta\Gamma_d$ measurement, and a first preliminary result is also discussed. Finally, in the Conclusion chapter, a summary of the obtained results is given.

The author of this thesis is one of the main proponents of all the ongoing measurements described in this document. In particular, the measurements described in Chapter 4 and 5 have being summarised in the analysis notes [7] and [8]. The specific contributions to the analyses by the author of this thesis are highlighted in each chapter.

Table of Contents

CHAPTER 1

Theory of neutral b-mesons decay in the Standard Model and beyond

This chapter gives an overview of the basic theoretical concepts and formulas that are needed for the measurement of the *CP*-violating phase ϕ_s via $b \to c\bar{c}s$ transitions and the decay width differences in the $B_{s/d}$ systems $\Delta\Gamma_{s/d}$. In particular, after a brief introduction of SM particle theory, we will mainly focus on the *CP* violation in the quark sector. In this context, a detailed description of the charged currents and the CKM matrix will be given, followed by the discussion of the phenomena of mixing and decay of neutral mesons.

1.1 The Standard Model

The SM of particle physics is a renormalisable quantum field theory that encompasses so far the most complete description of elementary particles and their interactions, except for gravity. The interactions are generated by requiring the SM Lagrangian to be invariant under local gauge transformations of the $SU(3) \times SU(2) \times U(1)$ symmetry group. The related group generators, known as *gauge bosons*, describe spin 1 massless particles that mediate the three fundamental interactions: the strong, weak and electromagnetic forces. The invariance of the Lagrangian under the local transformations of $SU(3)$, $SU(2)$ and $U(1)$ symmetry groups leads to the conservation of a charge for each specific interaction, according to the Noether's theorem [9]. The $SU(3)$ group describes the strong interaction by quantum chromodynamics (QCD) theory and the related conserved charge is called *color*. It can take the quantum numbers red, green and blue, as well as the corresponding anti-colours. The $SU(3)$ group has eight generators, called gluons (G^i_μ, with $i = 1 - 8$). The electroweak interaction, the unification of electromagnetic

and weak forces, is instead described by the theory based on the $SU(2) \times U(1)$ group and the conserved charges are the weak isospin and the hyper-charge, with gauge bosons called W^i_μ ($i = 1, 2, 3$) and B_μ respectively. The spontaneous symmetry breaking, implemented through the Brout-Englert-Higgs mechanism [10–12] of the $SU(2) \times U(1)$ group, results in the combination of the electroweak massless gauge bosons into three massive spin 1 particles responsible of the weak interactions, known as $W^\pm$ and Z^0, and one massless electromagnetic boson called photon, γ. The visible matter of the Universe is constituted of fermions characterised by having half-integer spin. They, together with their anti-particles, are described in the theory as harmonic excitations of fermionc fields and are classified into two main groups: leptons and quarks. The first has no color charge so can not interact via strong interaction. Both the species are organised in generations: each generation is composed of doublets, which carry opposite weak isospin, and a particle of a higher generation has greater mass with respect to the corresponding particle of the previous one. While considering particles the doublets which carry weak isospin are left-handed, for the anti-particles the weak isospin is carried only by right-handed doublets. For the quark sector, the three generations are (u, d), (c, s), (t, b), called *up*, *down*, *charm*, *strange*, *top* and *bottom*, respectively. The first element of each generation is generally called up-type quark and has an electric charge of $+2/3e$, while the second is known as a d-type quark with electric charge equal to $-1/3e$. All the particles described by the SM, with their mass, electric charge, spin and interactions are summarised in Tab. 1.1. Colour charged particles, like individual quarks and gluons, can not be separated, because of the so-called confinement. This results in the observation of only neutral colour objects, known as hadrons. Traditionally, hadrons can be subdivided into two groups: mesons, consisting of a quark-antiquark pair and baryons, containing three quarks. The SM Lagrangian consists of three parts and contains all the information to describe the dynamics of the field and of the elementary particles

$$\mathcal{L}_{SM} = \mathcal{L}_{kinetic} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}. \tag{1.1}$$

The $\mathcal{L}_{kinetic}$ describes the dynamics of Dirac fermions

$$\mathcal{L}_{kinetic} = i\overline{\psi}(D^\mu \gamma_\mu)\psi, \tag{1.2}$$

where the interaction terms are obtained introducing the covariant derivative D^μ, that imposes the gauge invariance of the theory

$$D^\mu = \partial^\mu + ig_s G^\mu_a L_a + igW^\mu_b \sigma_b + ig_2 B^\mu Y, \tag{1.3}$$

with g_s, g, and g_2 corresponding to the coupling constants of the G^μ, W^μ and B^μ fields and L_a and σ_b being the Gell-Mann and Pauli matrices, respectively.
The $\mathcal{L}_{Higgs}$ encloses the Higgs scalar field and its potential, responsible for masses of the $SU(2)$ gauge bosons, thought the mechanism of spontaneous symmetry breaking

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2, \tag{1.4}$$

Name	Symbol	Mass	Spin	Electric charge	Interactions
Higgs					
Higgs boson	h	126 GeV	0	0	W
Gauge bosons					
photon	γ	0	1	0	
W bosons	$W^{\pm}$	80.4 GeV	1	± 1	W
Z boson	Z	91.2 GeV	1	0	W
gluons	g	0	1	0	S
Leptons					
electron	e	0.5 MeV	1/2	-1	EM+W
muon	μ	106 MeV	1/2	-1	EM+W
tau	τ	1.78 GeV	1/2	-1	EM+W
electronic neutrino	ν_e	< 0.1 eV	1/2	0	W
muonic neutrino	ν_μ	< 0.1 eV	1/2	0	W
tauonic neutrino	ν_τ	< 0.1 eV	1/2	0	W
Quarks					
up	u	2.3 MeV	1/2	+2/3	EM+W+S
down	d	4.8 MeV	1/2	-1/3	EM+W+S
strange	s	95 MeV	1/2	-1/3	EM+W+S
charm	c	1.3 GeV	1/2	+2/3	EM+W+S
bottom	b	4.5 GeV	1/2	-1/3	EM+W+S
top	t	173 GeV	1/2	+2/3	EM+W+S

TABLE 1.1. List of SM elementary particles taken from [13] with their mass, spin, electric charge and interactions (W stands for weak, S for strong and EM for electromagnetic).

with ϕ a $SU(2)$ doublet composed of two complex scalar fields, ϕ^+ and ϕ^0.
The fermionic fields appear in the kinetic term in five representations: $Q_L^j = (u_L\ d_L)^j$ and L_L^j are, respectively, the quark and lepton left-handed fields, doublets of $SU(2)$, with weak-isospin $T = 1/2$ and third component $T_3 = \pm 1/2$; while u_R^j, d_R^j and l_R are, respectively, the up-type, down-type and lepton right-handed fields, singlets of $SU(2)_L$, thus with $T = 0$. It is important to note that all the fermionic fields are expressed in the interaction basis, thus they are interaction eigenstates. The interactions between the Higgs field and the fermions are described by the $\mathcal{L}_{Yukawa}$ term of the Lagrangian

$$-\mathcal{L}_{Yukawa} = Y_d^{ij}\overline{Q_L}^i\,\phi\,d_R^j + Y_u^{ij}\overline{Q_L}^i\,\widetilde{\phi}\,u_R^j + Y_l^{ij}\overline{L_L}^i\,\phi\,l_R^j + \text{h.c.}. \tag{1.5}$$

The elements Y^{ij} are arbitrary complex matrices, $\widetilde{\phi} = i\sigma^2\phi^\dagger$ and $h.c.$ stands for hermitian conjugate, where σ represents the Pauli matrix. After the spontaneous symmetry breaking, the Higgs doublet can be written as

$$\phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} \rightarrow \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v+H\end{pmatrix}, \tag{1.6}$$

where H is the scalar and real Higgs field, and the following mass terms for the fermion fields emerge from the Lagrangian

$$-\mathcal{L}^{mass}_{Yukawa} = M_d^{ij}\overline{d}_L^i\, d_R^j + M_u^{ij}\overline{u}_L^i\, u_R^j + M_l^{ij}\overline{l}_L^i\, l_R^j + \text{h.c.}. \tag{1.7}$$

In order to obtain the proper mass of the fermions, M_d^{ij} and M_u^{ij} have to be diagonalised. It is possible to find four unitary matrices, $V_{u,L}^{ij}$, $V_{u,R}^{ij}$, $V_{d,L}^{ij}$, $V_{d,R}^{ij}$, such that $\widetilde{M}^{ij}$ is diagonal:

$$\widetilde{M}_u^{ij} = V_{u,L}^{ik}\, M_u^{kl}\, V_{u,R}^{lj}, \tag{1.8a}$$

$$\widetilde{M}_d^{ij} = V_{d,L}^{ik}\, M_d^{kl}\, V_{d,R}^{lj}. \tag{1.8b}$$

We can now express the $\mathcal{L}_{Yukawa}$ as

$$\overline{d}^m_{Li}(\widetilde{M}^d_{ij})d^m_{Rj} + \overline{u}^m_{Li}(\widetilde{M}^u_{ij})u^m_{Rj} + h.c. = \tag{1.9a}$$

$$\overline{d}_{Li}\, V_{d,L}(V^\dagger_{d,L}\, M^d_{ij}\, V_{d,R})\, V^\dagger_{d,R}\, d_{Rj} + \overline{u}_{Li}\, V_{u,L}(V^\dagger_{u,L}\, M^u_{ij}\, V_{u,R})\, V^\dagger_{u,R}\, u_{Rj} + h.c. \tag{1.9b}$$

The quark mass eigenstates $u^{(m)}$, $d^{(m)}$ can be then derived as follows

$$u_L^{i,(m)} = V_{u,L}^{ij\,\dagger}\, u_L^j \quad , \quad u_R^{i,(m)} = V_{u,R}^{ij\,\dagger}\, u_R^j\,, \tag{1.10a}$$

$$d_L^{i,(m)} = V_{d,L}^{ij\,\dagger}\, d_L^j \quad , \quad d_R^{i,(m)} = V_{d,R}^{ij\,\dagger}\, d_R^j\,. \tag{1.10b}$$

As we will see in the next section, the complex nature of M_d^{ij} and M_u^{ij} matrices and the impossibility to diagonalise them with the same transformation leads to the couplings between different generations and to the *CP* violation in the quark sector, opening the way to the field of flavour physics.

1.2 *CP* violation in the Standard Model

The *CP* transformation is defined as the combination of the charge conjugation (C) and the parity (P) transformations and has the effect to transform particles into their anti-particles. In fact, the C-transformation changes the sign of all particles internal quantum numbers, while the P one reverses the direction of the Cartesian axes system:

$$P(x, y, z) = (-x, -y, -z), \tag{1.11a}$$

$$C(e^-) = e^+. \tag{1.11b}$$

Until 1957, these two transformations have been separately considered exact symmetries of all fundamental interactions. Later, studies on ^{60}Co decays carried on by C. S. Wu and co-workers [14] showed, for the first time, the P-transformation violation in β-decays, topic already under question by T. D. Lee and C. N. Yang [15]. The result was confirmed by R. L. Garwin and co-workers [16], with the additional observation of the C-symmetry violation

in mesons decays. These two fundamental transformations have been found to be drastically violated by the weak interactions, while there is no evidence of $P-$ or $C-$ symmetry violation in electromagnetic or strong interactions to date. In 1964, contrary to the expectation at the time, the violation of the *CP* transformation by the weak interactions in the neutral kaon decays has been observed, even if with a very small magnitude($\sim 10^{-3}$) [17]. This was a key turning point for the fundamental interaction physics sector, opening the way to a new and essential field of research. In fact it suggested the existence of a third generation of quarks, the b and the t, later discovered, and implied the possibility to distinguish matter from antimatter, being indicated eventually as one of the necessary conditions to explain the generation of the matter–antimatter asymmetry observed in the Universe. The only source of *CP* violation measured in the SM so far is a non-trivial phase in the so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix, as it will be explained in more details in the next section. Unfortunately, its measured value is too small to explain the observed matter-antimatter asymmetry, suggesting the existence of new sources of *CP* violation not described by the current SM theory. One possibility, highly investigated so far and subject of this thesis, is the presence of *CP*-violating interactions between new and SM particles, that change the magnitude of measurable observables of the SM.
In the next sections, an overview of the charged weak interactions in the quark sector is given, with special focus on the CKM matrix, with its related *CP* violating phase, and on the mixing process and decay of the neutral mesons.

1.2.1 Charged weak current and CKM matrix

Following the explanation given in Section 1.1, the charged weak interaction between left-handed quarks in the interaction basis can be expressed as

$$\begin{aligned}\mathcal{L}_{Kinetic,CC}(Q_L) &= i\overline{Q}^i_L\gamma_\mu\frac{i}{2}gW^\mu_b\sigma_b Q^i_L = \\ &= -\frac{g}{\sqrt{2}}\overline{u}^i_L\gamma_\mu W^{-\mu}d^i_L - \frac{g}{\sqrt{2}}\overline{d}^i_L\gamma_\mu W^{+\mu}u^i_L,\end{aligned}$$

using $W^+ = \frac{1}{\sqrt{2}}(W_1 - iW_2)$ and $W^- = \frac{1}{\sqrt{2}}(W_1 + iW_2)$. After the diagonalisation of the Yukawa matrices, we can now write the Lagrangian in terms of mass eigenstates $(u^{(m)}_L\ d^{(m)}_L)^i$ as

$$\mathcal{L}_{Kinetic,CC} = \frac{g}{\sqrt{2}}\overline{u}^{i(m)}_L(V^u_L V^{d\dagger}_L)_{ij}\gamma_\mu W^{-\mu}d^{j(m)}_L + \frac{g}{\sqrt{2}}\overline{d}^{i(m)}_L(V^d_L V^{u\dagger}_L)_{ij}\gamma_\mu W^{+\mu}u^{j(m)}_L. \tag{1.13}$$

The transformation from weak to mass eigenstates implies the introduction of a new term in the Lagrangian, the so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix, that allows the quark mixing between different generations

$$V_{CKM} = (V^u_L V^{d\dagger}_L)_{ij}. \tag{1.14}$$

By convention, the mass eigenstates and the interaction eigenstates for the up-type quarks are chosen to be equal, whereas the rotation only affects the down-type quarks

$$u_L^i = u_L^{i(m)}, \tag{1.15a}$$

$$d_L^i = V_{CKM} d_L^{i(m)}, \tag{1.15b}$$

that in the explicit form becomes

$$\begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d_L^{(m)} \\ s_L^{(m)} \\ b_L^{(m)} \end{pmatrix}. \tag{1.16}$$

Now it is interesting to verify what is the effect of applying the *CP* transformation to the charged weak Lagrangian; namely

$$\mathcal{L}_{Kinetic,CC} = \frac{g}{\sqrt{2}} \overline{u}_L^{i(m)} V_{CKM\,ij} \gamma_\mu W^{-\mu} d_L^{j(m)} + \frac{g}{\sqrt{2}} \overline{d}_L^{i(m)} V_{CKM\,ij}^* \gamma_\mu W^{+\mu} u_L^{j(m)}, \tag{1.17a}$$

$$\mathcal{L}_{Kinetic,CC}^{CP} = \frac{g}{\sqrt{2}} \overline{d}_L^{i(m)} V_{CKM\,ij} \gamma_\mu W^{+\mu} u_L^{j(m)} + \frac{g}{\sqrt{2}} \overline{u}_L^{i(m)} V_{CKM\,ij}^* \gamma_\mu W^{-\mu} d_L^{j(m)}. \tag{1.17b}$$

From the formulas above, one can conclude that the Lagrangian is unchanged under *CP* transformations only if $V_{CKM\,ij} = V_{CKM\,ij}^*$, thus the *CP* violation in the SM arises only if the nature of the CKM matrix is complex.
Another interesting result arises looking at the consequences of the unitarity of the CMK matrix

$$V_{CKM} V_{CKM}^\dagger = \mathbb{1}. \tag{1.18}$$

This leads to nine unitary relationships, six of which sum to zero (orthogonal relations) and can therefore be represented as unitary triangles in the complex plane. They have all the same area, that quantify the magnitude of the *CP* violation in the SM.
In the literature, there are many different parametrisations of the CKM matrix. A convenient one, introduced by Chau and Keung, uses the Eulero angles θ_{ij}, where i, j refer to the generations, and one irreducible phase δ. Indeed, even though the CKM matrix is a 3x3 matrix, it has only four free parameters. A general nxn matrix has $2n^2$ real parameters: n^2 values and n^2 phases. The unitarity condition implies n^2 constraints and since the phases can be rotated in order to obtain an overall phase, that is irrelevant, there are $2n-1$ removable phases. So, eventually, the CKM matrix has $(n-1)^2 = 4$ free parameters: one phase and three angles. In terms of these quantities the matrix can be written as

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\, e^{i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & -c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{1.19}$$

where the compact notation $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$ and $\delta = \delta_{CKM}$ has been used.

All the CKM matrix elements must be determined experimentally [13, 18–22] and the current knowledge of the CKM matrix elements [13] is reported in the following

$$|V_{CKM}| = \begin{pmatrix} 0.97401 \pm 0.00011 & 0.22650 \pm 0.00048 & 0.00361^{+0.00011}_{-0.00009} \\ 0.22636 \pm 0.00048 & 0.97320 \pm 0.00011 & 0.04053^{+0.00083}_{-0.00061} \\ 0.00854^{+0.00023}_{-0.00016} & 0.03978^{+0.00082}_{-0.00060} & 0.999172^{+0.000024}_{-0.000035} \end{pmatrix}. \tag{1.20}$$

The strength of the CKM elements seem to exhibit a hierarchy. The couplings are more intense between particles of the same generation (Cabibbo-favoured), and become weaker between particles of different generations (Cabibbo-suppressed), in particular between the first and third. This behaviour suggested to Wolfenstein [23] to parametrise the CKM matrix in power of the parameter $\lambda = |V_{us}|/\sqrt{|V_{us}|^2 + |V_{ud}|^2} = \cos\theta_c \simeq 0.22$ and as a function of A, ρ and η. Expanding the V_{CKM} up to $\mathscr{O}(\lambda^3)$ the following form can be obtained

$$V_{CKM} \sim \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathscr{O}(\lambda^4), \tag{1.21}$$

This parametrisation allows to notice that only two of the six unitarity triangles have terms with equal powers in λ, of the order of $\mathcal{O}(\lambda)^3$, making the study of the angles and the sides possible

$$V_{ud}^* V_{ub} + V_{cd}^* V_{cb} + V_{td}^* V_{tb} = 0, \tag{1.22a}$$

$$V_{us}^* V_{ub} + V_{cs}^* V_{cb} + V_{ts}^* V_{tb} = 0. \tag{1.22b}$$

These terms govern the dynamics in the B^0 and B_s^0 meson systems, respectively. The other four triangles are very flat, with one side much smaller then the other two. By convention, the unitary triangles are normalised such that one side has unit length and it points along the real axis. In particular, the triangle sides are divided respectively by $|V_{cd}V_{cb}|$ and $|V_{cs}V_{cb}|$. The apex is located by definition at $(\overline{\rho}, \overline{\eta}$, which are defined in terms of the Wolfenstein parameters ρ and η as follow

$$\overline{\rho} = \rho(1 - \frac{1}{2}\lambda^2) + \mathcal{O}(\lambda^4) \quad \overline{\eta} = \eta(1 - \frac{1}{2}\lambda^2) + \mathcal{O}(\lambda^4). \tag{1.23}$$

Fig. 1.1 shows the unitarity triangles in the complex plane with the aforementioned rotation applied. We can define the angles of the unitary triangles as

$$\alpha \equiv arg\Big[\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\Big] \quad \beta \equiv arg\Big[\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\Big] \quad \gamma \equiv arg\Big[\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\Big] \quad \beta_s \equiv arg\Big[-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\Big]. \tag{1.24}$$

Looking at the definitions of sides and angles of the unitarity triangles we can see that they can be extracted from measurable quantities. Therefore, in order to test the SM theory is fundamental

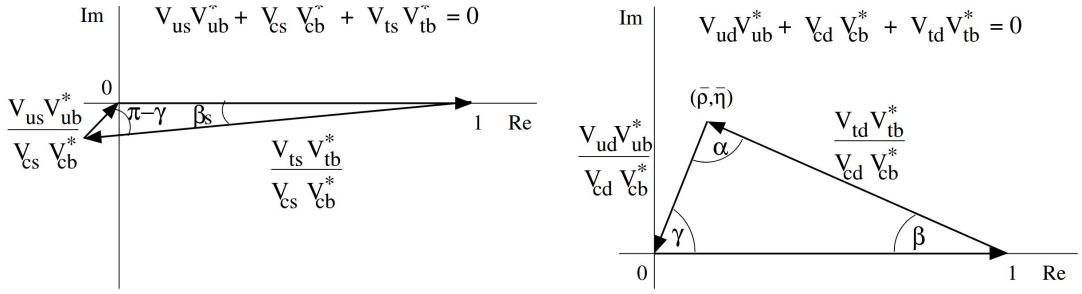

Figure 1.1: The unitarity triangle results from Eqs. 1.22b, for the B_s^0 system at left and for the B_d^0 one at right.

to measure the parameters of the unitarity triangles and, in particular, it is important to extract the sides and the angles in an independent and redundant way. Any disagreement between the angles and the lengths of the sides would necessarily be an evidence of new physics phenomena not described by the SM. In this context, the measurement of the *CP*-violating phase ϕ_s, one of the major topics of this thesis, is of extremely importance, because of its strong relation with the unitarity angle β_s, as it will be shown in Chapter 1.3.1.

1.2.2 Phenomenology of Mixing

The neutral mesons containing the quark b, c and s, can transform into their antiparticles and vice-versa, through flavour changing neutral current weak processes. This meson-antimeson transition occurs continuously in their time evolution, giving rise to the phenomenon commonly named *neutral meson mixing*. As a consequence of the fact that neutral mesons have a non-zero probability of oscillating into their antiparticles before decaying, neutral mesons and anti-mesons constitute a unique two-state system. The reason of the process lies in the fact that the weak interaction eigenstates do not coincide with the eigenstates of the free Hamiltonian, which governs the time evolution of the particles. The phenomenology of the process can be described as follows. Let's consider a generic meson state P^0. In the absence of neutral meson mixing, at time $t > 0$ the state can be expressed as

$$|P^0(t)\rangle = |P^0\rangle\, e^{-\frac{\Gamma t}{2}}\, e^{-\frac{imt}{2}}, \tag{1.25}$$

where m refers to the mass of the meson and $\Gamma = 1/\tau$ is the decay rate of the particle that ensures the exponential decay of the probability density function. This wave-function has to satisfy the time-dependent Schroedinger equation

$$i\frac{\partial}{\partial t}|P^0(t)\rangle = (m - \frac{i}{2}\Gamma)|P^0(t)\rangle, \tag{1.26}$$

defining the action of the *effective* Hamiltonian $\mathcal{H}$ as

$$\mathcal{H}|P^0(t)\rangle = (m - \frac{i}{2}\Gamma)|P^0(t)\rangle. \tag{1.27}$$

Because of the presence of the exponential decay term, the Hamiltonian is not hermitian, reflecting the non-conservation of the probability density function. As a consequence of the oscillation of a meson into its antiparticle, the time evolution of an initial P^0 must include both P^0 and $\overline{P}^0$ components

$$|P(t)\rangle = a(t)|P^0\rangle + b(t)|\overline{P}^0\rangle, \tag{1.28}$$

where the coefficients $a(t)$ and $b(t)$ contain information about the amplitude and phase of P^0 and $\overline{P}^0$ components at the time t. The time evolution of $|P(t)\rangle$, accordingly to Eq. 1.26, can be expressed as

$$i\frac{\partial}{\partial t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} \mathcal{H}_{aa} & \mathcal{H}_{ab} \\ \mathcal{H}_{ba} & \mathcal{H}_{bb} \end{pmatrix}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \tag{1.29}$$

where the 2×2 Hamiltonian matrix, $\mathcal{H}$, can be decomposed into the mass $\mathbf{M}$ and decay hermitian $\mathbf{\Gamma}$ matrices, $\mathcal{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma}$. Let's now require the invariance of $\mathcal{H}$, under the CPT transformation, where T transformation leads to the inversion of the currents flow direction, assumption motivated by conservation of the CPT symmetry in all the measurements performed up to date. This requirement results in $M_{11} = M_{22} \equiv M$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$. Namely, they correspond to the mass and the inverse of the lifetime of the particle. Because of the hermiticity of $\mathbf{M}$ and $\mathbf{\Gamma}$, the off-diagonal elements are complex-conjugated, *i.e.* $M_{21} = M_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$. The off-diagonal mass terms, also called the dispersive part of the mixing amplitude, describe $P^0 \longleftrightarrow \overline{P}^0$ transitions via virtual intermediate state. Conversely, the off-diagonal elements of the decay matrix $\mathbf{\Gamma}$ describe the amplitude change due to decays common to both the P^0 and $\overline{P}^0$ states. The physical eigenstates of $\mathcal{H}$, found diagonalising the *effective* Hamiltonian, can be expressed as the linear combination of the flavour eigenstates

$$|P_{1,2}\rangle = p|P^0\rangle \pm q|\overline{P}^0\rangle, \tag{1.30}$$

normalised by imposing $|q|^2 + |p|^2 = 1$ and where the coefficients p and q are defined such as:

$$\frac{p}{q} = -\sqrt{\frac{2M_{12}^* - i\Gamma_{12}^*}{2M_{12} - i\Gamma_{12}}}. \tag{1.31}$$

It is convenient to define some parameters that are strongly related to the flavour eigenstates: the difference in mass, Δm , and the difference in decay width, $\Delta\Gamma$, between the two physical eigenstates $P_{1,2}$. They are defined as following

$$\Delta m = m_1 - m_2 =\sim 2|M_{12}|\Big(1 - \frac{|\Gamma_{12}|^2}{8|M_{12}|^2}\sin^2\phi_{12}\Big), \tag{1.32a}$$

$$\Delta\Gamma = \Gamma_2 - \Gamma_1 =\sim 2|\Gamma_{12}|\cos\phi_{12}\Big(1 + \frac{|\Gamma_{12}|^2}{8|M_{12}|^2}\sin^2\phi_{12}\Big), \tag{1.32b}$$

where

$$\phi_{12} = \arg(\frac{-M_{12}}{\Gamma_{12}}). \tag{1.33}$$

It can be theoretically demonstrated that for the $B^0(d\bar{b})$ and $B_s^0(s\bar{b})$ systems

$$\Delta m_B \propto |V_{tb}^* V_{tq}|^2, \tag{1.34}$$

where $q = d, s$ depending on the fact that we are referring to B^0 or B_s^0 system.

The time evolution of a state created at $t = 0$ in its flavour eigenstates is then

$$|\psi_{P^0}(t)\rangle = g_+(t)\,|P^0\rangle + \frac{q}{p}\,g_-(t)\,|\overline{P}^0\rangle, \tag{1.35a}$$

$$|\psi_{\overline{P}^0}(t)\rangle = g_+(t)\,|\overline{P}^0\rangle + \frac{p}{q}\,g_-(t)\,|P^0\rangle, \tag{1.35b}$$

where $|\psi_{P^0}(t)\rangle$ and $|\psi_{\overline{P}^0}(t)\rangle$ are the states at time t of an initial $|P^0(t=0)\rangle$ or $|\overline{P}^0(t=0)\rangle$ respectively. The probability to measure, at a time $t > 0$, the same flavour state as the initial state is equal to $|g_+(t)|^2$, whereas the probabilities to measure a different flavour depend on the flavour of the initial state, provided that $|q/p| \neq 1$ and that $g_-(t)$ is a non-zero parameter, and are defined as below

$$|\langle\psi_{P^0}(t)|\overline{P}^0\rangle|^2 = \left|\frac{q}{p}\right|^2 |g_-(t)|^2, \tag{1.36a}$$

$$|\langle\psi_{\overline{P}^0}(t)|P^0\rangle|^2 = \left|\frac{p}{q}\right|^2 |g_-(t)|^2. \tag{1.36b}$$

Usually, the mass and width differences of the physical eigenstates are parametrised in units of the average decay width, through the two dimensionless mixing parameters x and y,

$$x = \frac{M_1 - M_2}{\Gamma}, \tag{1.37a}$$

$$y = \frac{\Gamma_1 - \Gamma_2}{\Gamma}. \tag{1.37b}$$

Thanks to this parametrisation, the evolution coefficients can be simply written as

$$|g_\pm(t)|^2 = \frac{1}{2}\,e^{-\Gamma t}\left(\cosh(y\Gamma t) \pm \cos(x\Gamma t)\right). \tag{1.38}$$

The mixing can take place only if x or y are different from zero. The x parameter governs the periodic oscillation between flavour eigenstates in the mixing process, while the y parameter modifies the exponential decay. The formalism described above is general and can be applied to the different flavours of neutral mesons, although with large differences in the phenomenology of mixing depending on the flavours of the particles. These different behaviours are caused by largely different interactions contributing to the matrix elements responsible for the transitions, M_{12} and Γ_{12}. Let's examine for example the x mixing parameter. If $x \ll 1$ the oscillations

are slow with respect to the time scale of the decay and the flavour is conserved with good approximation, this is the case of $D^0 - \overline{D}^0$ mixing, with $x = 0.003$. If $x \gg 1$ the oscillation are really fast compare to the lifetime and the mesons oscillate many times before decaying, this is the case of $B_s^0 - \overline{B}_s^0$ mixing, with $x = 26.7$. Finally, if $x \sim 1$ the time scale of the decay and the oscillation are approximately the same, which is the case of K^0 and B^0 mesons. The box diagrams responsible to the $\Delta F = 2$ neutral currents that cause the oscillation of B_q^0 mesons, with $q = d, s$, are shown in Fig. 1.3. Inside the FCNC loop, two up-type quarks (u, c, t) are exchanged. In the case of b-mesons, meaning any neutral mesons containing a b-quark, the most relevant contribution to mixing involves the exchange of two top quarks, because of the large breaking of the GIM mechanism caused by the large top-quark mass. The big difference in the mixing parameters for B_s^0 and B^0 mesons can be understood considering that the magnitude of V_{td} CKM element is much smaller than that of V_{ts}. Fig 1.2 shows the different oscillation behaviour in B^0 and B_s^0 mesons. Similar considerations can be done for the y parameter, which is found to be smaller for B^0 than for B_s^0 mesons. This can be explained by looking at the total branching fraction of B decays to final states that are also accessible to $\overline{B}$ decays. Since they are dominated by $b \to ccq$ transitions, which are Cabibbo favoured for B_s^0 decays $(q = s)$ and Cabibbo suppressed for B^0 decays $(q = d)$, the total branching fraction in the B_s^0 system in larger than that of the B_d^0 one, resulting in a larger y value.

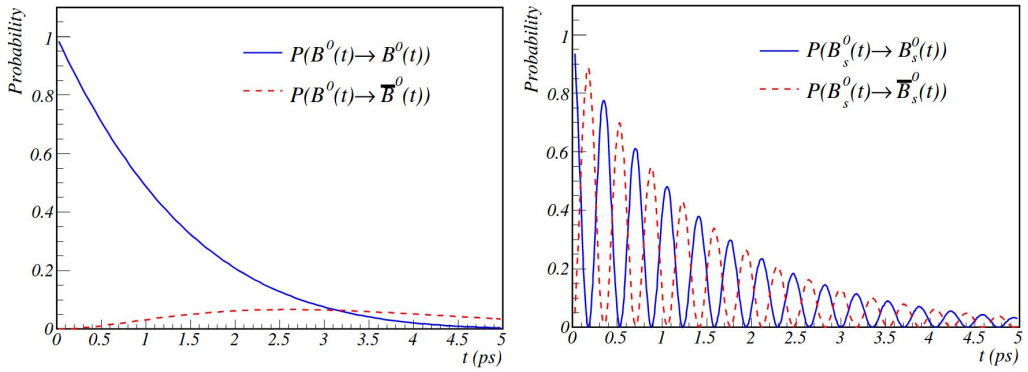

FIGURE 1.2. Probability for a produced B_d^0 (left) and B_s^0 (right) meson to oscillate in its relative antimeson (red) or to preserve its flavour quantum numbers (blue) as a function of time.

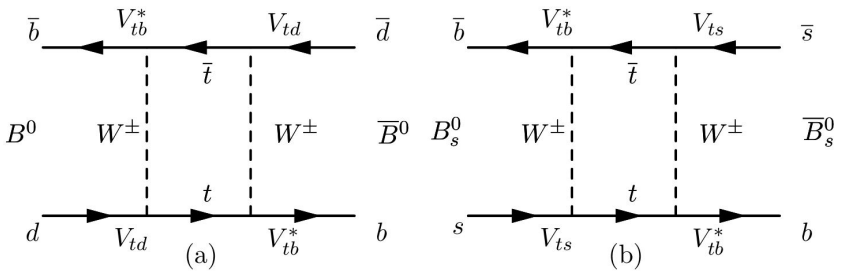

FIGURE 1.3. The lowest order SM contribution to (a) $B^0 - \overline{B}$ and (b) $B_s^0 - \overline{B}_s^0$ mixing. Diagrams with internal quarks swapped with the W are also possible. Diagrams taken from [1].

1.2.3 Classification of CP violation

In Sections 1.2.1 and 1.2.2, the origin of *CP* violation and the phenomenology of the neutral meson mixing in the SM theory have been introduced. It is now useful to define the following amplitudes, which describe the transitions of a meson P^0, or its CP-conjugated state $\overline{P}^0$, towards a generic final state f, or the *CP* conjugated $\overline{f}$

$$\mathcal{A}_{P^0\to f} = \langle f|\mathcal{H}|P^0\rangle, \quad \mathcal{A}_{\overline{P}^0\to f} = \langle f|\mathcal{H}|\overline{P}^0\rangle, \tag{1.39}$$

$$\mathcal{A}_{P^0\to \overline{f}} = \langle \overline{f}|\mathcal{H}|P^0\rangle, \quad \mathcal{A}_{\overline{P}^0\to \overline{f}} = \langle \overline{f}|\mathcal{H}|\overline{P}^0\rangle, \tag{1.40}$$

where $\mathcal{H}$ is the Hamiltonian responsible for the transition. Of special interest, in order to investigate the *CP* violation, is the case when both P^0 and $\overline{P}^0$ mesons can decay into the same final state. Then we can define a further quantity useful to parametrise *CP* violation, which is

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}. \tag{1.41}$$

The probability that the state P^0 produced at $t = 0$ decays in f at $t > 0$ is quantified by the time-dependent decay rate $\Gamma_{P^0\to f}(t) = |\langle f|\mathcal{H}|P^0(t)\rangle|^2$. A general expression of it can now be constructed for the meson and anti-meson as follow

(1.42a)
$$\Gamma_{P^0\to f}(t) = \left|g_+(t)A_f + \frac{q}{p}g_-(t)\overline{A}_f\right|^2 = |A_f|^2\Big(|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 + 2\mathcal{R}\Big[\lambda_f g_+^*(t)g_-(t)\Big]\Big),$$

(1.42b)
$$\Gamma_{\overline{P}^0\to f}(t) = \left|\frac{p}{q}g_-(t)A_f + g_+(t)\overline{A}_f\right|^2 = |A_f|^2\left|\frac{p}{q}\right|^2\Big(|g_-(t)|^2 + |\lambda_f|^2|g_+(t)|^2 + 2\mathcal{R}\Big[\lambda_f g_+(t)g_-^*(t)\Big]\Big).$$

The term proportional to $|A|^2$ is related to the decay occurring without mixing, whereas the terms proportional to $|A|^2\left|\frac{p}{q}\right|^2$ and $|A|^2\left|\frac{q}{p}\right|^2$ are associated with decays following a complete oscillation. The last term, proportional to $\mathcal{R}g^*g$, is instead related to the interference between

the two cases. Substituting Eq. 1.38 into the last definitions leads to the following equations for the decay rates of neutral mesons, known also as master equations

$$\Gamma_{P^0\to f}(t) = \\ = |A_f|^2\frac{e^{-\Gamma t}}{2}\Big((1+|\lambda_f|^2)\cosh\frac{1}{2}\Delta\Gamma t + 2\mathcal{R}\lambda_f\sinh\frac{1}{2}\Delta\Gamma t + (1-|\lambda_f|^2)\cos\Delta m t - 2\mathcal{I}\lambda_f\sin\Delta m t\Big),$$

$$\Gamma_{\overline{P}^0\to f}(t) = \\ = |A_f|^2\Big|\frac{p}{q}\Big|^2\frac{e^{-\Gamma t}}{2}\Big((1+|\lambda_f|^2)\cosh\frac{1}{2}\Delta\Gamma t + 2\mathcal{R}\lambda_f\sinh\frac{1}{2}\Delta\Gamma t - (1-|\lambda_f|^2)\cos\Delta m t + 2\mathcal{I}\lambda_f\sin\Delta m t\Big).$$

Commonly the master equations are expressed in a more compact way as

(1.44a)
$$\Gamma_{P^0\to f}(t) = |A_f|^2\frac{e^{-\Gamma t}}{2}(1+|\lambda_f|^2)\Big(\cosh\frac{1}{2}\Delta\Gamma t + D_f\sinh\frac{1}{2}\Delta\Gamma t + C_f\cos\Delta m t - S_f\sin\Delta m t\Big),$$

(1.44b)
$$\Gamma_{\overline{P}^0\to f}(t) = |A_f|^2\Big|\frac{p}{q}\Big|^2\frac{e^{-\Gamma t}}{2}(1+|\lambda_f|^2)\Big(\cosh\frac{1}{2}\Delta\Gamma t + D_f\sinh\frac{1}{2}\Delta\Gamma t - C_f\cos\Delta m t + S_f\sin\Delta m t\Big),$$

where

$$D_f = \frac{2\mathcal{R}\lambda_f}{1+|\lambda_f|^2} \qquad C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2} \qquad S_f = \frac{2\mathcal{I}\lambda_f}{1+|\lambda_f|^2}. \tag{1.45}$$

It is important to notice that the sinh- and sin- terms are related to the interference between the decays with and without mixing, which are fundamental for the topic of this thesis. Looking at the master equations it seems clear that we can fully describe the decay of a neutral meson if we know the expression of λ_f. Furthermore, it is evident that there is *CP* violation if the equations 1.44a and 1.44b are not identical, so in the case when either $|\lambda_f|$ or $\frac{q}{p}$ deviate from the unity. Phenomenologically, there are two main categories in which the *CP* violation can be studied: through direct decays and through neutral meson mixing. In particular, the *CP* violation phenomena are usually classified into three main areas: *CP* violation in the decay, *CP* violation in the mixing and *CP* violation in the interference between a decay with and without mixing.

CP violation in meson decays

It is the most intuitive type of *CP* violation, that occurs when the probability that a P^0 decays into the final state f is different from that of the decay $\overline{P}^0 \to \overline{f}$, namely

$$\frac{|A_f|}{|\overline{A_{\overline{f}}}|} \neq 1. \tag{1.46}$$

Considering that

$$CP|f\rangle = |\overline{f}\rangle = \eta_{CP}|f\rangle, \tag{1.47}$$

we can rewrite λ_f as a function of the *CP* conjugated amplitude as

$$\lambda_f = \eta_{CP} \frac{q}{p} \frac{\overline{A}_{\overline{f}}}{A_f}. \tag{1.48}$$

Supposing $\frac{q}{p} = 1$, *CP* violation in the decay implies $|\lambda_f| \neq 1$. This is the only type of *CP* violation that can occur in charged meson decays.
Eq. 1.46 shows a necessary but not sufficient condition to observe *CP* violation in the decay. In fact, to arrive at non-vanishing *CP*-violating observables, specific conditions for the transition amplitudes and their phases need to be met. In this context, two types of phases have to be distinguished: the *strong* and *weak* phases. The strong phase typically originates from gluon exchange in the final state. Since strong interactions are *CP* conserving, the strong phase is equal for two *CP* conjugate states. The weak phase comes from complex coupling constants in the Lagrangian and changes sign under *CP* transformation. Considering that observables are accessible only calculating $|A|^2$, in order to avoid the phase cancellation in the expectation value at least two decay amplitudes contributing to the same total one are needed

$$A_{P^0 \to f} = \sum_k |A_k| \, e^{i(\delta_k + \phi_k)} \tag{1.49}$$

$$A_{\overline{P}^0 \to \overline{f}} = \sum_k |A_k| \, e^{i(\delta_k - \phi_k)}. \tag{1.50}$$

From the difference in the expectation values, we can see that *CP* violation is possible thanks to the interference terms of the partial amplitudes,

$$|A_{P^0 \to f}|^2 - |A_{\overline{P}^0 \to \overline{f}}|^2 = -2 \sum_{k,l} |A_k||A_l| \sin(\delta_k - \delta_l) \sin(\phi_k - \phi_l). \tag{1.51}$$

Thus, in order to have *CP* violation directly in the decay process at least two decay amplitudes and different weak and strong phases are necessary. So, differently from the weak and the strong phase alone, relative strong or weak phases between partial amplitudes are physically meaningful.

To quantify the *CP* violation in the decay process the direct time-independent *CP* asymmetry can be defined as

$$\mathcal{A}_{CP}^{\text{dir}} = \frac{|A_{P^0 \to f}|^2 - |A_{\overline{P}^0 \to \overline{f}}|^2}{|A_{P^0 \to f}|^2 + |A_{\overline{P}^0 \to \overline{f}}|^2}. \tag{1.52}$$

CP violation in presence of mixing

In case of a neutral meson system, the *CP* violation phenomenology is enriched by the dynamics of the flavour mixing, described in the previous Sect. 1.2.2 for a generic neutral meson system $P^0 - \overline{P}^0$. In particular we can define two types of *CP* violating phenomena:

- ***CP* violation in mixing:** it implies that the probability of the transition $P^0 \to \overline{P}^0$ at time t is different from the *CP* conjugated process $\overline{P}^0 \to P^0$. Considering Eq. (1.35a), this can occur if

$$\left|\frac{q}{p}\right| \neq 1. \tag{1.53}$$

Assuming no direct *CP* violation, this leads as well to a modification of the absolute value of λ_f from unity. The resulting asymmetry, assuming $A_{\overline{f}} = \overline{A}_f = 0$ is given by

$$\mathcal{A_{CP}}^{mix} = \frac{\Gamma(\overline{P}^0 \to f) - \Gamma(P^0 \to \overline{f})}{\Gamma(\overline{P}^0 \to f) + \Gamma(P^0 \to \overline{f})} = \frac{|\frac{p}{q} g_-(t) A_f|^2 - |\frac{q}{p} g_-(t) \overline{A}_{\overline{f}}|^2}{|\frac{p}{q} g_-(t) A_f|^2 + |\frac{q}{p} g_-(t) \overline{A}_{\overline{f}}|^2} = \frac{1 - |\frac{q}{p}|^4}{1 + |\frac{q}{p}|^4}. \tag{1.54}$$

In the B^0- B_s^0- systems there is no evidence of *CP* violation in the mixing, because $\left|\frac{q}{p}\right| \sim 1$ for both within the experimental accuracy and theoretical expectation ([24]-[25]).

- ***CP* violation in the interference between a decay with and without mixing:** this type of *CP* violation can be investigated in decays to a final state accessible to both P^0 and $\overline{P}^0$. A powerful category are *CP* eigenstates, $f = \overline{f}$, where two amplitudes can contribute to the transition amplitude from the initial state P^0 and the final state f, in particular one related to the direct decay, $A(P^0 \to f)$, and that of the decay via mixing, $A(P^0 \to \overline{P}^0 \to f)$.

In this context the *CP* symmetry is violated if

$$\Gamma(P^0_{(\rightsquigarrow \overline{P}^0)} \to f)(t) \neq \Gamma(\overline{P}^0_{(\rightsquigarrow P^0)} \to f). \tag{1.55}$$

In order to quantify this type of *CP* violation, assuming that there is no *CP* violation in mixing, the *CP* asymmetry can be defined as

$$\frac{\Gamma_{P^0(t) \to f} - \Gamma_{\overline{P}^0(t) \to f}}{\Gamma_{P^0(t) \to f} + \Gamma_{\overline{P}^0(t) \to f}} = \frac{2C_f \cos\Delta m t - 2S_f \sin\Delta m t}{2\cosh\frac{1}{2}\Delta\Gamma t + 2D_f \sinh\frac{1}{2}\Delta\Gamma t}. \tag{1.56}$$

If we also assume no direct *CP* violation, the expression above can be considerably simplified as

$$\frac{-\mathcal{I}\lambda_f \sin\Delta m t}{\cosh\frac{1}{2}\Delta\Gamma t + \mathcal{R}\lambda_f \sinh\frac{1}{2}\Delta\Gamma t}. \tag{1.57}$$

Thus we can have *CP* violation in the interference between a decay with and without mixing even if it is the only source of *CP* violation, namely when

$$\mathcal{I}(\lambda_f) \neq 0 \tag{1.58}$$

or alternatively, $\phi_f = \arg(\lambda_f) \neq \{0, \pi\}$.

1.3 The *CP*-violating phase ϕ_s in $B_s^0 \to J/\psi h^+h^-$ decays

The *CP*-violating phase ϕ_s arises in the interference between the decay of a B_s^0 meson into a *CP* eigenstate via a $b \to c\bar{c}s$ transition and the same decay after the mixing. Its importance lies in its proportionality with the unitarity angle β_s, as explained soon in more details.
At the hadron colliders, like LHC, it is of particular interest measuring ϕ_s using the decay $B_s^0 \to J/\psi(1S)\phi(1020)$, with $J/\psi(1S) \to \mu^+\mu^-$ and $\phi(1020) \to K^+K^-$, known as the golden mode for this type of measurements. In fact, this decay mode has a high branching fraction compared to the exclusive branching fractions of B_s^0 decays, namely about $3\ 10^{-5}$ [13], and it has a clear experimental signature, thanks to the presence of muons and kaons in the final state. The second most important channel is $B_s^0 \to J/\psi(1S)\pi^+\pi^-$, where $J/\psi(1S) \to \mu^+\mu^-$, with a branching fraction of about $1\ 10^{-5}$ [13]. Also in this case, the full detectable final state (without neutrinos or neutral particles) and the presence of the muons make the decay signature very clean to be observed. In addition, these channels allow to measure $\Delta\Gamma_s$, Γ_s and Γ_H, the last two in $B_s^0 \to J/\psi\phi$ and $B_s^0 \to J/\psi\pi^+\pi^-$ respectively, providing the possibility to further study some B_s^0 mixing parameters.
In this thesis, the two most recent ongoing measurements of ϕ_s at the LHCb experiment will be discussed, exploiting the decays mentioned above. In particular, in the following sections it will be presented how we can theoretically access ϕ_s using these decay modes, as well as its relation with β_s, the current status of art, the impact of so-called penguin diagrams and possible New Physics contributions to ϕ_s.

1.3.1 The *CP* violating phase ϕ_s

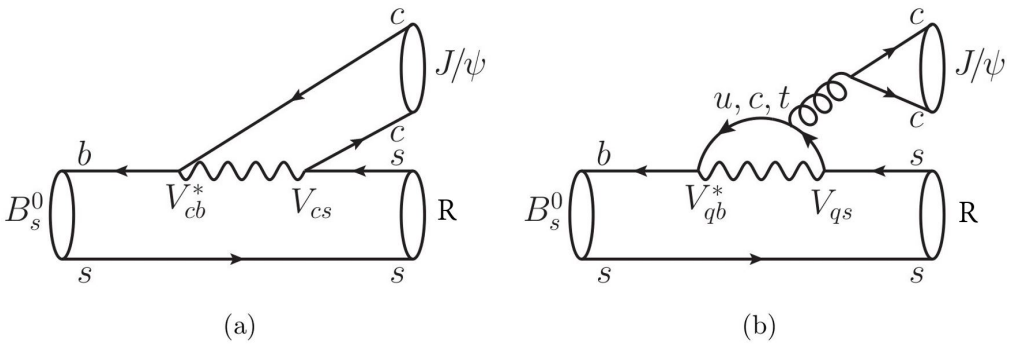

FIGURE 1.4. Feynman diagrams for the decay $B_s^0 \to J/\psi R(\to h^+h^-)$, with R a generic $s\bar{s}$ resonance. The dominant tree amplitude (a) and the higher order penguin contributions (b) are shown.

Fig. 1.4 shows the dominant and higher second order diagrams that contribute to the B_s^0 decay via a $b \to c\bar{c}s$ transition. As explained in Sec. 1.2.3, in order to fully describe the decay rate of a neutral meson we have to know what is the expression of the *CP* violating parameter

λ_f. The amplitude $A_{J/\psi hh}$, accordingly to Fig 1.4, can be expressed as follow

$$A_{J/\psi hh} = V_{cs}V_{cb}^* T + V_{us}V_{ub}^* P_u + V_{cs}V_{cb}^* P_c + V_{ts}V_{tb}^* P_t, \tag{1.59}$$

where T is the amplitude associated to the tree level decay, and P_q $(i = u, c, t)$ is referred to the so-called penguin decay topology. Considering that $V_{ts}V_{tb}^* = -V_{cs}V_{cb}^* - V_{us}V_{ub}^*$ (exploiting the unitarity triangle for the B_s^0 system), the amplitude can be written as

$$A_{J/\psi hh} = V_{cs}V_{cb}^*(T + P_c - P_t) + V_{us}V_{ub}^*(P_u - P_t) \sim V_{cs}V_{cb}^*(T + P_c - P_t) + \mathcal{O}(\lambda)^4, \tag{1.60}$$

neglecting the term proportional to $V_{us}V_{ub}^*$. In general, following what we saw in Sec 1.2.3, an amplitude can be written as a function of the weak and strong phases, ϕ_D and δ_f, as

$$A_f = |A_f| e^{i(\delta_f + \phi_D)}. \tag{1.61}$$

Remembering that the strong phase does not changed under *CP* transformation, contrary to the weak phase, the amplitude $\overline{A}_{J/\psi hh}$ can be defined as

$$\overline{A}_{J/\psi hh} = \eta_{J/\psi hh}\overline{A_{\overline{f}}} = \eta_{J/\psi hh}|A_f| e^{i(\delta_f - \phi_D)}, \tag{1.62}$$

$$\sim \eta_{J/\psi hh} V_{cs}^* V_{cb}(T + P_c - P_t) + \mathcal{O}(\lambda)^4 \tag{1.63}$$

where we are assuming no direct *CP* violation.
Consequently, the amplitude ratio $\overline{A}_f/A_f$ is given by

$$\frac{\overline{A}_{J/\psi hh}}{A_{J/\psi hh}} = -\eta_{J/\psi hh}\frac{V_{cs}^* V_{cb}}{V_{cs}V_{cb}^*} = -\eta_{J/\psi hh} e^{2i\phi_D}. \tag{1.64}$$

Let us now give a look at the q/p term. Using Eq. 1.31 and considering that for the B_s^0 system $\Gamma_{12} \ll \Delta m_{12}$, we can write it as

$$\frac{q}{p} \simeq -\sqrt{\frac{M_{12}}{M_{12}^*}}. \tag{1.65}$$

Then, the Eqs. 1.32a and 1.34 allow to obtain

$$\Delta m_{B_s^0} \sim 2|M_{12}| \propto |V_{tb}^* V_{ts}|^2, \tag{1.66}$$

obtaining, eventually, the following relation:

$$\frac{q}{p} = -\frac{V_{tb}^* V_{ts}}{V_{tb}V_{ts}^*} = -e^{-\phi_M}. \tag{1.67}$$

The combination of $\frac{\overline{A}_{J/\psi hh}}{A_{J/\psi hh}}$ and $\frac{q}{p}$ lead to:

$$\lambda_f = \eta_{J/\psi hh}\frac{V_{cs}^* V_{cb}}{V_{cs}V_{cb}^*}\frac{V_{tb}^* V_{ts}}{V_{tb}V_{ts}^*} = \eta_{J/\psi hh} e^{2i\phi_D - i\phi_M} = \eta_{J/\psi hh} e^{-i\phi_s}. \tag{1.68}$$

The phase ϕ_s, if we neglect the penguin contributions, is proportional to the angle β_s of the unitarity triangle of the B_s^0 system, namely

$$\phi_s = -2\beta_s. \tag{1.69}$$

This allows to measure in an independent way the angle β_s, parameter that is very well predicted in the SM, $\beta_s^{SM} = -0.0370^{+0.0007}_{-0.0008}$ rad [3]. This precise prediction makes the measurement of β_s via the *CP*-violating phase ϕ_s an excellent test of the *CP* violation in the SM theory.

1.3.2 Time-dependent decay rate

The time-dependent decay rates of the $B_s^0 \to J/\psi h^+h^-$ and its *CP* conjugated decay, as well as the experimental artefacts introduced in the decay rate by the detector and analysis procedure (see Chapter 3), are the core of the ϕ_s measurements described in this thesis. In terms of the expression of λ_f presented in Eq. 1.68, they can be written as

(1.70a)
$$\Gamma_{P^0\to f}(t) = |A_f|^2 e^{-\Gamma t}\frac{1}{1+C}\Big(\cosh\frac{1}{2}\Delta\Gamma t + \eta_{J/\psi hh}D_f\sinh\frac{1}{2}\Delta\Gamma t + C_f\cos\Delta m t - \eta_{J/\psi hh}S_f\sin\Delta m t\Big),$$

(1.70b)
$$\Gamma_{\overline{P}^0\to f}(t) = |A_f|^2\Big|\frac{p}{q}\Big|^2 e^{-\Gamma t}\frac{1}{1+C}\Big(\cosh\frac{1}{2}\Delta\Gamma t + \eta_{J/\psi hh}D_f\sinh\frac{1}{2}\Delta\Gamma t - C_f\cos\Delta m t + \eta_{J/\psi hh}S_f\sin\Delta m t\Big),$$

where the coefficients D and S are redefined in the following way

$$D_f = \frac{-2|\lambda_f|\cos(\phi_s)}{1+|\lambda_f|^2} \qquad C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2} \qquad S_f = \frac{2|\lambda_f|\sin(\phi_s)}{1+|\lambda_f|^2}. \tag{1.71}$$

Considering that $B_s^0 \to J/\psi(\to \mu^+\mu^-)h^+h^-$ is a four-body decay (four particles in the final state), the phase space description depends on a set of 16 kinematical variables, related to the four-momenta of the daughter particles. This number can anyway be reduced to five considering the constraints applied by the conservation of the four-momenta:

- four constraints due to the energy-momentum conservation;
- four constraints due to the known masses of the particles in the final state;
- removal of three angular variables due to the freedom to change the orientation of the B_s^0 meson in its rest frame, since it is a spinless particle.

Any set of five independent variables can be used to parametrise the phase space. In this thesis the Cabibbo-Maksymowicz (CM) parametrisation is chosen, where the phase space is fully described as a function of two invariant masses, $m(\mu^+\mu^-)$ and $m(h^+h^-)$, two polar angles and one azimuthal angle. Because of the presence of the J/ψ as the only narrow resonant structure

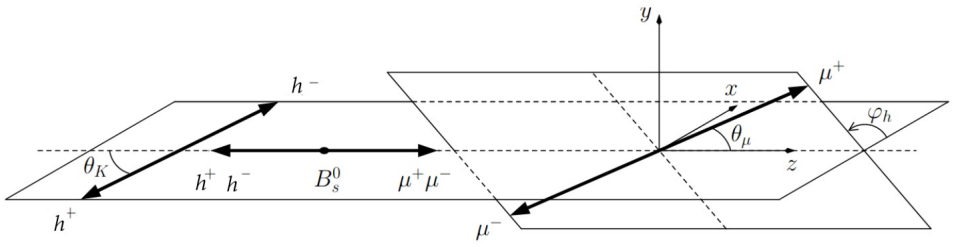

FIGURE 1.5. The three angles in the helicity formalism that describe the phase space of the $B_s^0 \to J/\psi h^+h^-$ system.

in the $\mu^+\mu^-$ system, the invariant mass of the two muons is constrained to the known mass of the J/ψ, removing another variable from the space phase parametrisation, for a total of four independent variables. The so-called helicity basis is chosen to define the three angles. Their definition is shown in Figure 1.5. We can first define the helicity axis as the direction of the $(\mu^+\mu^-)$ and (h^+h^-) momenta in the rest frame of the B_s^0 meson. Then the angle θ_μ (θ_h) is define as the angle between the positively charged muon (hadron) in the rest frame of the J/ψ ((h^+h^-) resonance) and the helicity axis. Then φ_h is defined as the angle between the direction of the negatively charged kaon and the direction of the positively charged muon. Looking at the decay rates defined in Eq. 1.70a and 1.70b, we can notice that they depend on the angular momentum l between the J/ψ and the h^+h^- resonance, through the dependence on the *CP* eigenvalue $\eta_{J/\psi hh} = (-1)^l$. Since this is strictly subordinated to the resonances in the final state, in the following paragraphs we will discuss the dependence on the orbital angular momentum, and its consequences, separately for $B_s^0 \to J/\psi\phi(\to K^+K^-)$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays. For that, it is very instructive to look at the time-dependent decay rate as a function of the angles. Just for illustrative purposes, the expression can be simplified by looking to one single angle θ [1]

$$\begin{aligned}\frac{d\Gamma}{dtd\cos(\theta)} \propto \frac{1}{2}(1+\cos^2\theta)(1-R_\perp)\times \\ \times\Big[(1+\cos\phi_s)e^{-\Gamma_L t}+(1-\cos\phi_s)e^{-\Gamma_H t}+\sin\phi_s\sin(\Delta m_s t)\Big] \\ +\sin^2\theta R_\perp\times\Big[(1-\cos\phi_s)e^{-\Gamma_L t}+(1+\cos\phi_s)e^{-\Gamma_H t}-\sin\phi_s\sin(\Delta m_s t)\Big],\end{aligned}$$

where $R_\perp$ is the fraction of the *CP*-odd component ($\eta_{J/\psi hh} = -1$) in the final state. The decay rate is the result of a combination of two exponentials, with decay widths equal to Γ_H and Γ_L. As an example, we can know consider the decay of a b-meson into a pure *CP*-odd final state, $R_\perp = 1$. Then the previous expression can be simplified as

$$\frac{d\Gamma}{dtd\cos(\theta)} \propto \sin^2\theta\times\Big[(1-\cos\phi_s)e^{-\Gamma_L t}+(1+\cos\phi_s)e^{-\Gamma_H t}-\sin\phi_s\sin(\Delta m_s t)\Big]. \tag{1.73}$$

The decay rate still depends on both mass eigenstates widths, since the mass eigenstates do not correspond to the *CP* eigenstates, because of the presence of *CP* violation. Anyway, taking into

consideration the ϕ_s very small magnitude, the *CP*-even component is highly suppressed and the sensitivity on the Γ_L determination is very low. Thus, if we want to precisely access the $\Gamma_s = (\Gamma_H + \Gamma_L)/2$, $\Delta\Gamma_s = (\Gamma_H - \Gamma_L)/2$ and Δm_s observables, decays in which the final state is largely accessible by both the *CP* components are more indicated.

Table 1.2: Definition of the coefficients a_k, b_k, c_k and d_k of the time-dependent functions h_k used in Eqs. (1.75) (1.76).

a_k	b_k	c_k	d_k
1	D_f	C_f	$-S_f$
1	D_f	C_f	$-S_f$
1	$-D_f$	C_f	S_f
$C_f \sin(\delta_\perp - \delta_{\|\|})$	$S_f \cos(\delta_\perp - \delta_{\|\|})$	$\sin(\delta_\perp - \delta_{\|\|})$	$D_f \cos(\delta_\perp - \delta_{\|\|})$
$\cos(\delta_{\|\|} - \delta_0)$	$D_f \cos(\delta_{\|\|} - \delta_0)$	$C_f \cos(\delta_{\|\|} - \delta_0)$	$-S_f \cos(\delta_{\|\|} - \delta_0)$
$C_f \sin(\delta_\perp - \delta_0)$	$S_f \cos(\delta_\perp - \delta_0)$	$\sin(\delta_\perp - \delta_0)$	$D_f \cos(\delta_\perp - \delta_0)$
1	$-D_f$	C_f	S_f
$C_f \cos(\delta_{\|\|} - \delta_S)$	$S_f \sin(\delta_{\|\|} - \delta_S)$	$\cos(\delta_{\|\|} - \delta_S)$	$D_f \sin(\delta_{\|\|} - \delta_S)$
$\sin(\delta_\perp - \delta_S)$	$-D_f \sin(\delta_\perp - \delta_S)$	$C_f \sin(\delta_\perp - \delta_S)$	$S_f \sin(\delta_\perp - \delta_S)$
$C_f \cos(\delta_0 - \delta_S)$	$S_f \sin(\delta_0 - \delta_S)$	$\cos(\delta_0 - \delta_S)$	$D_f \sin(\delta_0 - \delta_S)$

$B_s^0 \to J/\psi\phi(\to K^+K^-)$ The decay $B_s^0 \to J/\psi\phi$ is the decay of a pseudo scalar particle, so with spin equal to 0, to two vectors particles, with spin equal to 1. For the conservation of the total angular momentum, the angular momentum between the J/ψ and the ϕ mesons must be $l = 0, 1, 2$. Thus the final state $J/\psi\phi$ is a superposition of *CP*-even ($l = 0, 2$) and *CP*-odd ($l = 1$) components, according to $\eta_{J/\psi hh} = (-1)^l$. Therefore, three polarisation amplitudes contribute to the decay rate, $A_g = |A_g|e^{-i\delta_g + \phi_D}$, where the indices $g \in 0, \|, \perp$ refer to the longitudinal, transverse-parallel and transverse-perpendicular relative orientations of the linear polarisation vectors of the J/ψ and ϕ mesons, respectively. There is one additional component that contributes to the $B_s^0 \to J/\psi K^+K^-$ data analysed in this thesis. In fact, in addition to the ϕ resonance, an irreducible non-resonant component and the irreducible scalar $f_0(980)$ contribute to K^+K^- system. Since both these contributions are spinless, they are considered together as a *CP*- odd component, with amplitude A_S, called S-wave. Although it contributes minimally to the hadron system, the advantage in considering it is mainly to resolve the twofold ambiguity in the measurement of ϕ_s. In fact, if we consider only the $\phi(1020) \to K^+K^-$ contribution, the

Table 1.3: Amplitudes and related angular functions $f_k(\theta_\mu, \theta_h, \varphi_h)$ used in Eqs. (1.75) and (1.76).

A_k	f_k
$\lvert A_0\rvert^2$	$\cos^2\theta_K \sin^2\theta_\mu$
$\lvert A_{\lvert\lvert}\rvert^2$	$\frac{1}{2}\sin^2\theta_K(1-\cos^2\varphi_h \sin^2\theta_\mu)$
$\lvert A_\perp\rvert^2$	$\frac{1}{2}\sin^2\theta_K(1-\sin^2\varphi_h \sin^2\theta_\mu)$
$\lvert A_\perp A_{\lvert\lvert}\rvert$	$\sin^2\theta_K \sin^2\theta_\mu \sin\varphi_h \cos\varphi_h$
$\lvert A_0 A_{\lvert\lvert}\rvert$	$\sqrt{2}\sin\theta_K \cos\theta_K \sin\theta_\mu \cos\theta_\mu \cos\varphi_h$
$\lvert A_0 A_\perp\rvert$	$-\sqrt{2}\sin\theta_K \cos\theta_K \sin\theta_\mu \cos\theta_\mu \sin\varphi_h$
$\lvert A_S\rvert^2$	$\frac{1}{3}\sin^2\theta_\mu$
$\lvert A_S A_{\lvert\lvert}\rvert$	$\frac{2}{\sqrt{6}}\sin\theta_K \sin\theta_\mu \cos\theta_\mu \cos\varphi_h$
$\lvert A_S A_\perp\rvert$	$-\frac{2}{\sqrt{6}}\sin\theta_K \sin\theta_\mu \cos\theta_\mu \sin\varphi_h$
$\lvert A_S A_0\rvert$	$\frac{2}{\sqrt{3}}\cos\theta_K \sin^2\theta_\mu$

decay rate is invariant under the following transformation

$$(\phi_s, \Delta\Gamma_s, \delta_\perp, \delta_{||}, \delta_0) \to (\pi - \phi_s, -\Delta\Gamma_s, \pi - \delta_\perp, -\delta_{||}, -\delta_0). \tag{1.74}$$

The phase of the S-wave amplitude depends on the invariant mass of the K^+K^- system, as it will be explained in Chapter 3. This reflects the almost three body decay structure discussed above and the necessity to parametrise the decay rate also as a function of the (K^+K^-) invariant mass. This dependence of the S-wave component on the invariant mass of the K^+K^- system allows to single out one physically correct solution, resolving the ambiguity in the decay rate. Due to the presence of the S-wave component, starting from now we will refer to the $B_s^0 \to J/\psi\phi$ decay mode with $B_s^0 \to J/\psi K^+K^-$.

In order to optimally measure ϕ_s, these amplitudes have to be statistically separated. For this purpose we can look at the helicity angles defined above, that fully describe the angular distribution of the final state. Considering the four contributing amplitudes, the time-angular-dependent decay rate for a B_s^0 meson is given by the sum of ten terms, corresponding to the four squared polarisation amplitudes and their interference terms written as

$$\frac{d^4\Gamma(t)}{dt\, d\theta_K\, d\theta_\mu\, d\varphi_h} = \sum_{k=1}^{10} A_k h_k(t) f_k(\theta_K, \theta_{mu}, \varphi_h), \tag{1.75}$$

where the decay-time-dependent functions $h_k(t)$ are given as

$$h_k(t) = \frac{1}{1+C_f} e^{-\Gamma_s t}\Big(a_k \cosh\frac{\Delta\Gamma_s t}{2} + b_k \sinh\frac{\Delta\Gamma_s t}{2} + c_k \cos(\Delta m t) + d_k \sin(\Delta m t)\Big). \tag{1.76}$$

According to Eq. 1.70b, the signs of c_k and d_k should be reversed for a $\overline{B}{}^0_s$ meson. For each of the four amplitudes a common weak phase is considered, in order to use a single ϕ_s parameter to describe all polarisations. On the contrary, potentially different strong phases (δ_0, $\delta_\parallel$, $\delta_\perp$, δ_S) are taken into account, originating from different QCD interactions of the final states. Since only phase differences are observable, the four phases reduce to three phase differences that were chosen to be $\delta_\perp - \delta_0$, $\delta_\parallel - \delta_0$ and $\delta_S - \delta_0$, while the phase δ_0 is fixed to 0. The explicit expressions for the coefficients A_k, f_k, a_k, b_k, c_k and d_k are given in Table 1.2 and Table 1.3. Since $J/\psi K^+K^-$ is an admixture of *CP*-even and *CP*-odd components, this decay allows to measure, in addition to ϕ_s, Γ_s, $\Delta\Gamma_s$ and Δm_s.

$B^0_s \to J/\psi\pi^+\pi^-$ The decay $B^0_s \to J/\psi\pi^+\pi^-$ presents a complex resonant structure in the $\pi^+\pi^-$ mass and, therefore, to accurately separate the *CP* components of the final states a time-dependent amplitude analysis is necessary. All details on this approach can be found in Ref. [26].

The decay $B^0_s \to J/\psi\pi^+\pi^-$ can be treated as a resonant quasi-two-body decay in which the B^0_s decays into two intermediate states R_1 and R_2, each of which in turn decays into two final state particles. The resulting configuration of spin and angular momentum can be complicated so it could be useful to use the so-called *helicity formalism*. It allows a proper treatment of angular configurations taking into account the spin of final state and intermediate particles. The basic idea is that the decay amplitude can be factorised into two-body decay amplitudes, where for each decay a reference frame with the mother at rest can be defined

$$A_{a\to b(\to de)c(\to fg)} = A_{a\to bc} A_{b\to de} A_{c\to fg}. \tag{1.77}$$

Let's focus on the $a \to bc$ decay. We can define a coordinate system in which a is at rest and the z axis is directed as the spin quantisation axis, with J and M respectively its total angular momentum and its projection along the z axis. The two particles in the final state are created back to back in the a rest frame along a direction defined by a polar angle θ and an azimuthal angle φ. Taking advantage from the rotational invariance of helicities, the direction of the z axis of the new coordinate system of this frame, z', can be chosen to be aligned with the emitted particles. This two body state has the same total angular momentum J, due to its conservation, but in general the angular momentum projection M' is changed. In particular, since z' is directed along the momentum of the particles in the final state, the projections coincide with the b and c helicities, λ_b and λ_c, leading to a total projection of $M' = \lambda_b - \lambda_c$. Then, the amplitude can be expressed as

$$A_{a\to bc} = \sqrt{\frac{2J+1}{4\pi}}\, d^J_{M_a,\lambda_b-\lambda_c}(\theta,0) e^{i\lambda\varphi} \mathcal{A}_{\lambda_b,\lambda_c}, \tag{1.78}$$

where $\mathcal{A}$ is a dynamical factor which takes into account the quantum couplings to those specific helicities and contains all the QCD dynamics and the non-perturbative effects. The Wigner d functions are, instead, related to the probability for a state with angular momentum J of passing from the angular projection M to $M^{'}$ after a rotation.
A similar procedure can be used to build the amplitudes $A_{b\to de}$ and $A_{c\to fg}$, leading to the four-body decay total amplitude multiplying the single terms, according to Eq. 1.77.
As explained in more details in Ref. [27], the angular basis resulting from this formalism is equivalent to the helicity basis defined in Fig. 1.5.
The formalism described above can be used to write the decay rate for the $B_s^0 \to J/\psi\pi^+\pi^-$ analysis. Following the discussion in Section 1.2.2, it can generally be written as

(1.79)
$$\Gamma(t) = \mathcal{N}e^{-\Gamma_s t}\Big(\frac{|A|^2+|\overline{A}|^2}{2}\cosh\frac{\Delta\Gamma t}{2}+\frac{|A|^2-|\overline{A}|^2}{2}\cos\Delta mt-\mathcal{R}(A^*\overline{A})\sinh\frac{\Delta\Gamma t}{2}-\mathcal{I}(A^*\overline{A})\sin\Delta mt\Big),$$

where $\mathcal{N}$ is a normalisation term, $|A| = |\langle J/\psi\pi^+\pi^-|\mathcal{H}|B_s^0\rangle|$ and $|\overline{A}| = |\langle J/\psi\pi^+\pi^-|\mathcal{H}|\overline{B}_s^0\rangle|$.

The amplitudes can be now written using the helicity formalism

(1.80)
$$|A_f(m_{hh},\theta_h,\theta_l,\varphi)|^2 = \sum_{\alpha=\pm1}\Big|\sum_{\lambda,J}^{|\lambda|\leq J}\sqrt{\frac{2J+1}{4\pi}}\mathcal{H}_\lambda^J(m_{hh})e^{i\lambda\varphi}d^1_{\lambda,\alpha}(\theta_l)d^J_{-\lambda,0}(\theta_h)\Big|^2,$$

where $\lambda = 0,\pm1$ is the J/ψ helicity, $\alpha = \pm1$ is the helicity difference between the two muons, J is the spin of the $\pi^+\pi^-$ intermediate state, d is the part of the Wigner matrices that depends only on the polar angle and $\mathcal{H}_\lambda^J(m_{hh})$ is instead the dynamical factor containing the lineshape of the distribution of the $m(hh)$ spectrum. Since many intermediate states contribute to the $\pi^+\pi^-$ mass spectrum of the final $B_s^0 \to J/\psi\pi^+\pi^-$ four-body decay, all the possible interfering contributions are summed coherently, while all the processes which can be discriminated by a final state measurement (depending in this case on the helicities of the muons) have to be summed incoherently. The list of resonances used to describe the $\pi^+\pi^-$ mass spectrum in this analysis [6, 28] is presented in Table 1.4. In addition, a non-resonant component contributing to the $\pi^+\pi^-$ mass spectrum is also considered. As explained in Ref. [28], the largest component of the decay is the $f_0(980)$ resonance with the $f_0(1500)$ being almost an order of magnitude smaller. Because of the conservation of the total angular momentum, the J/ψ and $\pi^+\pi^-$ system have a relative angular momentum $l = 1,2,3$, depending on the spin of the $\pi^+\pi^-$ resonance. Ref. [28]

Table 1.4: Parameters of the resonances contributing to the $\pi^+\pi^-$ mass spectrum.

Resonance	Spin	Helicity	Resonance Formalism	Mass (MeV)	Width (MeV)	Source
$f_0(980)$	0	0	Flatté	Varied in the analysis		
$f_0(1500)$	0	0	BW	Varied in the analysis		
$f_0(1790)$	0	0	BW	1790^{+40}_{-30}	270^{+60}_{-30}	BES [29]
$f_2(1270)$	2	$0,\pm1$	BW	1275.5 ± 0.8	$186.7^{+2.2}_{-2.5}$	PDG [13]
$f_2'(1525)$	2	$0,\pm1$	BW	1522.2 ± 1.7	78.0 ± 4.8	LHCb [30]

shows that the *CP* content of the decay is consistent with being purely *CP*-odd ($l = 1, 3$), with the *CP*-even component ($l = 2$) limited to 2.3% at 95% CL. In this channel, the measurement of Γ_s would thus be affected by the lack of sensitivity on the Γ_L determination. For that reason, fixing Γ_L to the known value, the measurement of Γ_H can be instead performed, allowing a futher characterisation of the $B_s^0 - \overline{B}_s^0$ mixing phenomena.

Table 1.5: Definition of the coefficients $\Theta_{\lambda',\lambda}(\theta_l)$ as a function of the different J/ψ polarisation conditions.

λ	λ'	$\Theta_{\lambda',\lambda}(\theta_l)$
0	0	$sin^2\theta_l$
0	1	$\frac{1}{\sqrt{2}} sin\theta_l cos\theta_l$
0	-1	$-\frac{1}{\sqrt{2}} sin\theta_l cos\theta_l$
1	0	$\frac{1}{\sqrt{2}} sin\theta_l cos\theta_l$
1	1	$\frac{1}{2}(1 + cos^2\theta_l)$
1	-1	$\frac{1}{2} sin^2\theta_l$
-1	0	$-\frac{1}{\sqrt{2}} sin\theta_l cos\theta_l$
-1	1	$\frac{1}{2} sin^2\theta_l$
-1	-1	$\frac{1}{2}(1 + cos^2\theta_l)$

Defining:

$$\mathcal{H}_\lambda(m_{\pi^+\pi^-}, \theta_h) = \sum_J \sqrt{\frac{2J+1}{4\pi}} \mathcal{H}_\lambda^J(m_{hh}) d^J_{-\lambda,0}(\theta_h) \tag{1.81}$$

and

$$\Theta_{\lambda',\lambda'}(\theta_l) = \sum_{\alpha=\pm1} d^1_{\lambda',\alpha}(\theta_l) d^1_{\lambda,\alpha}(\theta_l), \tag{1.82}$$

the Eq. 1.80 results in

$$|A_f(m_{\pi^+\pi^-}, \theta_l, \theta_h, \varphi_h)|^2 = \sum_{\lambda',\lambda} \mathcal{H}_\lambda(m_{\pi^+\pi^-}, \theta_h) \mathcal{H}^*_{\lambda'}(m_{\pi^+\pi^-}, \theta_h) e^{i(\lambda-\lambda')\varphi_h} \Theta_{\lambda',\lambda}(\theta_l). \tag{1.83}$$

Table 1.5 lists the $\Theta_{\lambda',\lambda}(\theta_l)$ functions in the different J/ψ polarisation conditions.

The amplitude for $\overline{B}_s^0 \to J/\psi\pi^+\pi^-$ can be defined as:

$$|A_f(m_{\pi^+\pi^-},\theta_l,\theta_h,\varphi_h)|^2 = \sum_{\lambda',\lambda} \overline{\mathcal{H}}_\lambda(m_{\pi^+\pi^-},\theta_h)\overline{\mathcal{H}}^*_{\lambda'}(m_{\pi^+\pi^-},\theta_h)e^{i(\lambda-\lambda')\varphi_h}\Theta_{\lambda',\lambda}(\theta_l), \tag{1.84}$$

while the interference term can be written as

$$|A_f(m_{\pi^+\pi^-},\theta_l,\theta_h,\varphi_h)|^2 = \sum_{\lambda',\lambda} \overline{\mathcal{H}}_\lambda(m_{\pi^+\pi^-},\theta_h)\mathcal{H}^*_{\lambda'}(m_{\pi^+\pi^-},\theta_h)e^{i(\lambda-\lambda')\varphi_h}\Theta_{\lambda',\lambda}(\theta_l). \tag{1.85}$$

The term, $\mathcal{H}_\lambda(m_{\pi^+\pi^-},\theta_h)$, combines the amplitude $A_R(m_{\pi^+\pi^-})$, which describes the mass lineshape of the resonance R, in most cases is a Breit-Wigner function, with the B_s^0 resonance decay properties to form the expression

(1.86)

$$\mathcal{H}_\lambda(m_{\pi^+\pi^-},\theta_h) = h_\lambda^R\sqrt{2J_R+1}\sqrt{P_R P_B}F_B^{(L_B)}F_R^{(L_R)}A_R(m_{\pi^+\pi^-})\left(\frac{P_B}{m_B}\right)^{L_B}\left(\frac{P_R}{m_0}\right)^{L_R} d^J_{-\lambda,0}(\theta_h).$$

Here P_B is the J/ψ momentum in the B_s^0 rest frame, P_R is the momentum of either of the two hadrons in the dihadron rest frame, m_B is the B_s^0 mass, m_0 is the central mass of resonance R, J_R is the spin of R, L_B is the orbital angular momentum between the J/ψ and $\pi^+\pi^-$ system, and L_R the orbital angular momentum in the $\pi^+\pi^-$ decay, and thus is the same as the spin of the $\pi^+\pi^-$ resonance. $F_B^{(L_B)}$ and $F_R^{(L_R)}$ are the Blatt-Weisskopf barrier factors for B_s^0 and R resonance, respectively [31].

1.3.3 Impact of penguin diagram contribution

As mentioned in Section 1.3.1, the predominant diagram in $B_s^0 \to J/\psi h^+h^-$ decays via $b \to c\bar{c}s$ transitions is due to the tree level process, allowing the comparison between the measured observable ϕ_s and the theoretical prediction of $-2\beta_s$. However, this comparison is valid only if terms proportional to $V_{us}V_{ub}^*$, associated to doubly Cabibbo-suppressed penguin topologies, can be neglected in Eq. 1.60.

Indeed, the quantity actually measured is

$$\phi_s = -2\beta_s + \Delta\phi_s^{peng.}, \tag{1.87}$$

where $\Delta\phi_s^{peng.}$ is the contribution to ϕ_s due to penguin diagrams only.

For this reason, considering the increasing experimental precision of ϕ_s measurements, the impact due to penguin diagrams have to be determined. This is crucial, otherwise any deviation from the SM prediction of $-2\beta_s$ caused by penguin pollution could be wrongly interpreted as an evidence of NP contributions. Unfortunately, the penguin contributions are difficult to calculate, due to the non-perturbative nature of the QCD processes involved. However, in Refs. [32] and [33] a very interesting strategy to put constraints on $\Delta\phi_s^{peng.}$ has been proposed, which makes use of measurements of direct *CP* violation and branching fractions in related decay channels.

In particular, of special interest are the decays $B_s^0 \to J/\psi K^*(892)^0$ [34] and $B^0 \to J/\psi\rho^0$ [35]. In order to understand the reason behind the choice of these decay modes we have to look at their tree penguin diagrams. Fig. 1.6 shows, as an example, the tree and penguin diagrams

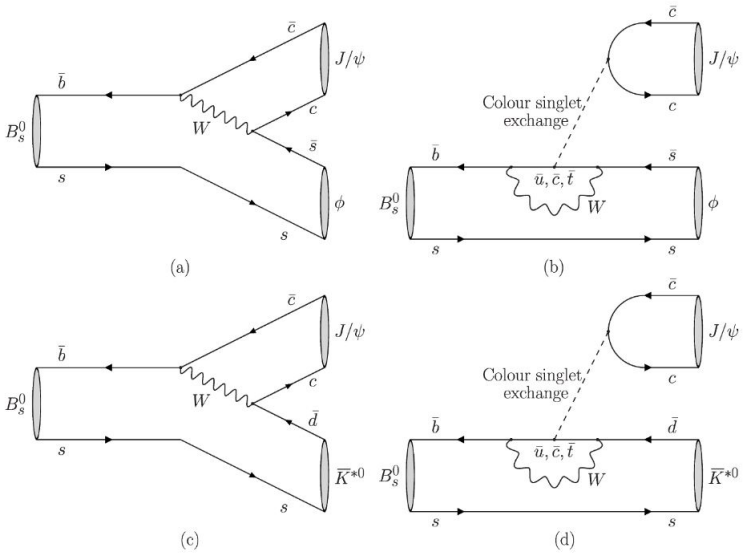

FIGURE 1.6. Decay topologies contributing to the $B_s^0 \to J/\psi\phi$ channel (a, b) and $B_s^0 \to J/\psi K^{*0}$ channel (c, d). The tree diagrams (a, c) are shown on the left and the penguin diagrams (b, d) on the right.

contributing to both $B_s^0 \to J/\psi\phi$ and $B_s^0 \to J/\psi K^{*0}$ decays. In the last mode, tree level and penguin amplitudes have the same order of magnitude, allowing the measurement of the penguin contribution, condition occurring also for $B^0 \to J/\psi\rho^0$ decay. Studying with more attention the diagrams, we can see that the first and the second orders in the penguin diagrams (that dominate in the creation of the $\Delta\phi_s^{peng.}$ shift) are due to an exchange of a c and u virtual quark for the $B_s^0 \to J/\psi\phi$ mode and of an u and c virtual quark for $B_s^0 \to J/\psi K^{*0}$, respectively. Thus, each diagram of the same order has the same magnitude for both the decay modes, even if involving different quarks. Equivalent considerations can be made for $B^0 \to J/\psi\rho^0$. Assuming SU(3) flavour symmetry, $B^0 \to J/\psi\rho^0$ and $B_s^0 \to J/\psi K^{*0}$ channels represent thus the counterparts of $B_s^0 \to J/\psi h^+h^-$ decays ($h = K, \pi$). This allow to consider the $\Delta\phi_s^{peng.}$ measured in these decays equal to the shift due to penguin diagrams in $B_s^0 \to J/\psi h^+h^-$ decays. The two analyses combined results, as a function of the polarisation of the final state, are:

$$\Delta\phi_{s,0}^{peng.} = 0.000^{+0.009}_{-0.011}\ (\text{stat})\ ^{+0.004}_{-0.009}\ (\text{syst})\ \text{rad},$$
$$\Delta\phi_{s,\|}^{peng.} = 0.001^{+0.010}_{-0.014}\ (\text{stat})\ \pm 0.008\ (\text{syst})\ \text{rad},$$
$$\Delta\phi_{s,\perp}^{peng.} = 0.003^{+0.010}_{-0.014}\ (\text{stat})\ \pm 0.008\ (\text{syst})\ \text{rad}.$$

These results show that penguin pollution in the measurement of ϕ_s is close to 0 within a precision of ~ 10 mrad, therefore so far any major deviation from $-2\beta_s$ can be interpreted as

NP effects. However, with the increase of the experimental accuracy of the ϕ_s measurement, the penguin contribution will need to be considered and an update of these results and potentially the inclusion of penguin contributions in the ϕ_s analysis will be necessary.

1.3.4 Polarisation-dependent decay rate

So far we have always made the assumption that $\lambda_f \equiv \eta_f \frac{q}{p} \frac{\bar{A}_f}{A_f}$ is independent of the polarisation of the final state f, which implies that the *CP*-violation is the same for all polarisation states. However, penguin pollution magnitude could be different as a function of the polarisation, as well as possible contributions due to NP phenomena. With the increasing of the amount of data, the measurement of different λ_f for each polarisation state is becoming a real possibility. Therefore, in addition to the baseline measurement with one common λ, the measurement of different λ_f for each polarisation state in the analyses presented in this thesis is under study, in order to check the assumption of no polarisation dependence. The expressions for the coefficients a_k, b_k, c_k and d_k for the polarisation-dependent case are given in Table 1.6.

Table 1.6: The angular and time-dependent functions used in Eqs. (1.75) and (1.76), for a polarisation-dependent ϕ_s and $|\lambda|$. Abbreviations for cosine and sine functions are used: $c_K = \cos\theta_K$, $s_K = \sin\theta_K$, $c_l = \cos\theta_l$, $s_l = \sin\theta_l$, $c_\phi = \cos\phi$, $s_\phi = \sin\phi$.

a_k	b_k	c_k	d_k
$\frac{1}{2}(1+\lvert\lambda_0\rvert^2)$	$-\lvert\lambda_0\rvert\cos(\phi_0)$	$\frac{1}{2}(1-\lvert\lambda_0\rvert^2)$	$\lvert\lambda_0\rvert\sin(\phi_0)$
$\frac{1}{2}(1+\lvert\lambda_{\lvert\rvert}\rvert^2)$	$-\lvert\lambda_{\lvert\rvert}\rvert\cos(\phi_{\lvert\rvert})$	$\frac{1}{2}(1-\lvert\lambda_{\lvert\rvert}\rvert^2)$	$\lvert\lambda_{\lvert\rvert}\rvert\sin(\phi_{\lvert\rvert})$
$\frac{1}{2}(1+\lvert\lambda_\perp\rvert^2)$	$\lvert\lambda_\perp\rvert\cos(\phi_\perp)$	$\frac{1}{2}(1-\lvert\lambda_\perp\rvert^2)$	$-\lvert\lambda_\perp\rvert\sin(\phi_\perp)$
$\frac{1}{2}\Big[\sin(\delta_\perp-\delta_{\lvert\rvert})-\lvert\lambda_\perp\lambda_{\lvert\rvert}\rvert$ $\sin(\delta_\perp-\delta_{\lvert\rvert}-\phi_\perp+\phi_{\lvert\rvert})\Big]$	$\frac{1}{2}\Big[\lvert\lambda_\perp\rvert\sin(\delta_\perp-\delta_{\lvert\rvert}-\phi_\perp)$ $+\lvert\lambda_{\lvert\rvert}\rvert\sin(\delta_{\lvert\rvert}-\delta_\perp-\phi_{\lvert\rvert})\Big]$	$\frac{1}{2}\Big[\sin(\delta_\perp-\delta_{\lvert\rvert})+\lvert\lambda_\perp\lambda_{\lvert\rvert}\rvert$ $\sin(\delta_\perp-\delta_{\lvert\rvert}-\phi_\perp+\phi_{\lvert\rvert})\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_\perp\rvert\cos(\delta_\perp-\delta_{\lvert\rvert}-\phi_\perp)$ $+\lvert\lambda_{\lvert\rvert}\rvert\cos(\delta_{\lvert\rvert}-\delta_\perp-\phi_{\lvert\rvert})\Big]$
$\frac{1}{2}\Big[\cos(\delta_0-\delta_{\lvert\rvert})+\lvert\lambda_0\lambda_{\lvert\rvert}\rvert$ $\cos(\delta_0-\delta_{\lvert\rvert}-\phi_0+\phi_{\lvert\rvert})\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_0\rvert\cos(\delta_0-\delta_{\lvert\rvert}-\phi_0)$ $+\lvert\lambda_{\lvert\rvert}\rvert\cos(\delta_{\lvert\rvert}-\delta_0-\phi_{\lvert\rvert})\Big]$	$\frac{1}{2}\Big[\cos(\delta_0-\delta_{\lvert\rvert})-\lvert\lambda_0\lambda_{\lvert\rvert}\rvert$ $\cos(\delta_0-\delta_{\lvert\rvert}-\phi_0+\phi_{\lvert\rvert})\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_0\rvert\sin(\delta_0-\delta_{\lvert\rvert}-\phi_0)$ $+\lvert\lambda_{\lvert\rvert}\rvert\sin(\delta_{\lvert\rvert}-\delta_0-\phi_{\lvert\rvert})\Big]$
$-\frac{1}{2}\Big[\sin(\delta_0-\delta_\perp)-\lvert\lambda_0\lambda_\perp\rvert$ $\sin(\delta_0-\delta_\perp-\phi_0+\phi_\perp)\Big]$	$\frac{1}{2}\Big[\lvert\lambda_0\rvert\sin(\delta_0-\delta_\perp-\phi_0)$ $+\lvert\lambda_\perp\rvert\sin(\delta_\perp-\delta_0-\phi_\perp)\Big]$	$-\frac{1}{2}\Big[\sin(\delta_0-\delta_\perp)+\lvert\lambda_0\lambda_\perp\rvert$ $\sin(\delta_0-\delta_\perp-\phi_0+\phi_\perp)\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_0\rvert\cos(\delta_0-\delta_\perp-\phi_0)$ $+\lvert\lambda_\perp\rvert\cos(\delta_\perp-\delta_0-\phi_\perp)\Big]$
$\frac{1}{2}(1+\lvert\lambda_S\rvert^2)$	$\lvert\lambda_S\rvert\cos(\phi_S)$	$\frac{1}{2}(1-\lvert\lambda_S\rvert^2)$	$-\lvert\lambda_S\rvert\sin(\phi_S)$
$\frac{1}{2}\Big[\cos(\delta_S-\delta_{\lvert\rvert})-\lvert\lambda_S\lambda_{\lvert\rvert}\rvert$ $\cos(\delta_S-\delta_{\lvert\rvert}-\phi_S+\phi_{\lvert\rvert})\Big]$	$\frac{1}{2}\Big[\lvert\lambda_S\rvert\cos(\delta_S-\delta_{\lvert\rvert}-\phi_S)$ $-\lvert\lambda_{\lvert\rvert}\rvert\cos(\delta_{\lvert\rvert}-\delta_S-\phi_{\lvert\rvert})\Big]$	$\frac{1}{2}\Big[\cos(\delta_S-\delta_{\lvert\rvert})+\lvert\lambda_S\lambda_{\lvert\rvert}\rvert$ $\cos(\delta_S-\delta_{\lvert\rvert}-\phi_S+\phi_{\lvert\rvert})\Big]$	$\frac{1}{2}\Big[\lvert\lambda_S\rvert\sin(\delta_S-\delta_{\lvert\rvert}-\phi_S)$ $-\lvert\lambda_{\lvert\rvert}\rvert\sin(\delta_{\lvert\rvert}-\delta_S-\phi_{\lvert\rvert})\Big]$
$-\frac{1}{2}\Big[\sin(\delta_S-\delta_\perp)+\lvert\lambda_S\lambda_{sec:phis\perp}\rvert$ $\sin(\delta_S-\delta_\perp-\phi_S+\phi_\perp)\Big]$	$-\frac{1}{2}\Big[\lvert\lambda_S\rvert\sin(\delta_S-\delta_\perp-\phi_S)$ $-\lvert\lambda_\perp\rvert\sin(\delta_\perp-\delta_S-\phi_\perp)\Big]$	$-\frac{1}{2}\Big[\sin(\delta_S-\delta_\perp)-\lvert\lambda_S\lambda_\perp\rvert$ $\sin(\delta_S-\delta_\perp-\phi_S+\phi_\perp)\Big]$	$-\frac{1}{2}\Big[-\lvert\lambda_S\rvert\cos(\delta_S-\delta_\perp-\phi_S$ $+\lvert\lambda_\perp\rvert\cos(\delta_\perp-\delta_S-\phi_\perp)\Big]$
$\frac{1}{2}\Big[\cos(\delta_S-\delta_0)-\lvert\lambda_S\lambda_0\rvert$ $\cos(\delta_S-\delta_0-\phi_S+\phi_0)\Big]$	$\frac{1}{2}\Big[\lvert\lambda_S\rvert\cos(\delta_S-\delta_0-\phi_S)$ $-\lvert\lambda_0\rvert\cos(\delta_0-\delta_S-\phi_0)\Big]$	$\frac{1}{2}\Big[\cos(\delta_S-\delta_0)+\lvert\lambda_S\lambda_0\rvert$ $\cos(\delta_S-\delta_0-\phi_S+\phi_0)\Big]$	$\frac{1}{2}\Big[\lvert\lambda_S\rvert\sin(\delta_S-\delta_0-\phi_S)$ $-\lvert\lambda_0\rvert\sin(\delta_0-\delta_S-\phi_0)\Big]$

1.4 The mixing parameter $\Delta\Gamma_d$

In order to test the SM theory and to study the potential presence of new physics phenomena, a useful class of observables are those whose predictions are vanishingly small compared to the experimental sensitivity. In fact, the measurement of a value very different from zero for these parameters leads directly to the observation of NP effects. The search for *CP* violation in neutral B^0 meson mixing itself falls into this class of observables, where the best null-test parameter is $\Delta\Gamma_d$, predicted to be very small [4]:

$$\Delta\Gamma_d^{SM} = (2.6 \pm 0.4) \times 10^{-3}\ \text{ps}^{-1}. \tag{1.89}$$

In order to examine how $\Delta\Gamma_d$ may be determined experimentally, let's consider the time-dependent decay rates of an initial $B^0(\overline{B}^0)$ meson to a final state f

$$\Gamma_{B^0\to f}(t) = |A_f|^2 e^{-\Gamma_d t}\frac{1}{1+C}\Big(\cosh\frac{1}{2}\Delta\Gamma t + D_f\sinh\frac{1}{2}\Delta\Gamma t + C_f\cos\Delta m t - S_f\sin\Delta m t\Big), \tag{1.90}$$

$$\Gamma_{\overline{B}^0\to f}(t) = |A_f|^2\Big|\frac{p}{q}\Big|^2 e^{-\Gamma_d t}\frac{1}{1+C}\Big(\cosh\frac{1}{2}\Delta\Gamma t + D_f\sinh\frac{1}{2}\Delta\Gamma t - C_f\cos\Delta m t + S_f\sin\Delta m t\Big). \tag{1.91}$$

Assuming that B^0 and $\overline{B}^0$ are produced in equal numbers, the total untagged* decay rate becomes

$$\Gamma_{B^0(\overline{B}^0)\to f}(t) = \Big[e^{-\Gamma_L t}|\langle f|B_L\rangle|^2 + e^{-\Gamma_H t}|\langle f|B_H\rangle|^2\Big] = \tag{1.92a}$$

$$|A_f|^2 e^{-\Gamma_d t}\frac{1}{1+C}\Big(\cosh\frac{1}{2}\Delta\Gamma t + D_f\sinh\frac{1}{2}\Delta\Gamma t\Big), \tag{1.92b}$$

where the terms depending on $\Delta m t$ cancel in the sum. Because of the mixing phenomenon, the above expression is governed by two exponential functions, with lifetime equal to Γ_H and Γ_L, respectively. The decay rate can be anyway described with a single exponential function, with lifetime equal to the so-called *effective lifetime* τ^{eff}, expressed as

$$\tau^{eff} = \frac{\int_0^\infty t\Gamma_{B\to f}(t)\,dt}{\int_0^\infty \Gamma_{B\to f}(t)\,dt}. \tag{1.93}$$

Considering that substitution $y_d = \Delta\Gamma_d/2\Gamma_d \ll 0$, the effective lifetime expression can be simplified as [36]

$$\tau_{eff} = \frac{1}{\Gamma_d}\frac{1}{1-y_d^2}\Big(\frac{1+2D_f y_d + y_d^2}{1+D_f y_d}\Big). \tag{1.94}$$

As a function of the final state, the decay occurs with different combinations of the two mass eigenstates (D_f depends on the final state) and hence, the effective lifetime varies. Noting that $e^{X(1+\epsilon)} \sim e^X(1+X\epsilon)$, any attempt to measure y_d from a single decay mode will suffer from large systematic uncertainties (ϵ) on the decay time distribution. Therefore the most suitable decay modes must have large branching fractions, clear signature in the detector and large, well-known values of D_f. These considerations suggest to use decays via $b\to c\bar{c}s$ transitions, with J/ψ mesons in the final state.

Following what is suggested in Ref. [37], as well the considerations made above, y_d can be determined from the difference in decay rates between decays to *CP*-eigenstates and decays to flavour-specific final states. In this thesis, we use as decay to a *CP*-eigenstate $B_d^0\to J/\psi K_s^0$ and as a decay to a flavour-specific final state $B_d^0\to J/\psi K^*(892)^0$. In the next paragraphs, the decay rates of the two decay modes named above are presented.

*The untagged term indicates a sample composed from both the B_s^0 and $\overline{B}_s^0$ decays, where the information about the initial flavour of the meson can not be inferred by looking at the final states. Otherwise, the decay rate is known as tagged. More details on this topic are given in Chap. 3

$B_d^0 \to J/\psi K^*(892)^0$ The decay $B_d^0 \to J/\psi K^*(892)^0$, with $K^*(892)^0 \to K^+\pi^-$, is the decay of a B_d^0 meson to a flavour-specific final state. It means that the final state is only accessible by the decay of the B^0 meson, becoming therefore reachable for a $\overline{B}^0$ only through oscillation. For these decays $\overline{A}_f = A_{\overline{f}} = 0$, resulting in $D_f = 0$. So, the total decay rate can be expressed as:

$$\tau^{eff}_{B_d^0 \to J/\psi K^*(892)^0} = \frac{1}{\Gamma_d}\frac{1}{1-y_d^2}(1+y_d^2). \tag{1.95}$$

$B_d^0 \to J/\psi K_s^0$ The decay $B_d^0 \to J/\psi K_s^0$, with $K_s^0 \to \pi + \pi^-$, is the decay of a B_d^0 meson to a *CP* eigenstate. In order to know the expression for D_f, the *CP*-violating parameter λ_f has to be calculated. Following the same approach used in Section 1.3.1, λ_f can be written as:

$$\lambda_f = \Big(\frac{q}{p}\Big)_{B^0}\Big(\eta_{J/\psi K_s^0}\frac{A_{\overline{J/\psi K_s^0}}}{A_{J/\psi K_s^0}}\Big). \tag{1.96}$$

For the conservation of the total angular momentum, considering that B^0 and K_s^0 are spin-0 particles while the $J\psi$ has spin equal to 1, the resonances in the final state have a relative angular momentum $l = 1$, resulting in:

$$\eta_{J/\psi K_s^0} = (-1)^l = -1. \tag{1.97}$$

Consequently, λ_f is defined as:

$$\lambda_f = -\Big(\frac{q}{p}\Big)_{B^0}\Big(\frac{\overline{A}_{J/\psi\overline{K}^0}}{A_{J/\psi K^0}}\Big)\Big(\frac{p}{q}\Big)_{K^0} = -\Big(\frac{V_{tb}^*V_{td}}{V_{cb}^*V_{cd}}\frac{V_{tb}V_{td}^*}{V_{cb}V_{cd}^*}\Big), \tag{1.98}$$

and the following form for D_f can be derived:

$$D_f = \cos(2\beta), \tag{1.99}$$

where the β definition is given in Eq. 1.24. So, the total decay rate can be expressed as:

$$\tau^{eff}_{B_d^0 \to J/\psi K_s^0} = \frac{1}{\Gamma_d}\frac{1}{1-y_d^2}\Big(\frac{1+2\cos(2\beta)y_d+y_d^2}{1+\cos(2\beta)y_d}\Big). \tag{1.100}$$

Determination of $\Delta\Gamma_d$ A measurement of the $\Delta\Gamma_d$ parameter is possible, describing the theoretical decay rates of the two modes as a single exponential function with lifetime equal to Eqs. 1.95 and 1.100, respectively. Two major methods can be used. The first consists to perform a binned ratio R of the two untagged decay rates, described as

$$R = e^{-t/\left(\tau^{eff}_{B_d^0\to J/\psi K_s^0} - \tau^{eff}_{B_d^0 \to J/\psi K^*(892)^0}\right)} = e^{-t\dfrac{(1+\cos(2\beta)y_d)(1+y_d^2)}{1+2\cos(2\beta)y_d+y_d^2}}. \tag{1.101}$$

The second exploits a simultaneous fit of the two decay rates. More details about the methods can be found in Chap. 6. The strategy to use two channels has very important advantages from

an experimental point of view. In fact, the two decay modes are topologically similar, so many systematic uncertainties related to the reconstruction of the particles and of the vertex of the decay are the same. Considering that the measurement is mainly sensitive to the differences between the two decay modes, this allows to have a less important impact on the measurement due to these common effects.
However, one possible systematic effect that has to be taken into account is the potential asymmetry, $A_P(B^0)$, between the production rates of B^0 and $\overline{B}^0$. If non zero, the cancellation of terms depending to Δmt in the sum of Eqs. 1.90 and 1.91 would not be exact, then the untagged decay rate of 1.92 would include an additional factor proportional to:

$$A_P(B^0)\Big(C_f\cos(\Delta mt) - S_f\sin(\Delta mt)\Big). \tag{1.102}$$

1.5 Possible New Physics contributions to B mixing

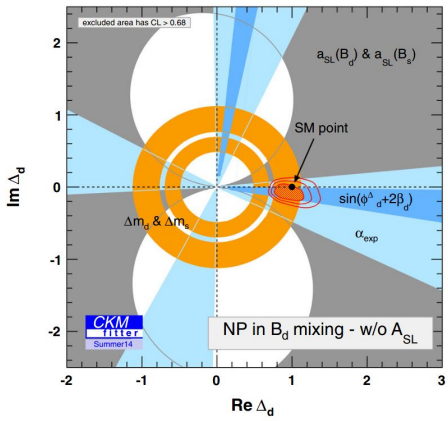

FIGURE 1.7. Global fit of possible New Physics contributions to $B^0_d - \overline{B}^0_d$ meson mixing [38].

The NP contributions to the $B^0_q - \overline{B}^0_q$ system, with $q = d, s$, can be parametrised in the form of two complex quantities Δ_q and Λ_q:

$$M^q_{12} = M^{q,SM}_{12}|\Delta_q|e^{i\phi^{\Delta}_q} \qquad \Gamma^q_{12} = \Gamma^{q,SM}_{12}|\Lambda_q|e^{i\phi^{\Lambda}_q}, \tag{1.103}$$

with four real degrees of freedom. We can write ϕ_s and $\Delta\Gamma_q$ as a function of NP parameters as:

$$\phi_s = -2\beta_s + \phi^{\Delta}_s - \delta_s \qquad \Delta\Gamma_q = (\Delta\Gamma_q)_{SM}|\Lambda_q|\frac{\cos\Big(\phi^{q,SM}_{12} + \phi^{\Delta}_q - \phi^{\Lambda}_q\Big)}{\cos(\phi^{q,SM}_{12})}, \tag{1.104}$$

where δ_s is the shift of ϕ_s due to either SM penguin diagrams or NP contributions in the decay process. While the NP contributions to ϕ_s are due only to contributions on the M_{12} parameters, the $\Delta\Gamma_q$ can change because of NP contributions in both M_{12} and Γ_{12}. We can have NP only in M_{12} simply considering $\Delta F = 2$ processes. In this context, because of the

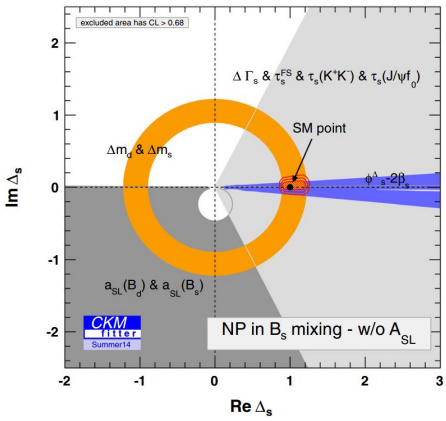

FIGURE 1.8. Global fit of possible New Physics contributions to $B_s^0 - \overline{B}_s^0$ meson mixing [38].

so-called Grossman bound, $\Delta\Gamma_q$ can only be decreased by NP contributions. On the contrary NP in $\Delta F = 1$ processes can contribute positively to the absorptive part Γ_{12}^q, because both B_q and $\overline{B}_q$ can decay to a final state through flavour dependent new physics interactions.
The scenario of $\Delta F = 2$ NP processes has been studied in extensions of the CKM fit of the SM ([38]). Figs. 1.7 and 1.8 show that the SM-like scenario is very likely but data still allow sizeable NP contributions in both B_d^0 and B_s^0 sectors up to $30\% - 40\%$ at the 3σ level, representing a strong motivation for a more precise determination of B_d^0 and B_s^0 mixing parameters. The boundaries on Δ_s are dominated by constraints due to the Δm_s and ϕ_s measurements. This last parameter is so far dominated by statistical uncertainty, resulting in the importance of further determinations of ϕ_s.
Large values of $\Delta\Gamma_q$ are possible only in a special class of models that have flavour dependent couplings and new light particles in the $B_q - \overline{B}_q$ mixing box diagram ([39]). In particular, the third generation leptoquark model predicts a large enhancement of $\Delta\Gamma_q/\Gamma_q$ over the SM range. The value of $\Delta\Gamma_s/\Gamma_s$ can be as high as 0.51, which is restricted by experimental 95% C.L. bounds from direct measurements. So improvements in the $\Delta\Gamma_s/\Gamma_s$ measurements will hence either detect new physics, or will bound the coupling constant of the model. Finally, the value of $\Delta\Gamma_d/\Gamma_d$ can be enhance by a factor 5 with respect the prediction of the SM. This effect could explain the discrepancy between the semileptonic charge asymmetries a_{sl}^d and a_{sl}^s obtained by the D0 Collaboration, determined by measuring the like-sign dimuon charge asymmetry ([40],[41],[42],[43]) and the current experimental bounds [44]. In fact, in addition to the semileptonic asymmetries, $\Delta\Gamma_d/\Gamma_d$ has been identified as another source that contributes to the like-sign dimuon charge asymmetry [45]. If all NP effects in the dimuon asymmetry are due to $\Delta\Gamma_d$, then an enhancement factor of 6 with respect to its SM value is required [46]. On the other hand,if there are also NP contributions in the semileptonic asymmetries, then the enhancement factor in $\Delta\Gamma_d$ can be smaller, making the precise determination of this parameter crucial.

CHAPTER 2

The LHCb experiment

The LHCb experiment [47] is one of the most important experiments in the world performing high precision measurements in the heavy-flavour physics sector. It is located at the CERN, near Geneva, where it exploits proton-proton collisions from the Large Hadron Collider (LHC) [48]. In the following chapter a description of the LHCb detector is given, as well as a brief overview of the LHC collider and of the b-hadronic production inside it.

2.1 The Large Hadron Collider

LHC [48], with a length of approximately 27 km, is the world's largest particle accelerator. It is hosted in a circular tunnel about 100 m underground, which previously housed the Large Electron Positron (LEP) accelerator. It consists of a ring of superconducting magnets and accelerating structures that guide and boost two beams of high-energy particles in opposed directions inside two parallel vacuum pipes. It operates at cryogenic temperatures, about 1.9 K, to maximise the field strength of the superconducting NbTi magnets, with intensities up to 8 T. The LHC is designed to accelerate mainly protons, but it is also capable of accelerating heavy ions, making heavy ion physics studies possible. It is the last part of an accelerating chain, schematically illustrated in Fig. 2.1. The acceleration process starts with the linear accelerator LINAC2, where the protons reach energies of about 50 MeV. They are then injected in the circular accelerator PSB (Proton Synchrotron Booster), boosted up to an energy of 1.4 GeV. The proton acceleration continues in the Proton Synchroton (PS) and the Super Proton Synchroton (SPS), where they are accelerated up to an energy of 26 GeV and 450 GeV respectively, ready to be finally injected into LHC. Here, the particles are not distributed continuously in the beams, but grouped together in 2808 *bunches* each one filled with up to $1.15 \cdot 10^{11}$ protons. The time

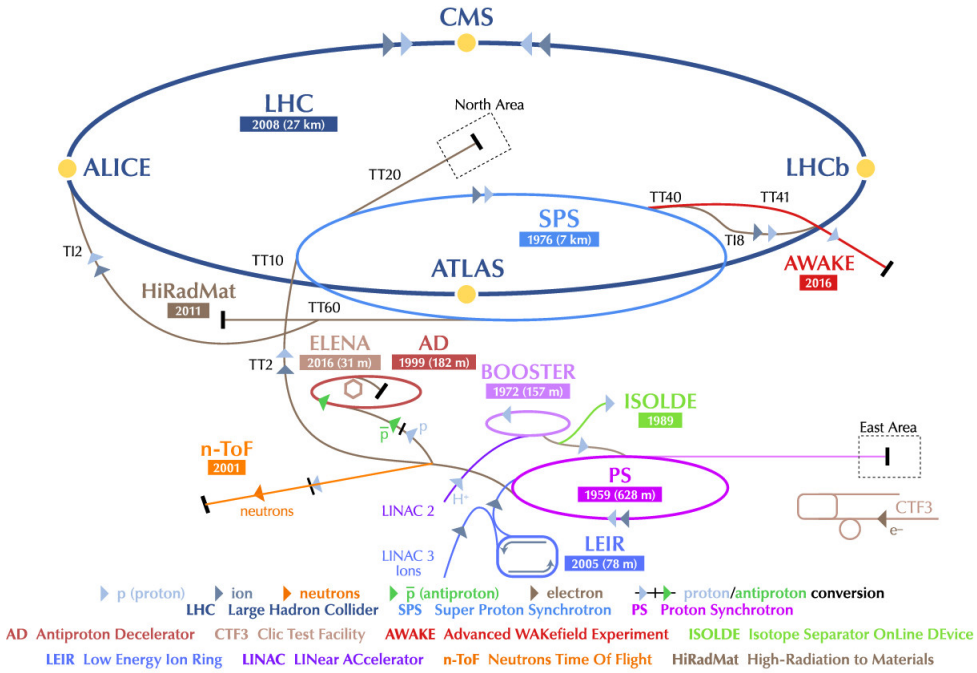

FIGURE 2.1. The CERN accelerator complex [49].

distance between two consecutive bunches is 25 ns (or multiple of 25 ns), corresponding to a maximum bunch crossing frequency of 40 MHz. The beams collide with a centre-of-mass energy of 14 TeV in four interaction points, in which are located the four major LHC experiments: ATLAS, CMS, ALICE and LHCb. The ATLAS [50] and CMS [51] experiments are general purpose detectors focused on the study of the electroweak and Higgs sector of the SM and on direct search of NP, in particular searching for production of beyond SM particles. The ALICE [52] experiment mainly investigates the properties of QCD at high density regimes in heavy ion collisions, with a particular focus to the phase transition to the quark–gluon plasma. The LHCb [47] experiment is dedicated to measuring the properties of b- and c-hadron decays and will be described in greater detail below. The LHC performs at a nominal instantaneous luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, though the LHCb and the ALICE experiments operate in a lower luminosity condition, in order to limit the number of interactions per bunch-crossing.

2.2 b-hadron production at the LHC

Heavy flavour production in proton-proton (pp) collisions at LHC can be described by the QCD-parton model. As a result of the strong interaction between partons in the incoming hadrons heavy $q\bar{q}$ pairs are created, mainly due to flavour creation processes, i.e., quark-antiquark annihilation, $q\bar{q} \to b\bar{b}$, and gluon-gluon fusion, $gg \to b\bar{b}$. Fig. 2.2 shows the leading-order Feynman diagrams of these production processes.

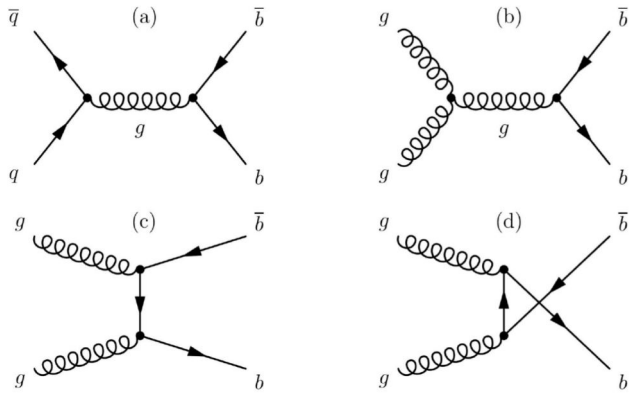

FIGURE 2.2. Examples of Feynman diagrams for $b\bar{b}$ pairs production. In the leading-order diagrams pairs are created through quark-antiquark annihilation (a) and gluon fusion (b,c,d).

The cross section of $b\bar{b}$ pairs production depends on the polar angle to the beam axis, because of the parton distribution functions of quarks and gluons inside the protons. Looking at the parton distribution functions at the center-of-mass energy of LHC, $Q^2 = 10^4$ GeV^2, in Fig. 2.3, it is clear these partons can carry very different momentum fraction of the protons, resulting in the production of a $b\bar{b}$ system that is likely boosted in either forward or backward direction along the beam line. Fig. 2.4 shows the polar angle distribution for the b and $\bar{b}$ quarks produced in proton-proton interactions at the LHC at $\sqrt{s} = 14$ TeV. They are produced mainly in the same direction close to the beam line.

The $b\bar{b}$ production cross section in pp collisions at the center-of-mass energy of 13 TeV has been measured by the LHCb experiment to be [55]:

$$\sigma(pp \to H_b X)(13 \text{ TeV}) = (144.3 \pm 1 \pm 21)\ \mu b \tag{2.1}$$

After the production the b and $\bar{b}$ quarks hadronise to form a bound state. The probability to hadronise in one of the different beauty hadrons, i.e. B^+ ($u\bar{b}$), B^0_d ($ud\bar{b}$), B^0_s ($s\bar{b}$) and b-barions, like the Λ^0_b (udb), can be inferred by looking at the measurement of the fragmentation fractions, which values are reported in Table 2.1. The B^+ and B^0_d mesons are assumed to be produced in equal amount, the B_c production is neglected and the sum of the fractions is constrained to unity.

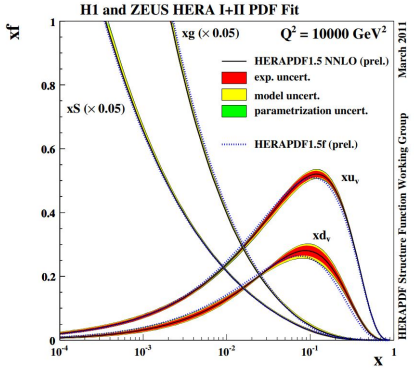

FIGURE 2.3. Simulation of polar angles distribution of the produced $b\bar{b}$ pairs at $\sqrt{s} = 14$ TeV. The LHCb acceptance is highlighted in red.

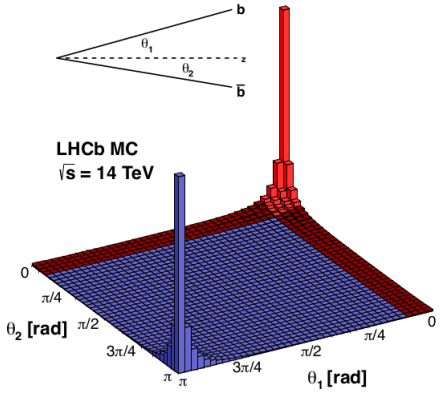

FIGURE 2.4. Production cross-section of $b\bar{b}$ pairs as a function of their polar angle to the beam axis, for proton-proton collisions simulated with Pythia [53] at a centre-of-mass energy of 14 TeV. The LHCb acceptance is highlighted in red. The plot is taken from Ref. [54]

b hadron fraction	value
B^+ or B^0_d	0.405 ± 0.006
B^0_s	0.101 ± 0.004
$b-$ barions	0.089 ± 0.012

TABLE 2.1. The measured b hadrons production fraction at the LHC [56]

2.3 The LHCb detector

The LHCb experiment (Large Hadron Collider beauty), as anticipated before, is one of the biggest detectors installed in the LHC accelerator. Thanks to the large production cross section for $b\bar{b}$, Eq. (2.1), in the LHCb acceptance that corrisponds approximately to 25% of the total pairs produced [55], and the excellent performance of the detector, LHCb is a fundamental

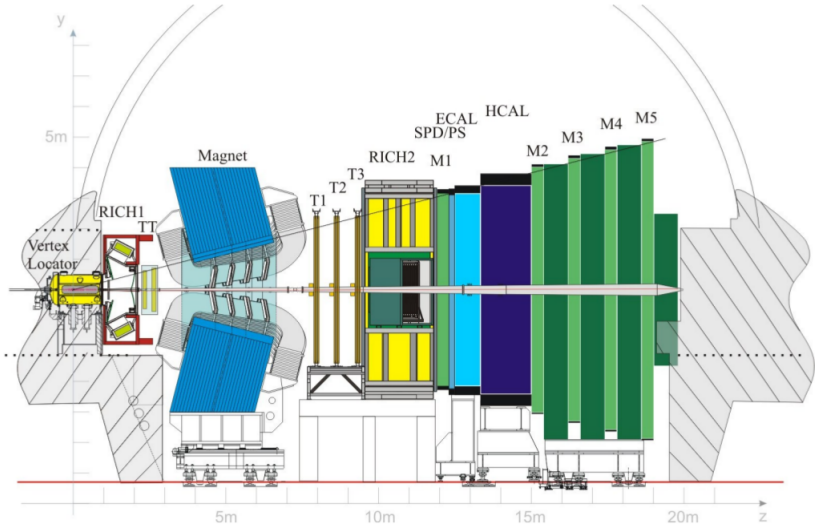

FIGURE 2.5. Scheme of the vertical section of the LHCb detector.

experiment performing high-precision tests on heavy flavour physics. With a wide physics program focused both in the beauty and charm sectors, the LHCb studies are mostly based on *CP* violation and decay properties (via angular and time-dependent analyses), and on the search for NP in rare decays. To maximise the number of $b\bar{b}$ pairs detectable in the LHCb acceptance, the detector is built as a single-arm spectrometer in the forward region with respect to the proton-proton interaction, and its vertical section is shown in Fig. 2.5. The coordinate system of LHCb is right-handed: the z axis follows the direction of the beam and goes from the point of interaction towards the muon chambers, while the y axis is vertically directed and pointing upwards. The particles produced in the *pp* interactions are deflected on the x-z plane by the magnet, called bending plane. The layout was designed to cover the angular region from approximately 10 mrad to 300 (250) mrad in the x-y (y-z) plane, translated as a pseudorapidity coverage of $2 < \eta < 5$, where the pseudorapidity is defined as

$$\eta = -\ln(\tan(\theta/2)), \tag{2.2}$$

and θ is the angle with respect to the beam axis.

The LHC instantaneous luminosity can be defined in terms of the machine parameters as [57]

$$\mathcal{L} = N_p^2 n_b f / A_{\perp}^{\text{eff}}, \tag{2.3}$$

where N_p, n_b, f and $A_{\perp}^{\text{eff}}$ are, respectively, the number of protons in a single bunch, the number of colliding bunches, the collision frequency and the effective transverse colliding area, which can be estimated from the overlap of the beam spatial distributions. As already anticipated, LHCb works at a luminosity of $4 \cdot 10^{32}\,\text{cm}^{-2}\text{s}^{-1}$, lower than the nominal operating value of LHC. In

order to meet the different experimental needs in term of luminosity, a levelling technique is used, tuning the transverse separation of the beams that constantly redefines the effective colliding area [58]. The lower luminosity presents several advantages; the radiation damage is reduced and the collisions between bunches occur mainly through single or double *pp* interactions, decreasing the number of primary vertices (PV) producing hadrons containing the b quark. This involves a simpler reconstruction of the PV and of the particle decay vertex (SV), guaranteeing a better reconstruction of the tracks. In view of achieving the aforementioned physics objectives, LHCb is built to have the following fundamental characteristics:

- a highly performing trigger system, optimised for the detection of b-hadrons;
- an excellent PV reconstruction system, essential for studying the oscillations of b-mesons and their *CP* violation;
- an excellent particle identification, to reduce the presence of wrong-identified particles, essential in the measurement of rare decays.

LHCb consists mainly of two types of detection systems:

- a tracking system composed of a vertex detector called VErtex LOcator (VELO), of the Tracker Turicensis (TT) and of T1, T2, T3 stations;
- an identification system, which includes two Cherenkov effect detectors (RICH1 and RICH2), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL), and finally five muon detection chambers (M1-M5).

These components are described in greater detail in the following paragraphs.

2.3.1 Dipole Magnet

In order to measure the momentum of the charged particles produced in the experiment, LHCb is equipped with a warm dipolar magnet [59] placed between the TT chambers and the first tracking station T1. The geometry of the magnet, with an opening angle of ±250 mrad vertically and ±300 mrad horizontally, has been chosen to take into account the acceptance of the experiment and it is shown in Fig 2.6. It is formed by two trapezoidal coils of 50 tons of pure conductor Al-99.7 cables bent at 45 degrees with respect to the x-z plane, so as to become wider as the z coordinate increases. With a nominal current of about 5.85 kA and a total resistance of 130 $m\Omega$ (at 20° C), it generates a magnetic field of about 1 T with a non-uniformity of the order of 1%, parallel to the y axis. Thus the charged particles are curved only in the x-z plane. In order to reduce the presence of systematic effects caused by a possible left-right asymmetry in the detector, the direction of the magnetic field is repeatedly reversed during the data taking. This is a crucial element for performing high-precision measurements of *CP* violation, and the

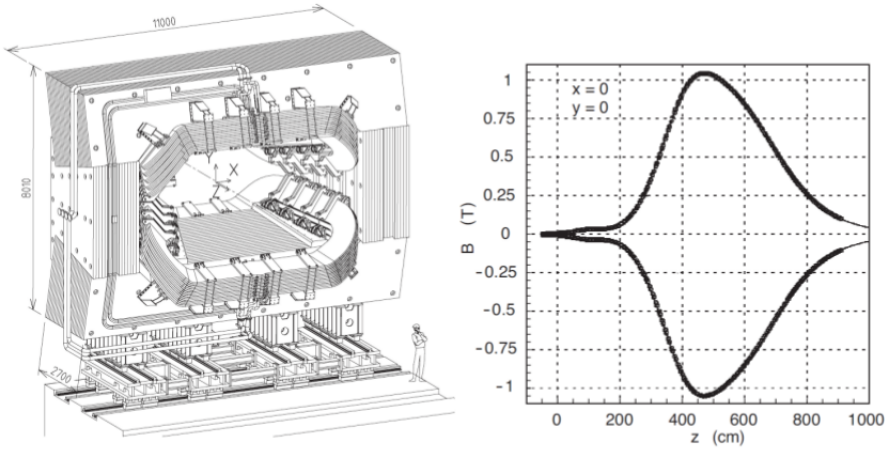

FIGURE 2.6. Drawing of the LHCb magnet (left) and the magnetic field along the z axis (right).

two direction conditions are named *MagUp* and *MagDown*. The region subject to the field ranges approximately from $z = 2.5$ m to $z = 7.95$ m.

2.3.2 Tracking system

The tracking system is one of the most crucial detector systems of the LHCb experiment. It is composed of the vertex detector (VELO) and four tracking stations: the Tracker Turicensis (TT), located between RICH-1 and the LHCb dipole magnet, and the T1-T3 stations, located over 3 metres between the magnet and RICH-2. Its purpose consists of recording the energy deposited in the detector material by the charged particles during their passage through LHCb, in order to reconstruct the trajectories and estimate the momenta.

2.3.2.1 VELO

The Vertex Locator (VELO) [60] has the purpose of accurately measuring the vertices of decay and production of unstable particles containing mainly b or c quarks. For this reason, it is positioned around the interaction vertex, extending from z = −18 cm to z = 80 cm. It consists of two identical rows of half-moon-shaped silicon sensors, each 0.3 mm thick, placed around the beam, which can move in a radial direction with respect to it, along the x-z plane. This allows the VELO to be positioned at a distance of 35 mm from the beam during the injection and acceleration phase, in order to protect the detector but also to prevent it from interfering with LHC operations. Once a stable phase has been obtained, in which the beams have reached the desired speed, the detector is brought closer, to a distance of about 7 mm. Each half is made up of 21 semicircular modules, consisting of pairs of silicon micro-strip planes, resistant to

radiation. In the interaction with the VELO, the particles produced in the pp collision generate electron-hole pairs, which give rise to measurable electrical signals. To allow a unambiguously identification of the position of the interaction point, the pairs of planes are formed by micro-strips that respectively give a response depending on the radial distance (r) from the beam and on the angular position (ϕ) in the x-y plane, with respect to the x axis. The information on the z coordinate is simply obtained from the position of the modules that gave rise to the electrical signal. The sensors are mounted on a support and contained in a vacuum vessel, which maintains the vacuum inside (about 10^{-4} mbar), providing also a separation from the Ultra High Vacuum of LHC (10^{-8} mbar), through a 0.3 mm thick layer of aluminium, the so-called RF foil, which faces the beam. Additional modules are installed upstream and constitute the pile-up veto system, used to measure the backward track multiplicity and to detect events with multiple primary interactions.
To ensure the alignment of the modules with the beam axis, the two halves of each module are slightly overlapped, as shown in the image on the left in Fig. 2.7. Given the multiplicity of tracks, it is necessary that the VELO has a micrometric precision. For this reason, the individual strips have dimensions ranging from 40 μm to 100 μm, decreasing the resolution as you move away from the beam. In addition, the modules were installed at such distances as to ensure the crossing of at least 3 of them by the particles produced in the PV, with pseudo-rapidity observable in LHCb. This guarantees a more precise identification of the vertex.
The PV is determined with a resolution of 13 μm in the transversal (x and y) direction and 71 μm in the longitudinal (z) direction. The resolution degrades to $\sim$150 μm for secondary vertices, due to the smaller number of tracks. The impact parameter, i.e. the distance between the track's point of closest approach to the PV and the PV itself, is measured with a resolution of 44 μm in the transversal direction for particles with transverse momentum of 1 GeV/c, that reduces to 15 μm for larger transverse momentum [61].

2.3.2.2 Tracking stations

The tracking stations, as already mentioned, consists of the Tracker Turicensis (TT), located between RICH-1 and the LHCb dipole magnet, and the T1-T3 stations, located over 3 metres between the magnet and RICH-2. Since the density of tracks scales with the inverse of the square of the beam axis distance, two detector technologies are employed: silicon microstrip detectors for the Silicon Tracker and straw-tube drift chambers for the Outer Tracker.

Silicon Tracker The Silicon Tracker (ST) [62] is composed of the TT station and the inner part of the T1-T3 stations, called the Inner Tracker (IT), as shown in Fig. 2.9. It consists of silicon microstrips sensors, with a strip pitch of about 200 μm, a hit resolution of 50 μm and a fast time response. Each of these detectors has four detection layers in an x-u-v-x arrangement, as shown in Fig. 2.8. The two x layers are composed of vertical detector elements, and the u

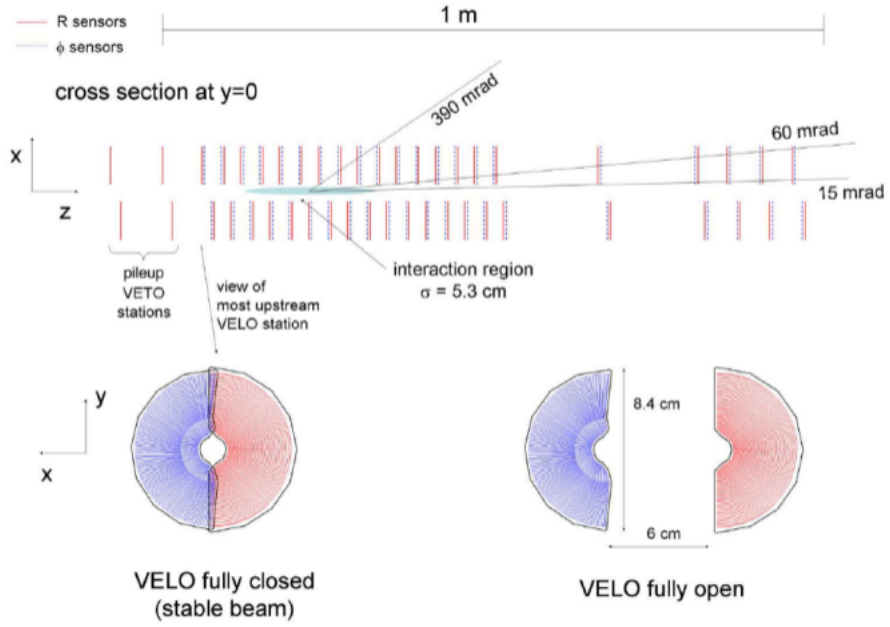

FIGURE 2.7. Scheme of the VELO layout, with R sensors in blue and Φ sensors in red. A representation of the two conditions of VELO fully closed and fully open is also shown (bottom).

and v layers are composed of elements rotated by a stereo angle of $-5°$ and $+5°$, respectively. This configuration allows for a measurement of both the x and y coordinates of a traversing particle, also getting the best hit resolution.

- **Tracker Turicensis**: the TT consists of a system of silicon detectors placed upstream of the magnet, which has the function of determining the position of the tracks after the VELO. In particular, it has the fundamental purpose of providing information for the reconstruction of decay products of long lived particles like K_S^0 and Λ, which mostly decay after the VELO, and of low momentum tracks, which do not reach the other tracking stations due to the deviation induced by the magnetic field. Its size is 150 cm wide and 130 cm high with a detection area of about 8.4 m^2.

- **Inner Tracker**: the IT covers a roughly 120 cm wide and 40 cm high cross-shaped region in the centre of the T1-T3 tracking stations, downstream of the magnet. Here the occupancy is higher ($5 \cdot 10^5$ cm^{-2} s^{-1}) and therefore a higher spatial resolution is required. Each station is arranged in four boxes surrounding the beam pipe, each one containing four layers of detectors in the x-u-v-x arrangement. About 20% of all produced charged tracks traversing the tracking system passes through this region.

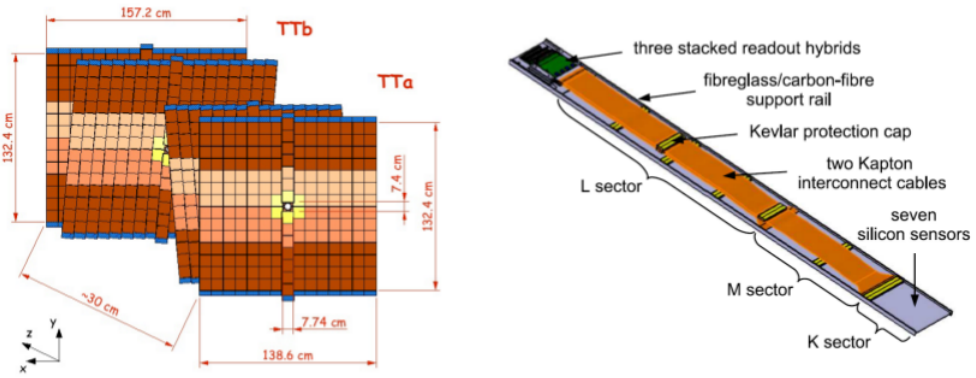

FIGURE 2.8. Left: view of the TT detector with the readout electronics in blue and the other readout sectors. Right: View of the TT module with three readout sections.

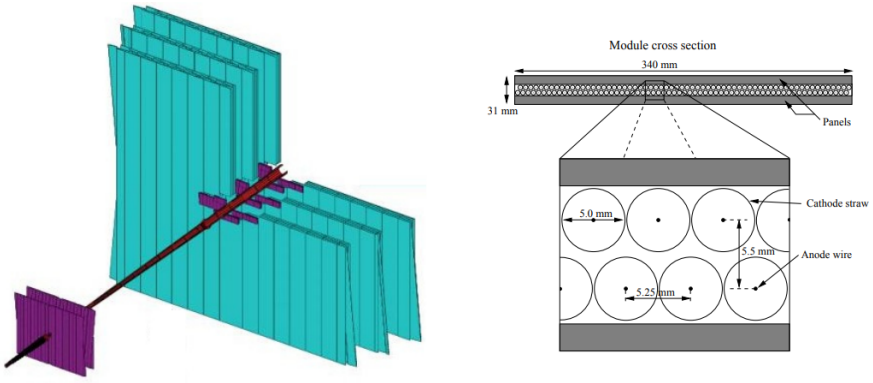

FIGURE 2.9. Left: scheme of the tracking system, with the OT (light blue), the beam pipe (brown) and the ST (violet). Right: section of a straw-tube in an OT module.

Outer Tracker The Outer Tracker (OT) [63] covers the remaining outer regions of the T1-T3 stations. It is a drift chamber detector, designed as an array of gas-tight straw tube modules. Each stations is composed of four modules in the x-u-v-x configuration, and each module contains two monolayers of drift tubes, vertically oriented, with diameters of 5 mm and filled with a gas mixture of CO_2 (∼30%) and Ar (∼70%). A 25 μm thick anode wire is located at the centre of each straw, as shown in Fig. 2.9. The drift time across the tube is less than 50 ns and the spatial resolution is about 200 μm.

2.3.3 Particle Identification system

The identification system has the purpose of distinguishing the different particles species produced in pp collision. Considering that b-hadrons decay into final states containing mainly pions, kaons, protons and muons, distinguishing these types of particles with great precision

is an extremely important requirement in LHCb. In particular, considering that pions are the particles mostly produced in LHC, their distinction is of extraordinary importance to reduce the combinatorial background. For the same reason, it is crucial to identify the muons, largely produced in the decays involving the $J\psi$ in the final states. Such a function is performed by the RICH detectors and muon chambers, respectively. The detection of neutral hadrons and photons and the measurement of the energy of the particles is finally left to the two calorimeters.

2.3.3.1 Ring Imaging Cherenkov detectors (RICH)

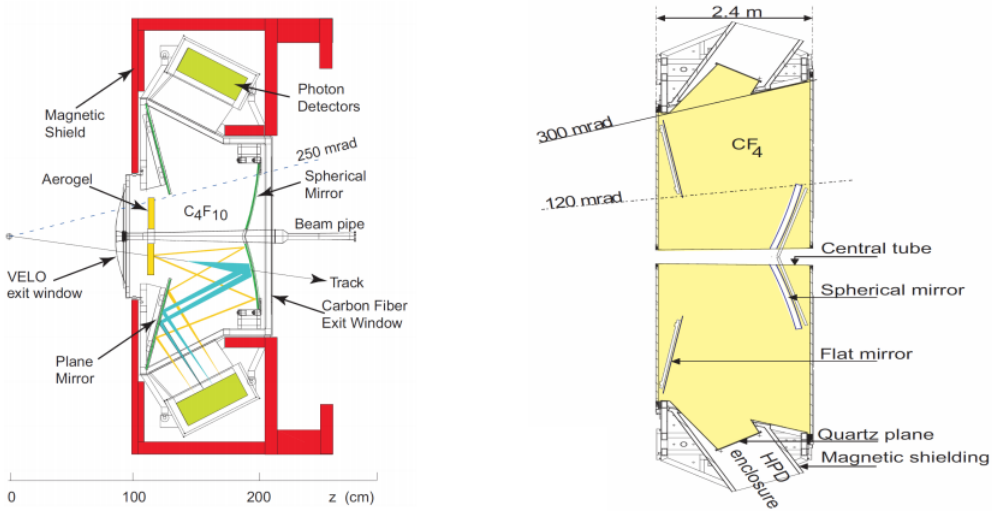

FIGURE 2.10. Schematic side view of the RICH-1 (left) and RICH-2 (right) detectors.

The RICH (Ring Imaging Cherenkov) detectors [64], which are schematically shown in Fig. 2.10, have the main purpose to separate kaons, pions and protons. They identify particles by exploiting the cone of light produced by the Cherenkov effect, in an overall momentum range of $(1 - 150)$ GeV/c. An example of the rings derived from the cones of light produced by the Cherenkov effect is shown Fig. 2.11 (right). If a charged particle, of mass m and momentum p, crosses a medium with refractive index n, with speed greater than that of the light in that medium, it emits photons along a cone, whose angular aperture (θ_c) depends on p and m according to the relation:

$$\cos\theta_c = \frac{1}{n}\sqrt{1 + \left(\frac{mc}{p}\right)^2} \tag{2.4}$$

The separation of the different particle types obtained with the RICH detectors is shown in Fig. 2.11 (left) [65]. In order to cover the entire range of momenta of $b-$ and $c-$ hadrons decay products, it is necessary to use different types of medium, called radiators. For this reason,

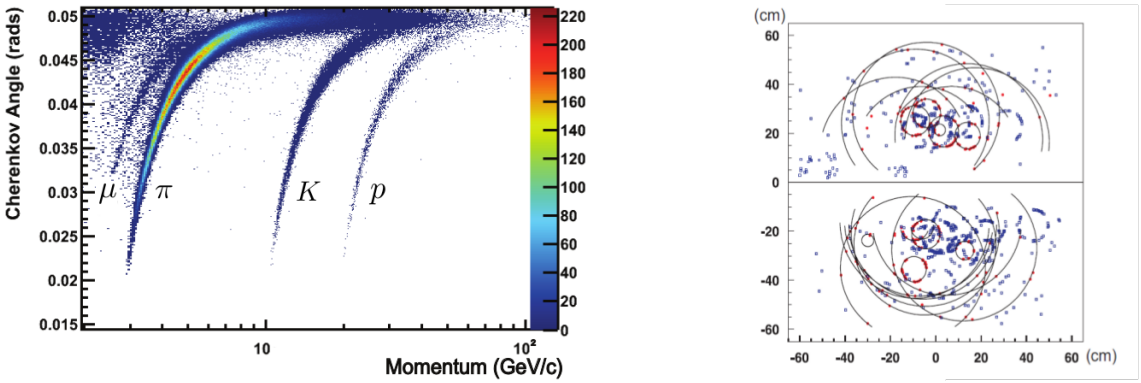

FIGURE 2.11. Left: Cherenkov angle versus particle momentum [65]. Right: a typical LHCb event in RICH-1 with Cherenkov rings interpolation.

LHCb is composed of two RICHs. The first, RICH-1, is positioned immediately after the VELO and consists of gaseous C_4 F_{10} ($n = 1.0014$) which has the role of covering the range of small momenta, $1 - 60$ GeV/c. The second, RICH-2, is located after the magnet. It is composed of CF_4 ($n = 1.0005$) and it is sensitive to particles with large impulse, up to $p \sim 100\,\text{GeV}/c$. The RICH-1 has a higher angular acceptance, ± 25 mrad to ± 300 mrad horizontally and ± 250 mrad vertically, compared to the RICH-2, for which the acceptance ranges go from ± 15 mrad to ± 120 mrad horizontally and ± 100 mrad vertically . Both RICH detectors are equipped with a mirror system to focus the Cherenkov light on Hybrid Photon Detectors (HPDs), placed out of acceptance and shielded from the magnetic fields. The measurement of the angular opening, combined with the information on the momentum obtained from the tracking system, allows the identification of the particle.

2.3.3.2 Calorimeter system

The calorimeter system [66] is located after the RICH-2 and consists of four stations: a Scintillator Pad Detector (SPD), a PreShower detector (PS), an electromagnetic (ECAL) and a hadronic (HCAL) calorimeter. The purpose is the identification of electrons, photons and neutral hadrons and the measurement of their transverse energy. All stations are made up of scintillating planes interspersed with absorption materials. A particle passing through the calorimeters produces a shower, composed of charged and neutral particles that release their energy inside the system, in the absorbers or in the active materials. The overall quantity of light produced in the scintillation materials depends on the total energy released. The scintillation light is transmitted to photomultipliers (PMTs), through WaveLength-Shifting (WLS) fibres that change the wavelength, adapting it to the absorber of the PMTs. The system composed of SPD and PS consists of a matrix of cells, which are composed of a 15 mm thick scintillator and a 12 mm-thick

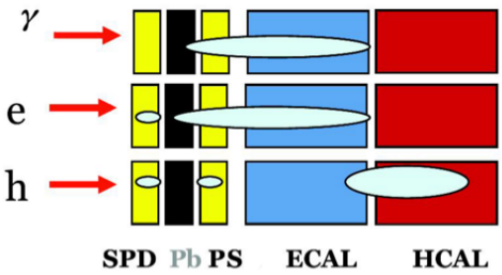

FIGURE 2.12. Schematic representation of the separation between photons, electrons and hadrons within the calorimeter system.

layer of lead. The SPD purpose is to identify charged tracks, while the PS discriminates photons and pions from electrons, in a fast way to use the information at hardware trigger level that will be explained in more details in the following section. Next, the electromagnetic calorimeter (ECAL), consisting of 4 mm thick scintillators and 2 mm thick layers of lead, measures the energy of electromagnetic showers produced by electrons and photons with the excellent resolution of

$$\frac{\sigma(E)}{E} = \frac{9\%}{\sqrt{E}} \oplus 0.8\%,$$

with E expressed in GeV, where the first term is due to the statistical uncertainty on the energy deposit, while the second is a constant contribution. The hadronic calorimeter (HCAL) is a sampling calorimeter, with the scintillating tiles arranged parallel to the beam pipe. It is composed of 16 mm thick iron tiles and 4 mm thick of scintillator layers, for a total extension along z of 1.6 m. The resolution in energy is

$$\frac{\sigma(E)}{E} = \frac{69\%}{\sqrt{E}} \oplus 9\%,$$

with E expressed in GeV.

The entire calorimeter system allows to perform a separation between electrons, photons and hadrons as shown in Fig. 2.12. In particular:

- photons do not release energy in the SPD station, but interact with the lead converter, creating an electromagnetic shower in the PS and ECAL;
- electrons behave like the photons at the PS and ECAL level, but also leave hits in the SPD detector;
- hadrons are typically Minimum Ionising Particles (MIPs) in the SPD, PS, and in the ECAL and basically all the energy is released in the HCAL.

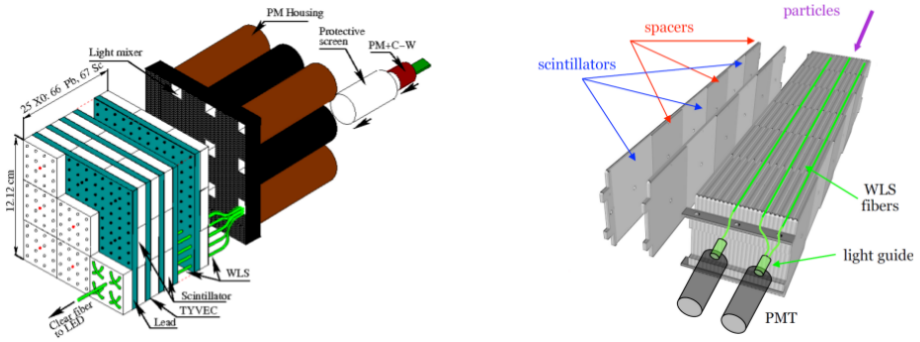

FIGURE 2.13. Illustrations of the ECAL cell (left) and HCAL cell (right) structures.

2.3.3.3 Muon system

The b hadrons often decay into final states containing muons, so LHCb must be equipped with a very efficient muon identification system. The muon detector [67–69] is placed downstream of the calorimeters, since the muons are the only charged particles that can pass through the calorimeters as minimum ionising particles with a low energy loss. It consists of 5 stations, each of which is divided into 4 regions, R1-R4 (R1 is the closest to the beam and R4 the furthest away). The size of each region, and its segmentation, increases with the distance from the beam, in order to keep constant the flow of particles intercepted by each region with the ratios 1:4:16:64, as shown in Fig. 2.14. The first station (M1) is placed between the RICH-2 and the SPD and it has the main purpose of measuring the transverse momentum, used in the hardware Trigger. The remaining stations are located at the end of LHCb and are interspersed with iron absorbers, 80 cm thick, which have the function of blocking hadrons. The whole detector is composed of 1380 chambers, 1368 Multi-Wire Proportional Chambers (MWPCs) and 12 triple GEM detectors (Gas Electron Multiplier). In particular each station is equipped with 276 MWPCs: 12, 24, 48 and 192 in R1, R2, R3 and R4 respectively, with the exception of the M1R1 region, which uses 12 GEMs, since they allow a higher granularity and a better response for high particles occupancy, which in this region is about 500 kHz/cm^2.

2.3.4 Trigger and data processing algorithms

The actual bunch crossing rate is of approximately 30 MHz and does not allow to write in a storage the information related to each pp collision event. The main purpose of the trigger is therefore to select the events of interest for the analyses carried out in the LHCb experiment, in order to decrease the rate with which the events are recorded, compatibly with the current storage rate for offline analyses of 5-12 KHz [70, 71]. In LHCb, the trigger system [72, 73] is composed of two main levels: an hardware trigger, called Level-0 (L0) and a software trigger, called High Level Trigger (HLT).

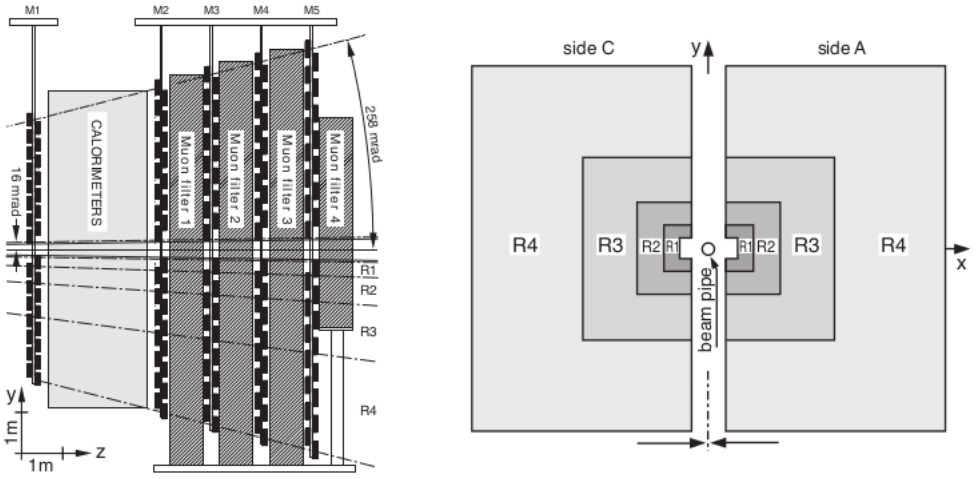

FIGURE 2.14. Side (left) and front (right) view of the LHCb Muon System. The R1-R4 projective scaling is also represented.

2.3.4.1 Level-0 trigger

The L0 trigger is a hardware trigger that reduces the frequency of the events from 40 MHz to 1 MHz. Specifically, it reduces the event rate to 450 kHz for events where charged hadrons have triggered, 400 kHz for those where the trigger is due to muons, and 150 kHz for those where the trigger is due to electrons and photons. The trigger works synchronously with bunch crossing and uses general event information provided from the different sub-detectors to perform a selection. In particular, the B and D decays of interest produce secondary particles with large transverse momentum and transverse energy. Depending on the nature of the particles produced, and therefore with which detectors of the tracking and calorimetry systems they interacted, an event can overcome one or more trigger lines operating in parallel: the L0-calorimeter trigger and the L0-muon trigger;

L0 calorimeter trigger

The L0 calorimeter uses the transverse energy in 2×2 cell blocks in ECAL and HCAL, defined as $E_T = \sum_{c=1}^{4} E_c \sin\theta_c$, where c is the cell index, E_c the energy deposit in each cell and θ_c the angle between the z axis and the line joining the cell to the nominal detector interaction region. Adding the information coming from PS and SPD the trigger performs the selection and flag the objects as "photon", "hadron" or "electron" according to the following definitions:

- Photon trigger (`L0Photon`): defined as the candidate with the highest ECAL only, E_T^{ECAL}, with also corresponding PS hits in front of the cluster and no hits in the aligned SPD cells. A typical threshold is $E_T > 2.7\,\text{GeV}$.

- Hadron trigger (`L0Hadron`): defined as the candidate with the highest E_T^{HCAL}. If there is an aligned cluster with the highest E_T^{ECAL} in ECAL the two values are summed. A typical threshold is $E_T > 4\,\text{GeV}$.

- Electron trigger (`L0Electron`): similar to `L0Photon` but with the additional requirement of at least one hit in the aligned SPD cells.

It is important to remark that the L0 thresholds varied a lot during the running period, following the evolving filling scheme of the LHC machine.

L0 Muon trigger

The muon trigger searches instead for muon candidates with transverse momentum greater than a given threshold, using the information of all the five muon stations. The hits in the muon stations are processed in parallel to search for the muon tracks. Then, the track direction is used to estimate the transverse momentum of the muon candidate, assuming it is coming from the pp interaction point and a single kick in the magnetic field. The calculated p_T, with a resolution of 20%, is then used to take the final decision:

- Muon trigger (`L0Muon`): the candidate with a p_T greater than a given threshold, typically about $2.5 - 3\,\text{GeV}/c$.

- Di-Muon trigger (`L0DiMuon`): the square root of the product of the two largest p_T has to be greater than a given value, typically about $1.5\,\text{GeV}/c$.

2.3.4.2 High Level Trigger

The events passing the hardware selection are processed by a software trigger (HLT), which works asynchronously with respect to the bunch crossing rate. It makes two successive selections to the L0 level, called HLT1 and HLT2.

HLT1

The HLT1 trigger further reduces the rate from 1 MHz to about 30 kHz, performing a partial reconstruction of the events, combining information from the VELO, the tracking system and the muon chambers. The event is processed on different HLT1 trigger selection lines, i.e. a series of parallel selection algorithms. If the L0 stage is passed, as first step VELO tracks [74] and PV are reconstructed, and the information from the VELO are matched with those of the tracking stations. This allow a refined measurement of the transverse momentum p_T with respect to the L0 trigger and an estimation of the particle charge. The tracks are then fitted with a bi-directional Kalman filter [75] in order to account for effects like multiple scattering and ionisation energy loss, further refining the track parameters. Then the impact parameters of

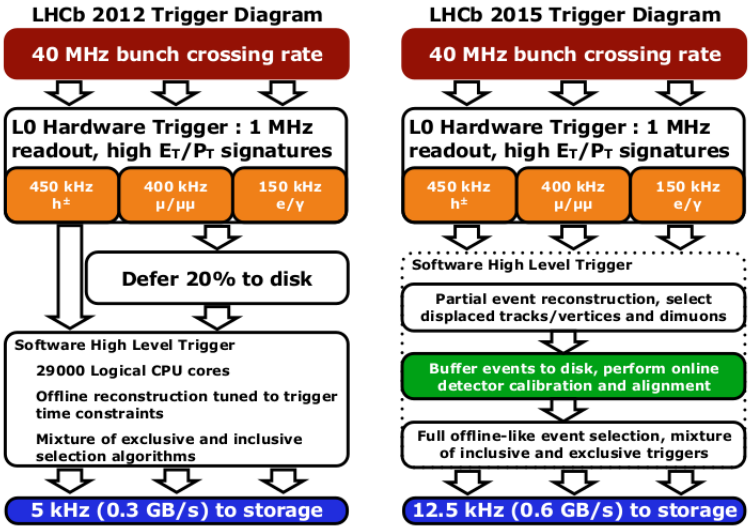

FIGURE 2.15. The LHCb trigger schemes for Run 1 (left) and Run 2 (right).

the tracks with respect to the PV are determined and the secondary vertices are reconstructed. Finally, the invariant mass and directions of flight are calculated and a first physics selection can be performed. Having a richer set of parameters with respect to L0, simple multivariate classifiers can be used to get an improved selection performance with respect to simple cuts. This is very useful for selecting one and two tracks combinations (topological triggers) [76] with an associated good quality vertex. In order to identify the muons, VELO and T-station tracks are compared with muon chamber hits. The parameter of the combined tracks are used to apply cuts on the momentum, transverse momentum and the Impact Parameter χ^2, which is the χ^2 difference between a PV fit with and without the considered track. If the event is selected by at least one HLT1 line, it can be further processed by the second level, HLT2.

HLT2

In HLT2, the events are totally reconstructed, including also information from RICH and calorimeters, after a precise real time alignment and calibration of the detector. The frequency is reduced to approximately 5-12.5 kHz, at which the data can be stored. Similarly to the HLT1 case, the HLT2 trigger is composed of a series of selection algorithms (lines) that are applied in parallel to the events and are mainly divided into two categories: the inclusive topological lines, in order to select also partially reconstructed decays, and exclusive lines that select specific final states. The HLT processing flow for Run 1 and Run 2 is shown in Fig. 2.15.

2.4 Event reconstruction at LHCb

2.4.1 The LHCb software

The information coming directly from the detector, from the readout electronics response, is saved in files called RAW. The subsequent data reconstruction and analysis are performed using different applications based on a common framework, called GAUDI [77]. The main purpose of this framework is to define a set of common interfaces for the various applications, keeping them decoupled from each other. The most important applications necessary to perform an analysis in LHCb are briefly described below:

- BRUNEL [78]: is the LHCb event reconstruction application, run centrally by the LHCb computing team. Its main purpose is the reconstruction of charged tracks by combining the hits from the tracking system, performing the so-called off-line reconstruction. In addition, it associates the information obtained from RICH, calorimeters and the muon system to the reconstructed tracks.

- DAVINCI [79]: is the LHCb physics analysis application. It applies particle identification hypotheses on the tracks already present at reconstruction level, creating the stable particles used to identify the decay chain of interest. The particle tracks are then used by the decay tree fitter (DTF) [80] algorithm to perform a simultaneous fit of all the vertices of the decay chain, allowing their finishing compared to the HLT vertex reconstruction. In addition, it allows the application of a loose selection on the available event information of the decay chain.

- MOORE [81]: is the LHCb application that applies the software trigger algorithms, performing the so-called on-line reconstruction. Since during the analysis process both the off-line and the on-line reconstruction applications are used, they share most of the reconstruction and selection algorithms, in order to keep the reconstructions as much as possible similar. Nevertheless, due to timing requirements in the software trigger, some simplifications are necessary.

- GAUSS [82]: is the LHCb event simulation program. The procedure of simulation first simulates the proton-proton collisions and the particle decays using the Pythia program [83] and the EvtGen program [84] respectively. Subsequently the Geant4 [85] program is used to simulate the interaction of the generated particles with the LHCb detector.

- BOOLE [86]: is the program that reproduces the digital response of the LHCb experiment, taking into account also possible dead channels, electronic noise or cross-talk, keeping therefore the same output format as the real data coming from the detector.

After the reconstruction, the full event information is saved in files called "Data Summary Tape" (DST), which typically takes around 150 kB of disk space per event and hundred of TB

in total. Allowing LHCb users direct access to the whole event information would be prohibitive in terms of disk space and computing resources, therefore a subsequent centralised procedure called *stripping* is implemented [87]. It consists of a series of offline loose selection algorithms, called stripping lines, with the aim of selecting the events of interest. The stripping lines are grouped into *streams* according to the type of analysis to which they are dedicated, and centrally run over the DST files. The events selected from a particular stripping line are called *candidates* and are saved in files usable by analysts in formats like DST or μDST. The last format stores only the information of the candidate of interest and discard the raw event, allowing to save further disk space.

2.4.2 Trigger decisions: TIS, TOS, Dec

In LHCb all the information used in the reconstruction, like detector channels, logical pads, etc. has a unique numerical identifier, called LHCbID. Once an event is accepted by the stripping selection, the LHCbIDs of the decay chain can be compared to the ones of the trigger objects. This allows to perform various checks, among others the measurement of the efficiency of a specific trigger requirement directly from data (TIS-TOS method [88]). Three main categories can be defined [88]:

- Trigger On Signal (TOS): events for which the presence of the signal of interest is sufficient to fire the trigger line, regardless of the rest of the event. More precisely, a candidate is considered TOS with respect to a trigger line if the LHCbIDs of the final state particles of the trigger accepted decay overlap for more than 70% with the LHCbIDs of the final state particles of the offline signal candidate.

- Trigger Independent of Signal (TIS): events for which the trigger line is fired regardless of the presence of the signal, leads to a signal sample that is unbiased with respect to that particular trigger requirements. In this case the LHCbIDs of the triggered decay overlap less than 1% compared to those of the offline candidate.

- Trigger Decision (Dec): all the events that have passed the selection of triggers, regardless their classification as TIS or TOS.

2.4.3 Main processing algorithms

In the following sections two main algorithms running at HLT and offline reconstruction level are explained: the tracks reconstruction and the particle identification algorithm.

2.4.3.1 Tracks reconstruction

In the LHCb environment, the track reconstruction is one of the most fundamental and challenging data processing algorithm, because of the high track density of typical pp collisions.

In fact, the main purpose is not only to reconstruct the tracks from the decays of interest, but also all tracks coming from the primary vertex, in order to define it precisely. This is crucial, among others, for the determination of the b-hadron lifetime. The track reconstruction consists of two separate phases: the pattern recognition or track finding (assignment of clusters to the tracks) and the track fit (determining the optimal track parameters). Different types of tracks are defined. The track reconstruction algorithms and the various track types are described in more detail in the next paragraphs.

Track types Depending on their origin vertex and their momentum, the charged tracks have different trajectories through the detector, leading to the following classification, as illustrated in Fig. 2.16 [74] :

- **VELO Tracks**: tracks leaving at least three hits in the r-sensors and at least three hits in the ϕ-sensors of the VELO. Since the magnetic field in the VELO region is negligible, no information on the charge or momentum of the track can be inferred but only the direction of the particle. Since the reconstruction of a track inside the VELO is an important step of the analyses presented in this thesis, the dedicated algorithm FastVELO [89] will be briefly presented.
  It is based on two main assumptions, necessary to allow the algorithm to be fast enough in order to be used also in the trigger. Firstly, because of the negligible magnetic field inside the VELO, charged-particle trajectories can be approximate to straight lines. Secondly, since most of the particles produced in a pp collision come from the interaction point, the particles originates from the PV and therefore the projections of their R coordinates along the z axis must follow a straight line. The reconstruction starts using only the information of the R sensors, searching for three or four hits building a straight line in the R - z projections. Then, additional hits that correspond to the straight line extrapolation are added to these preliminary tracks. Only at this stage, hits from the φ sensors are associated to the tracks, completing the reconstruction procedure.
  If possible, these tracks are extended to the other subdetectors, then forming upstream or long tracks, defined below. If not, they correspond generally to particles generated at high polar angles or which go backwards with respect to the detector. These tracks are useful for precise reconstruction of the primary vertices.

- **Upstream Tracks**: VELO tracks with additional TT hit information. They have generally a low momentum and are therefore bended out of the detector acceptance by the magnetic field. Passing through the RICH1, they emit Cherenkov photons if $p > 1\,\text{GeV}/c$ and can be used to study the background in the particle identification algorithms in the RICH. Because the deflection between the VELO and TT due to the magnetic field is limited, the momentum resolution is $\delta p/p \sim 15\%$.

- **T Tracks**: tracks leaving at least one hit in both the x and stereo-layers for each of the T-stations. If possible, they are extended to long tracks or downstream tracks. If not, they are mainly generated from secondary interactions in the material of the upstream detectors and are used for the RICH2 pattern recognition and for the tracking station internal alignment.

- **Downstream Tracks**: T-tracks with additional TT hit information. These tracks are mainly generated by long lived particles like K_S^0 and Λ. The momentum resolution is $\delta p/p \sim 0.43\%$, comparable to that of long tracks.

- **Long Tracks**: tracks leaving hits in the VELO and in all the T-stations and thus have the most accurate momentum measurement, $\delta p/p \sim 0.5\%$ at momentum less than $10\,\text{GeV}/c$, up to 1% at momentum of $200\,\text{GeV}/c$. Typically they are the most important tracks used for physics analyses.

Pattern recognition The pattern recognition is organised hierarchically and consists in several algorithms running in a sequence to find the various track types defined above.

- **VELO seeding**: the first step is the search for combinations of clusters which are already sufficient to define tracks. The clusters are then fitted by a straight line in 3D, forming the so-called VELO seed. These are used as input to further tracking algorithms.

- **Forward tracking**: VELO tracks are used as an input and extended in the T stations by using a Hough Transformation approach [90]. This corresponds to the main algorithm for the reconstruction of Long tracks, with an efficiency of about 90%.

- **T seeding**: the information collected in all the T stations are combined to reconstruct segments of tracks, under the hypothesis of parabolic trajectories. Initially only the x layers are used, and the $u - v$ layers are later added to reject random hit combinations.

- **Track matching**: the VELO and T seeds are taken as an input and matched to reconstruct a Long track. Thanks to this algorithm another 5% of Long tracks is found.

- **VELO/T tracking**: the VELO/T seeds that are not used in other pattern recognition algorithms are kept as VELO/T tracks.

- **Up/Downstream tracking**: VELO/T seeds are propagated forward/backward, if there are at least three hits from the TT detector matching them.

Track fitting

- **Kalman fitting** A bi-directional Kalman filter [75] fit is applied. This allows to take into account secondary interactions or energy loss, giving a better estimation of the track

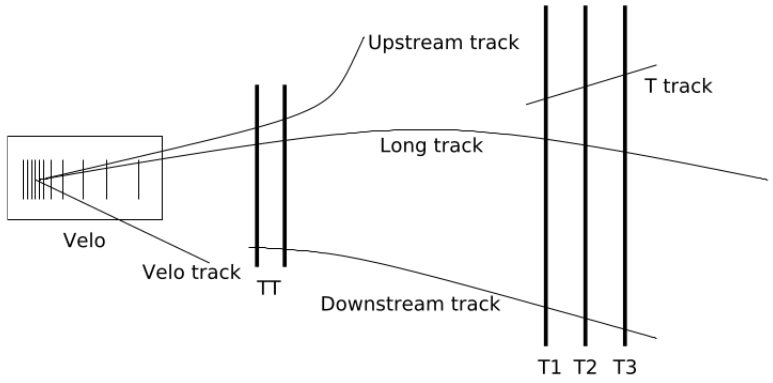

FIGURE 2.16. Schematic view of the different track types defined in the LHCb, shown in the $x - z$ plane.

parameters. The quality of each established trajectory is defined by the χ^2 per degree of freedom.

- **Clone killing**: some tracks could be reconstructed partially from the same hits. If the percentage of hits shared is greater than a certain threshold the tracks are called clones. In this case the tracks with the smaller number of hits are rejected by the clone-killing algorithm [91]. Whether two or more tracks have the same number of hits, the highest quality one (based on its χ^2) is retained.

- **Ghost removing**: some fake tracks can also be reconstructed from random combination of hits, creating the so called *ghost* tracks. They mostly arise from wrong associations between the VELO and T tracks which is caused by the large extrapolation distance when traversing the magnet. A neural network algorithm has been developed to decrease the number of ghost tracks. It uses information coming from the track fit, the track kinematics, and the number of measured hits in the tracking stations versus the number of expected hits and its output variables are used to quantify the probability that a track is a ghost, allowing the removal of them from the final data sample.

After the reconstruction, a tracking efficiency can be computed as the probability of fully reconstructing the trajectory of a particle leaving hits in the full tracking system. The average efficiency is above 96% for muons or hadrons that have passed through the full tracking system without hadronic interactions, in the momentum range $5 - 200\,\text{GeV}/c$ and within the pseudorapidity range of LHCb acceptance. As it will be shown in Chap. 6, this efficiency depends on the lifetime of the mother particle from which muons and hadrons come from, causing a distortion of the related decay-time distribution. As this work is focused on time-dependent analyses of b-meson decays, it is one of the most important experimental effects taken into account in this thesis.

2.4.3.2 Particle Identification

Particle identification in LHCb, i.e. the assignment of a mass hypothesis to the tracks, is provided by using information from different detector systems: the calorimeter system, the two RICH detectors and the muon stations. In particular, a likelihood based algorithm is used to calculate the probability associated with a specific mass hypothesis for each track produced in a pp collision. Consider a generic produced particle of type k, characterised by n experimental observables x_i, for example track parameters, energy deposit in the calorimeters etc. We can evaluate the probability that a particle with the measured observables x_i is a particle of type k as:

$$\mathscr{L}^k = \prod_i^n p_i^k(x_i), \tag{2.5}$$

where $p_i^k(x)$ are in general probability density functions extrapolated from simulations or from data calibration samples. In LHCb is common to express the probability using the log-likelihood form, in order to use sums instead of products:

$$\ln \mathscr{L}^k = \sum_i^n \ln p_i^k(x_i).$$

We can construct likelihoods for the different detector systems and combine them into a global likelihood. The simpler way is to multiply them, accordingly to a specific mass hypothesis, as:

$$\mathscr{L}(K) = \mathscr{L}^{RICH}(K) \cdot \mathscr{L}^{Calo}(K) \cdot \mathscr{L}^{Muon}(\bar{\mu}),$$
$$\mathscr{L}(\pi) = \mathscr{L}^{RICH}(\pi) \cdot \mathscr{L}^{Calo}(\pi) \cdot \mathscr{L}^{Muon}(\bar{\mu}),$$
$$\mathscr{L}(\mu) = \mathscr{L}^{RICH}(\mu) \cdot \mathscr{L}^{Calo}(\bar{h}, \bar{e}) \cdot \mathscr{L}^{Muon}(\mu),$$

where the bar indicates the hypothesis of not being that particle. Then the hypothesis about the particle nature can be compared. In particular we can define the global delta-log-likelihood (DLL) variable, to quantify how much a hypothesis is more likely than another:

$$DLL_{h_1/h_2} = \ln \mathscr{L}(h_1) - \ln \mathscr{L}(h_2)$$

, where h_1 and h_2 refer to different mass hypotheses of the track considered. Commonly, pions are used as h_2 particles, since they are the most produced particles in pp collisions, representing also the most common source of background in the measurements of physics parameters. The $DLL_{h/\pi}$ is often called `PIDh`.
Another more powerful strategy is to combine the sub-detectors information with that obtained from the tracking system using a multivariate algorithm, for example a neural network. This allows to exploit simultaneously all the information acquired by the LHCb detector and to directly translate them into a single probability value for each particle hypothesis, achieving better PID performances. This kind of variables are called `ProbNN` and are largely used in the analysis described in the following chapters.

CHAPTER 3

Analysis strategy

The analyses presented in this thesis aim to measure *CP*- violating and mixing parameters of B_s^0 and B^0 systems: the *CP*-violating phase ϕ_s, the decay widths Γ_s and Γ_H and the decay-widths differences $\Delta\Gamma_s$ and $\Delta\Gamma_d$. As shown in Chap. 1, all these parameters require a measurement of a time-dependent decay rate. Additionally, for the first four variables, the flavour at production of the B_s^0 mesons must be known, process commonly called *tagging*, as well as an angular analysis has to be performed in order to separate the *CP*-components of the final state. These measurements are very challenging, because many distorting effects caused by the detector or the analysis procedure itself can wrongly affect the measured values. In order to perform a precise measurement, it is therefore mandatory to have a good understanding of the experimental effects that can potentially affect the decay rate, the angular distributions and the tagging performance, and to correct for them.
In this chapter, an overview of the analyses strategies developed to perform these measurements are presented, with an important focus on the description of the different distorting sources and their corrections.

3.1 Decay-time of a particle and its experimental distortions

The time-dependent decay rate, as explained in Chap. 1, gives information about the probability that a particle produced at time $t = 0$, in this case a b-meson, decays into final state f at a certain time $t > 0$. The time t in which a particle decays is called the decay-time and it is given by

$$t = L\frac{m}{|\vec{p}|}, \tag{3.1}$$

where L is the flight distance travelled by the b-meson, calculated as the difference in position between the production vertex of the b-meson and its decay vertex, p is the reconstructed three-momentum and m is the b-meson reconstructed invariant mass. Because of the limited spacial and momentum resolution of the LHCb detector, the quantities used to determine the decay-time suffer from experimental uncertainties, that are the source of the decay-time resolution effect. Additionally, the decay-time distribution can be distorted by decay-time dependent efficiency effects, introduced by the detector or the analysis procedure, thus affecting the measurements of the observables of interest. Their typical shapes are shown in Fig. 3.1.

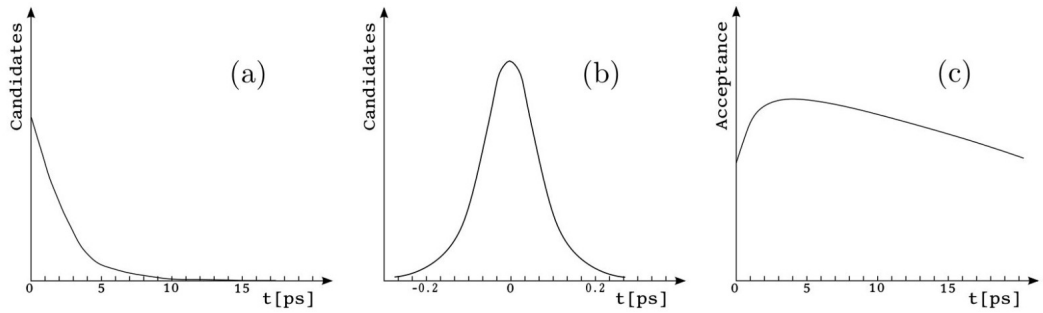

FIGURE 3.1. Picture of a typical (a) decay-time distribution, (b) decay-time resolution, and (c) decay-time dependent efficiency of a b-meson. Figure taken from [1].

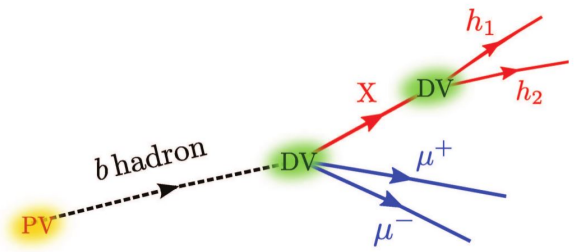

FIGURE 3.2. Schematic picture of a b-meson decay chain. It is produced at the PV and it decays into a $J\psi \to \mu^+\mu^-$ and another resonance $X \to h_1 h_2$. The J/ψ and X decay vertices, DV, are highlighted. Figure taken from [1].

Since the decay-time of a particle can be calculated by its reconstructed momentum, its production and decay vertices and its reconstructed mass, a precise determination of these quantities is necessary. Fig. 3.2 shows a schematic picture of a b-meson decay chain, supposing it is produced in the pp interaction point and that it decays promptly into a J/ψ and another resonant state X. Both these resonances decay into a pair of particles; in particular, we are interested in the J/ψ decay into two opposite charged muons, while the X resonance decays into two opposite charged hadrons (pions or kaons). The picture shows also the decay vertex of

the two resonances, DV. Since the J/ψ decays immediately after its production, its decay vertex coincides with the decay vertex of the b-meson. On the contrary, depending on its nature, the X resonance could have a not negligible lifetime, resulting in a DV detached with respect to its production vertex.
The LHCb framework allows to precisely reconstruct the decay chain in the presence of intermediate states by using the method called *decay tree fitter* (DTF) [80]. The algorithm performs a fit of the total decay chain, with a simultaneous determination of the position, momentum and decay-time of all particles involved. If a resonant state has a reconstructed decay length much smaller than its decay vertex resolution, the DTF will not determine its decay-time. The free parameters of the fit are the positions of the vertices and the momenta of all particles, while the mass of final state particles is assigned according to the particle hypothesis in the decay chain. During the fit, the four-momentum conservation at each vertex is required. The power of this tool lies in its ability to take into account correlations and dependencies between the decay chain parameters, performing a very precise determination of the variables of interest. The global fit to the particle parameters is performed using a Kalman Filter and the resulting Chi-squared per degrees of freedom, $\chi^2_{DTF}/nDoF$, can be used to partially remove decays that originate from wrong particle combinations, the so-called *combinatorial background*. The DTF method also ensures the determination of the uncertainty related to the reconstructed parameters, as the decay-time. This feature is very important, as it will be shown in the next section. In addition, it allows to put kinematical or geometrical constraints during the fit, such as constraints on the reconstructed mass of the intermediate resonances, or constraint on the origin vertex position of the b-mesons. This feature allows to obtain a better determination of some parameters, like the reconstructed mass of the b-meson or its decay-time.

3.1.1 Decay-time resolution

The impact of the decay-time resolution changes depending on the measurement to be performed. In particular it is not a dominant effect in the determination of $\Delta\Gamma_s$, Γ_H, Γ_s and $\Delta\Gamma_d$, while it has a big impact on the measurement of ϕ_s. In fact it smears neutral b-mesons oscillation pattern, whose amplitude depends on the mixing induced *CP*-violating phase. Assuming the presence of a Gaussian resolution contribution, characterised by a standard deviation σ_t and mean equal to zero, the smearing effect can be quantified by considering the convolution of the time-dependent decay rate with the resolution function

$$\Gamma(t)^{reco} = \Gamma(t)^{true} \otimes \frac{1}{\sqrt{2\pi}\,\sigma_t} e^{-\frac{(t)^2}{2\sigma_t^2}}. \tag{3.2}$$

The resolution effect can be more easily studied in the frequency domain, as suggested by Ref. [92]. Exploiting the convolution theorem, the convolution in the frequency domain can be

expressed as

$$FT\Big(\Gamma(t)^{true} \otimes \frac{1}{\sqrt{2\pi}\,\sigma_t} e^{-\frac{(t)^2}{2\sigma_t^2}}\Big)(\nu) = FT\Big(\Gamma(t)^{true}\Big)(\nu) \times FT\Big(\frac{1}{\sqrt{2\pi}\,\sigma_t} e^{-\frac{(t)^2}{2\sigma_t^2}}\Big)(\nu), \tag{3.3}$$

where FT is the Fourier transformation. The real part of the FT of the resolution function leads to a damping factor, D, of the amplitude of $\Gamma(t)$ in the frequency domain, namely

$$FT\Big(\frac{1}{\sqrt{2\pi}\,\sigma_t} e^{-\frac{(t)^2}{2\sigma_t^2}}\Big)(\nu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sigma_t)^2}{2}\nu^2}. \tag{3.4}$$

Greater is the frequency, greater is the amplitude reduction. If we consider only the oscillating part of the time-dependent decay rate, $\Gamma(t) = \cos(\Delta m t)$, the FT is non zero only at one frequency value, Δm, and the damping factor D is

$$D = e^{-\frac{(\sigma_t)^2}{2}(\Delta m)^2}. \tag{3.5}$$

In the more realistic case of an exponential suppression on top of the oscillating part of the decay rate, $\Gamma(t) = e^{-\Gamma t}\cos(\Delta m t)$, the real part of its FT is not characterised any more by a single frequency component, but by a frequency range centred at the characteristic frequency of the oscillation, Δm, to which corresponds the amplitude peak. However, in order to obtain the impact of the resolution, the damping factor can be nevertheless evaluated at the frequency of the amplitude peak, resulting in the same equation written above.

$$D = e^{-\frac{(\sigma_t)^2}{2}(\Delta m)^2} \tag{3.6}$$

This derivation will be useful for the determination of the decay-time resolution effect in the $B_s^0 \to J/\psi K^+K^-$ analysis in Chap.4

The damping factor is often called *dilution*, since its effect in the time domain results in a dilution of the amplitude of the b-mesons oscillation. If there is no decay-time resolution effect $D = 1$, while it leads to a smaller and smaller amplitude the greater is the decay time uncertainty. Being due to the limited resolution of the LHCb detector, the decay-time resolution depends on the phase space of the particles involved in the decay and typical values for b-meson decays in the LHCb experiment are in the range of $40 - 50$ fs, with a corresponding dilution of $0.7 - 0.8$. It becomes clear that an accurate evaluation of this effect and its correction is fundamental to obtain a precise measurement of ϕ_s.

In the context of $\Delta\Gamma$ or Γ measurements, given that these quantities appear in the time-dependent decay rate as parameters of the oscillating terms and not as amplitudes, the dilution effect is less important.

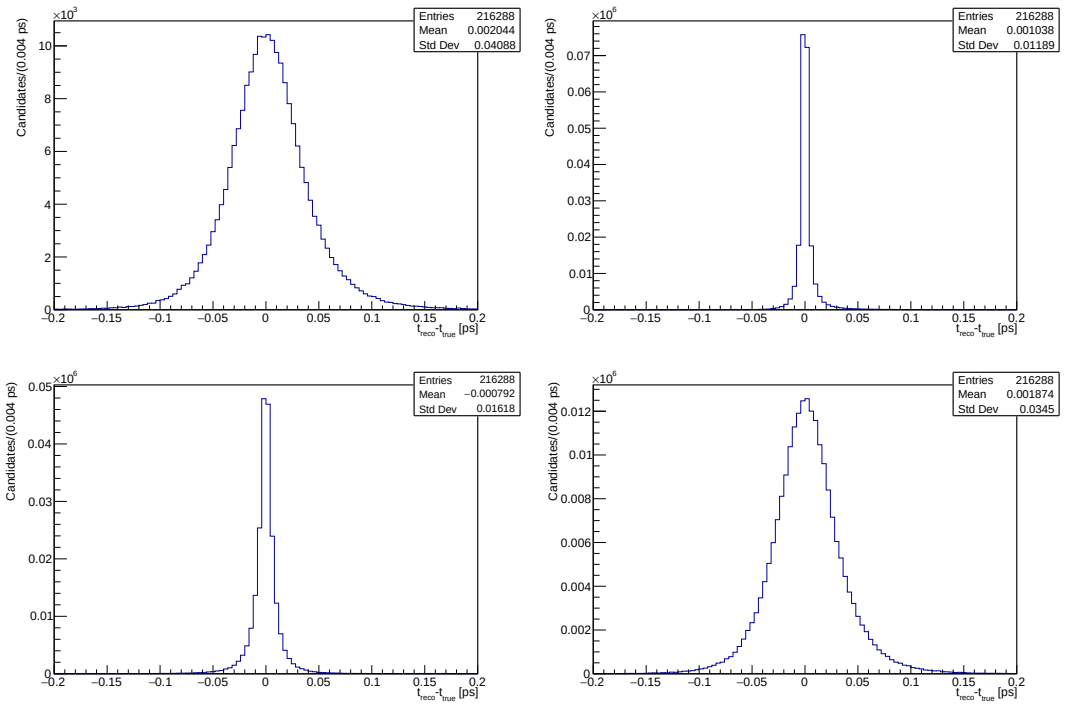

FIGURE 3.3. Decay-time resolutions obtained from a simulated sample of $B_s^0 \to J/\psi\pi^+\pi^-$ decays, where the reconstructed decay time is calculated using Eq. 3.1. On top, the plots show the resolution obtained when using only reconstructed quantities (left) or when using the reconstructed value for the momentum and true values for the other variables (right). On bottom, the plots show the resolution obtained when using true quantities except in the calculation of the flight distance, where the reconstructed PV position (left) or reconstructed DV position (right), is used. The B_s^0 mass is fixed to the nominal value [13].

We can now study more in detail what are the main sources of the decay-time resolution and what is the best strategy to correct for them. The decay-time uncertainty is due to the uncertainty on the flight distance measurement, σ_L, and on the momentum determination, σ_P

$$\sigma_t^2 = \left(\frac{m}{p}\right)^2 \sigma_L^2 + \left(\frac{t}{p}\right)^2 \sigma_P^2. \tag{3.7}$$

The resolution of the decay length depends on the production and decay vertex resolutions of the b-meson, while the resolution of the reconstructed mass and of the momentum depend on the momentum resolution of the particles in the final state. Fig. 3.3 shows, as an example, the decay-time resolution obtained from the simulated $B_s^0 \to J/\psi\pi^+\pi^-$ sample simply comparing the so-called true and the reconstructed B_s^0 meson decay-times. The true decay-time is obtained by using the generated values for all the quantities in Eq. 3.1. In order to isolate the dominant contribution to the decay-time resolution, the other plots in Fig. 3.3 show the resolution obtained when using in turn the reconstructed value of the momentum, of the PV position and of the

DV position, while using the generated information for the other inputs. It is possible to see that the decay-time resolution is dominated by the contribution due to the secondary vertex resolution, in particular due to the resolution along the collision axis

$$\sigma_t \sim \frac{m}{p}\sigma_L. \tag{3.8}$$

It is important to note that not all of the b-mesons have the same decay-time resolution. Let's consider two mesons with the same decay time, but with different momentum: the particle with higher momentum flies more than the other, leading to a smaller decay-time relative uncertainty. Furthermore, b-mesons that originate from a pp collision, where a lot of particles are created, have a smaller PV uncertainty with respect to the case in which less particles are produced in the pp collision. These examples show clearly that, to obtain a precise determination of ϕ_s, the best strategy is to consider the effect of a per-event decay-time resolution. A first approximate estimation of a per-event decay-time resolution can be obtained from the b-meson decay-time uncertainty, achievable performing a DTF fit of its decay chain. As we will see in detail in Chaps. 4 and 5, this quantity has to be calibrated in order to get the exact decay-time resolution effect that could then be included in the final fit of the experimental decay rate distributions.

3.1.2 Decay-time and angular acceptances

As anticipated in the first part of this chapter, the particle reconstruction performed by the LHCb experiment and the analysis selection efficiencies generally depend on the decay-time of the b-meson and on the angular distributions of its decay products. It is worth noting that only the shape of these efficiencies is important, not the overall scale, since the latter can be absorbed in a overall normalisation term to which the measured parameters are insensitive. For this reason we will refer to any decay-time or angular efficiency as decay-time or angular acceptance, $Acc(t)$ or $Acc(\Omega)$. Chapters 4, 5 and 6 will give a detailed overview about the different sources of decay-time and angular acceptances affecting the decay modes under study. Here, the most common effects and the general strategies developed to correct for them are discussed.
The angular and decay-time acceptances are generally factorisable effects, so they can be studied separately. This assumption has been investigated in the $B_s^0 \to J/\psi K^+K^-$ analysis, see Chap. 4. As shown in Fig. 3.1 (c), typically the decay-time acceptance behaviour is different at low and high decay-time values, due to the different sources that contribute in the two ranges. The acceptance at low decay-time is generally caused by trigger requirements used to select the decays of interest produced in the pp collisions. The HLT lines often require the separation of the b-meson decay vertex from the PV, in order to reject the combinatorial background events that are mainly created from tracks coming from the PV. A b-meson with a short decay time decays near the PV and therefore do not pass these trigger requirements. The acceptance at high values is instead mainly due to the pattern recognition algorithm used to find tracks in the VELO. As briefly described in Section 2.4.3.1, this algorithm is based on the assumption that the tracks

originate from the interaction point. Thus, tracks which are significantly displaced from the PV are reconstructed less efficiently. If a b-meson flies significantly before decaying, its decay products are more displaced from the PV and are more easily rejected from the reconstruction algorithm, leading to the observed drop of the acceptance at high value of the decay-time. Considering that the track reconstruction is already performed in the high level trigger, the created acceptance can not be removed offline and has to be studied during the analysis.

The sources of the angular acceptances are more difficult to understand looking directly to their shapes, since the helicity angles have a complex definition. They are mostly due to the geometrical acceptance of the LHCb detector and to the implicit selection on particle momentum made by the magnet.

In general the acceptance effects can be studied using simulated MC samples, properly corrected to describe the data as much as possible. This strategy is followed for the determination of the angular acceptances. However, the usage of these samples has some limitations, because they do not exactly reproduce all the effects present in the real data collected by the LHCb experiment. If these effects are big and/or the variables of interest are really sensitive to them, a common and better strategy is to employ a data driven method. This approach is used to treat the decay-time acceptances in the analyses described in this thesis. A second decay channel is commonly used, the so-called control mode, topologically very similar to the main decay under study and with a very well known decay-time distribution. Indeed, if the decay-time resolution is properly taken into account, the decay-time acceptance of the control mode, $Acc_{control}$, can be evaluated, since any deviation from the well known decay-time distribution is due to the decay-time acceptance. Thanks to the topological similarity, the control mode decay-time acceptance is a very good approximation of the one of our decay of interest, Acc_{main}, allowing its determination. In addition, small differences between control and main decay channels (R) could be studied in MC simulation and taken into account in the acceptance determination as a small correcting factor

$$Acc_{main}(t) = Acc_{control}(t) \times R_{MC}(t). \tag{3.9}$$

3.1.3 Decay-time bias

During the injection of protons inside the LHC, the two halves of the VELO are moved along the x-axis of the experiment, in order to protect the detector until the beams reach stable conditions. They are then closed to take data during collisions and an online alignment is performed at each closure. In order to monitor the quality of the alignment, the PV position determined using information coming from opposite sides of the VELO can be used. The left and the right sides or the forward and the backward ones are exploited. If there were a misalignment in the x-axis, Δx, the PV resulting by using the left or the right side of the detector would be different, as shown in the left plot of Fig. 3.4. The presence of a misalignment along the x-axis would affect also the determination of the PV position along the z-axis, as it is visible in the right plot of

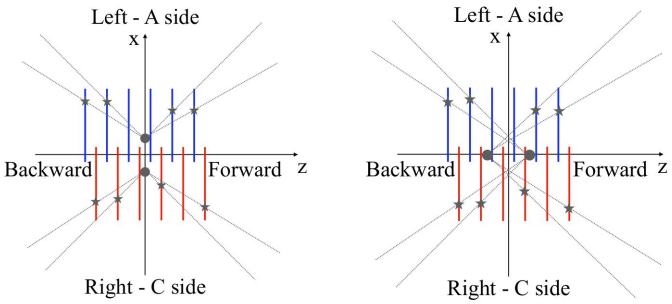

FIGURE 3.4. Reconstruction of the same primary vertex with tracks flying in the opposite sides of the detector. Figure taken [93].

Fig. 3.4, when only the forward or backward tracks are used for the PV determination. The bias along z, Δz, depends on the angle θ between the tracks and the beam axis and it can be written as

$$\Delta z = \frac{\Delta x}{\tan\theta}. \tag{3.10}$$

Considering that the PV is determined by using all the VELO tracks coming from the pp collision, the presence of a possible misalignment on both sides would not affect sensibly its position. On the contrary, the decay vertex of a b-meson is determined using only forward tracks and therefore its position could be biased. This effect can be immediately translated into a bias of the measured decay time and could have an important impact on the determination of CP or mixing parameters.

Especially in 2017 and 2018 data taking periods, a not negligible VELO misalignment has been observed [94]. Two main strategies can be used to account for this effect. A first estimation of the decay-time bias can be obtained from a sample of fake b-meson decays, where the decay products come from the PV. These candidates have, by definition, a *true* decay time equal to zero. Therefore, if the mean of their decay-time distribution is different from zero, it could be used as an estimation of the decay-time bias. Anyway, this value has to be calibrated in order to take into account the different phase space between the real signal decay and the fake one. For this purpose, samples of simulated signal and fake events are produced, where the same known misalignments of the VELO are introduced.

Another strategy is to update the parameters used to align the VELO, extracting the corrected ones from the difference in the biased PV positions obtained using only the left or the right tracks. Then, the data could be reprocessed offline using the correct parameters, giving directly the unbiased values of decay-time. This method requires a big computational effort, because the entire reconstruction procedure, normally performed online by the LHCb experiment, has to be rerun for all the collision events collected in the affected years.

3.2 Flavour tagging

In order to measure mixing-related *CP*-violating observables in neutral-meson systems, a crucial aspect is the determination of the flavour of the b-meson at production. This procedure is called *tagging*. As discussed in Chap. 1, the decay rate of a b-meson depends on the initial flavour of the meson, i.e. if it has been produced as a B or a $\overline{B}$. Because of the mixing phenomenology in the neutral systems, the flavour at production could be different with respect to the flavour at the decay. This aspect adds a new stage in the complexity of these types of analyses and, therefore, LHCb has developed some dedicated tools able to provide, for each reconstructed b-meson decay, its nature at production and the probability that this estimation is wrong. Major details related to the tagging algorithms and their performances could be found in Refs. [95], [96], [97] and [98].

The LHCb flavour tagging tools are based on two taggers categories, represented in Fig. 3.5.

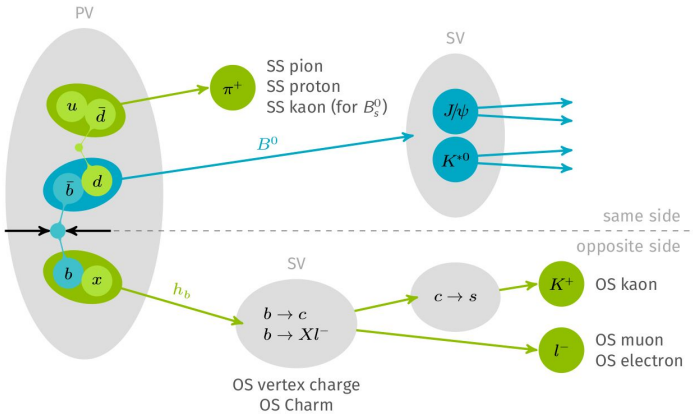

FIGURE 3.5. Schematic representation of the FT algorithms available at LHCb: the top half contains the signal B^0 decay and the same-side (SS) tagger, the bottom half shows the opposite-side (OS) taggers. Figure taken from Ref. [98].

The categories are called opposite-side (OS) taggers and the same-side (SS) taggers. Both the algorithms look for a specific type of particle generated in the pp collision, which has high probability to be correlated in charge with the flavour of the b-meson of interest, referred as *signal*. The OS taggers exploit the fact that b quarks are predominantly produced in $b\bar{b}$ pairs and try to infer the flavour at production of the signal b-meson looking for its correlation with the charge of the particles originated from the decay of the respective opposite-side b-meson. In particular, the algorithm is optimised to search for a kaon (OSK) or a lepton (OSμ or OSe) coming from a $b \to c \to s$ decay chain or for a lepton coming from a semileptonic b decay. In addition, the OS$_{vtx}$ tagger exploits the charge of the inclusive secondary vertex, reconstructed from the opposite b-hadron decay products. Since the hadronisation and the decay of the opposite-side b-meson is independent of the signal B meson, these algorithms can be used for

both B_d^0 and B_s^0. On the other hand, the SS taggers exploit the information related to light mesons, often produced in association to the hadronisation of the signal b-meson. In particular, they search for a pion (SSπ) or a proton (SSp) in case of a signal B_d^0 meson, and for a kaon (SSK) if the signal of interest is a B_s^0 meson.
Both the taggers types try to determine the flavour of the b-meson using multivariate algorithms, trained on flavour specific decays, where the flavour at decay is uniquely defined by the flavour of the decay products. In particular, reconstructed decays of $B^+ \to J/\psi K^+$ are used to train the OS taggers directly on data, while simulated events of $B_s^0 \to D_s^- \pi^+$ are exploited in the SSK tagger. The fast oscillations of the B_s^0 meson makes the training on data practically impossible. Several kinematical and geometrical information on the b-mesons and the tagging particles are used in the training. The output of these algorithms is transformed into a tagging decision q with values 1, 0 or -1 depending if the meson is tagged as B, untagged or tagged as $\overline{B}$.
The performance of the tagging algorithms is a crucial point in the ϕ_s analyses because, as it happens for the decay-time resolution, it causes a dilution of the b-meson oscillations. To study the impact of the tagging performance in the ϕ_s determination we can look at the time-dependent *CP* asymmetry defined in Eq. 1.57. Considering that $\Delta\Gamma$ is small, it can be simplified as

$$\mathcal{A}_{CP}^{true}(t) \propto \sin(\phi_s)\sin(\Delta m_s t). \tag{3.11}$$

Let us consider a number N of reconstructed signal b-meson decays. Because of the non perfect performance of the tagging algorithms, only N' mesons are tagged, while the other N_U remain untagged. In addition, not all the taken tagging decisions are correct and, among the tagged mesons, N_R are tagged correctly while N_W are not. Then the tagging efficiency ϵ_{tag} can be defined as

$$\epsilon_{tag} = \frac{N_R + N_W}{N}, \tag{3.12}$$

and the probability that the tag decision is wrong, called mistag probability, is

$$\omega = \frac{N_W}{N_R + N_W}. \tag{3.13}$$

Considering the application of flavour tagging algorithms, the measured number of mesons initially produced as B and $\overline{B}$ are then

$$N_B(t) = (1-\omega)N_B^{true}(t) + \omega N_{\overline{B}}^{true}, \tag{3.14}$$

$$N_{\overline{B}}(t) = (1-\omega)N_{\overline{B}}^{true}(t) + \omega N_B^{true}, \tag{3.15}$$

where N_B^{true} and $N_{\overline{B}}^{true}$ are the real number of reconstructed B and $\overline{B}$ mesons. The consequent measured *CP* asymmetry is

$$\mathcal{A}_{CP} = \frac{N_B(t) - N_{\overline{B}}(t)}{N_B(t) + N_{\overline{B}}(t)} = (1-2\omega)\mathcal{A}_{CP}^{true}. \tag{3.16}$$

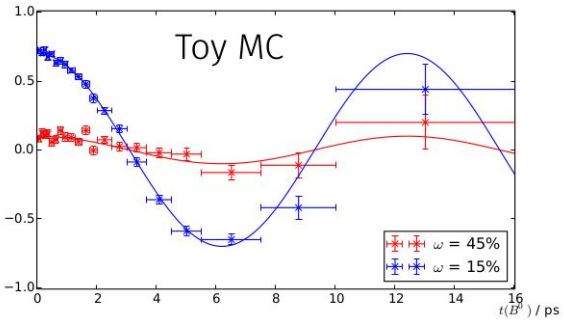

FIGURE 3.6. Changing of the *CP* asymmetry as a function of different mistag probabilities. Figure taken from Ref. [99].

The amplitude of the measured asymmetry is decreased proportionally to a dilution term $\mathcal{D} = (1 - 2\omega)$ directly depending on the mistag probability ω, as shown in Fig. 3.6. If $\omega = 0$ (perfect tagger), $\mathcal{D} = 1$ and the true asymmetry is not diluted by the tagging process. If $\omega = 0.5$ (random tagger), $\mathcal{D} = 0$ and it is not possible to measure the asymmetry and ϕ_s. Greater the dilution effect, smaller is the precision of the ϕ_s measurement. In fact, supposing that ω is perfectly known, the statistical uncertainty on the true *CP* asymmetry is

$$\mathcal{A}_{CP}^{true} = \frac{\mathcal{A}_{CP}^{meas}}{\mathcal{D}}, \quad \sigma_{\mathcal{A}_{CP}^{true}} \propto \frac{1}{\sqrt{\epsilon_{tag} N \mathcal{D}}}. \tag{3.17}$$

This means that the effective size of a tagged sample is reduced by the factor $\mathcal{D}^2$. The combined quantity $\epsilon_{tag}\mathcal{D}^2$ is called the effective tagging power, ϵ_{eff}, and it is a measurement of the tagging performance, being used as a figure of merit to be maximised during the training and development of the flavour tagging algorithms.

While the mistag probability ω can be easily determined for charged b-mesons samples, when the flavour tagging algorithms are applied to non-flavour specific decays, such as $B_s^0 \to J/\psi h^+ h^-$ decays, it is not possible to measure directly the mistag probability and the algorithms provide an event-by-event estimated mistag probability η, defined in a range [0, 0.5]. Since the tagging performance depends on the kinematics of the signal b-decays and on the requirements applied selecting the flavour specific decays used in the training, it is important to calibrate the predicted mistag rate to obtain the true mistag rate ω. This calibration will be shown in Chap. 4.

CHAPTER 4

Measurement of CP-violating phase ϕ_s with $B_s^0 \to J/\psi K^+K^-$ decays

In this chapter the ongoing measurement of the CP-violating phase ϕ_s in the golden channel $B_s^0 \to J/\psi K^+K^-$ is presented, exploiting the events collected by the LHCb experiment during the data taking period 2015-2018. As already mentioned in the previous chapters, this measurement requires a flavour-tagged time-dependent and angular analysis, with a fundamental focus on the correction of the resolution, acceptance and tagging effects.
After a brief summary of the current state of the art of ϕ_s, $\Delta\Gamma_s$ and Γ_s measurements, the candidates selection and the study of the experimental effects are explained in detail. Finally the preliminary results and the most important sources of systematic uncertainties will be shown.
The analysis presented in this thesis, considering its importance and its difficulty, is performed by a rather large group of scientists inside the LHCb Collaboration and is currently being summarised in the analysis note *"Measurement of $B_s^0 \to J/\psi K^+K^-$ with the full Run 2 data"* [7]. For completeness, all the analysis steps are presented in this thesis, while more details are given for the aspects to which I contributed the most. In particular, I am responsible for the development of the whole selection procedure of the $B_d^0 \to J/\psi K^0(892)^*$ and $B^+ \to J/\psi K^+$ control modes, in both data and simulation, essentials for the measurement of the decay-time acceptance and for the calibration of the OS tagging algorithms, respectively. The $B_d^0 \to J/\psi K^0(892)^*$ is besides used to validate the angular acceptance correction procedure.

4.1 Current status of ϕ_s, $\Delta\Gamma_s$ and Γ_s measurements

The LHCb experiment plays a key role in the measurement of the *CP*-violating phase ϕ_s, the decay width Γ_s and the decay width difference $\Delta\Gamma_s$, thanks also to the large number of decay modes used: $B_s^0 \to J/\psi K^+K^-$ with K^+K^- mass region around the $\phi(1020)$ [100], [5], $B_s^0 \to J/\psi K^+K^-$ in the mass region above the $\phi(1020)$ [101], $B_s^0 \to \psi(2S)\phi$ [102], $B_s^0 \to J/\psi\pi^+\pi^-$ [103], [6] and $B_s^0 \to D_s^+D_s^-$ [104]. Figure 4.1 shows the 68% confidence level regions in the ϕ_s versus $\Delta\Gamma_s$ plane for all the analyses, as well as the result of a combined fit together with the SM predictions. The luminosity values of 3 fb^{-1} and 4.9 fb^{-1} refer respectively to the Run 1 only and Run 1 plus 2015-2016 data taking periods. Note that in the $B_s^0 \to J/\psi\pi^+\pi^-$ and $B_s^0 \to D_s^+D_s^-$ analyses $\Delta\Gamma_s$ is not measured, then they are represented by vertical bands. The LHCb combined values are [5]

$$\phi_s^{LHCb} = -0.041 \pm 0.025 \text{ rad}, \tag{4.1}$$

$$\Gamma_s^{LHCb} = 0.6563 \pm 0.0021 \text{ ps}^{-1}, \tag{4.2}$$

$$\Delta\Gamma_s^{LHCb} = 0.0813 \pm 0.0048 \text{ ps}^{-1}. \tag{4.3}$$

The Heavy Flavour Averaging Group (HFLAV) [2] combined the LHCb combination results

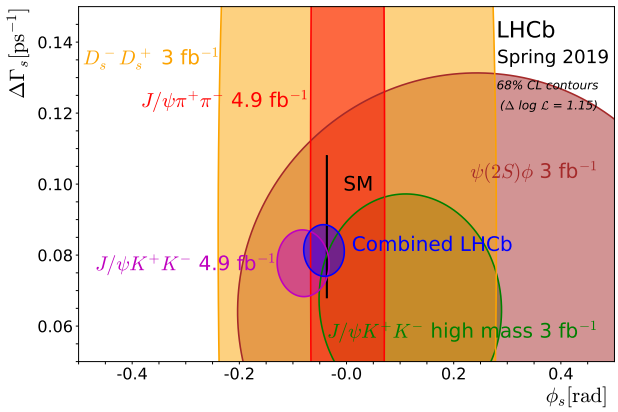

FIGURE 4.1. 68% confidence level regions in the ϕ_s versus $\Delta\Gamma_s$ plane for all the analyses performed by the LHCb experiment, as well as the result of a combined fit and the SM predictions, indicated by the thin black rectangle. The luminosities 3 fb^{-1} and 4.9 fb^{-1} refer respectively to the Run 1 only and Run 1 plus 2015-2016 data taking periods.

with analyses of $B_s^0 \to J/\psi K^+K^-$ performed by the CDF [105], D0 [106], ATLAS [107], [108], [109] and CMS [110], [111] experiments, taking into account the correlation matrix between all

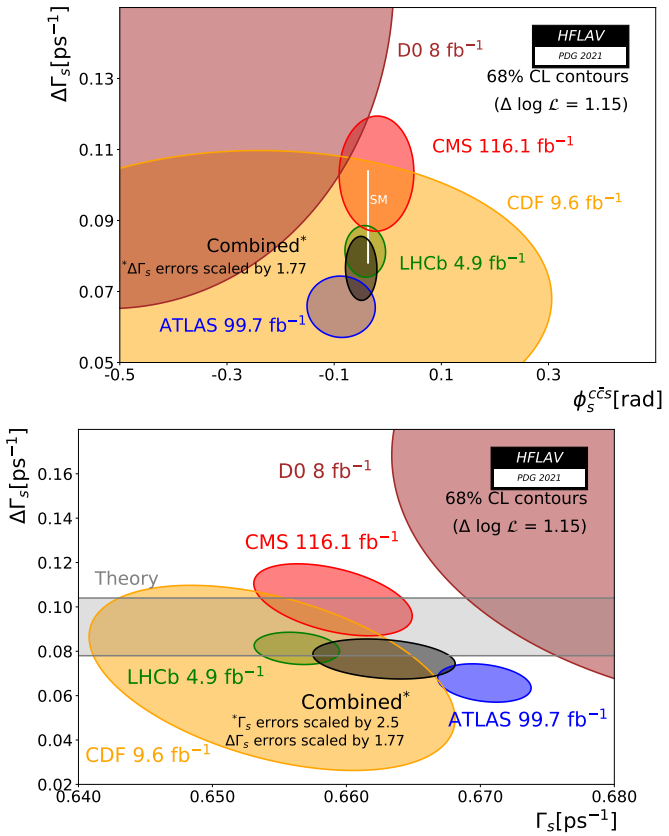

FIGURE 4.2. 68% confidence level regions in the ϕ_s vs $\Delta\Gamma_s$ plain on top and Γ_s vs $\Delta\Gamma_s$ plain on bottom for the different experiments, as well as the combined contour. The amount of collected data per experiment is shown, together with the ϕ_s and $\Delta\Gamma_s$ SM predictions, indicated by the thin white rectangle in the top plot and with the grey one in the bottom plot. Plots taken from Ref. [2].

physics parameters in each analysis. Fig 4.2 shows the 68% confidence level regions in the ϕ_s vs $\Delta\Gamma_s$ plane on top and Γ_s vs $\Delta\Gamma_s$ plane on bottom for the different experiments, as well as the combined contour. The amount of collected data per experiment is shown, together with the ϕ_s and $\Delta\Gamma_s$ SM predictions, indicated by the thin white rectangle in the top plot and with the grey one in the bottom one. The prediction for ϕ_s is taken as the indirect determination of $-2\beta_s$ via a global fit to experimental data within the SM, $-2\beta_s = -0.0370^{+0.0007}_{-0.0008}$ rad [3], while the SM prediction for $\Delta\Gamma_s$ is 0.091 ± 0.013 ps^{-1} [4]. The theoretical prediction based on the SM of Γ_s is given in comparison to Γ_d, $\Gamma_s/\Gamma_d = 0.999 \pm 0.002$ [112], since in the ratio many theoretical uncertainties cancel.

The combined measured results are

$$\phi_s^{HFLAV} = -0.050 \pm 0.019 \text{ rad}, \tag{4.4}$$

$$\Delta\Gamma_s^{HFLAV} = +0.082 \pm 0.005 \text{ ps}^{-1}, \tag{4.5}$$

$$\Gamma_s^{HFLAV} = 0.6597 \pm 0.0026 \text{ ps}^{-1}, \tag{4.6}$$

that are consistent with SM predictions. The LHCb experiment provides the most precise combined measurement of ϕ_s, $\Delta\Gamma_s$ and Γ_s up to date. It is interesting to note that the value of Γ_s shows an intriguing tension between the two most precise experiments, LHCb and ATLAS, of about 4σ. Further measurements from both the experiments are thus needed to understand if the discrepancy is due to statistical fluctuations or not.

4.2 Selection of $B_s^0 \to J/\psi K^+K^-$ signal candidates

The B_s^0 does not release hits in the LHCb detector and its identification is performed by the reconstruction of its decay products.
Figure 4.3 shows the topology of a $B_s^0 \to J/\psi K^+K^-$ decay. The B_s^0 mesons are produced

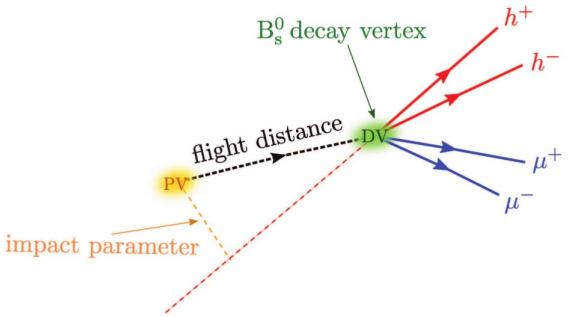

FIGURE 4.3. Schematic view of the $B_s^0 \to J/\psi K^+K^-$ decay topology. The signal meson is produced at the PV and the intermediate resonances J/ψ and ϕ immediately decay at the DV into the final state particles. Scheme modified from Ref. [1].

from the hadronisation of a b or $\bar{b}$ quark created in the pp interaction. They typically travel for a flight distance of few centimetres before decaying into a J/ψ and a K^+K^- pair. The J/ψ, in this particular decay chain, decays almost instantaneously into a pair of oppositely charged muons with a branching fraction of $BR(J/\psi \to \mu^+\mu^-) = (5.961 \pm 0.033)$ % [13]. The K^+K^- pair comes mainly from the ϕ resonance, with a branching fraction of (49.2 ± 0.5) % [13], plus a non resonant component. The total branching fraction of $B_s^0 \to J/\psi K^+K^-$ decay, with $J/\psi \to \mu^+\mu^-$ and K^+K^- coming from the $\phi(1020)$, is about $3 \cdot 10^{-5}$ %.
In a pp interaction a lot of particles are produced, mainly pions and other light particles. Then the first step of the analysis is the reconstruction and the selection of $B_s^0 \to J/\psi K^+K^-$ decays,

referred as signal candidates, among all the particles produced in the collision. For this purpose a selection strategy has been developed and optimised in order to obtain an almost pure sample of $B_s^0 \to J/\psi K^+K^-$ decays without loosing too much statistics in the process.
Four main steps are needed:

- a trigger selection, discussed in Section 2.3.4, centrally performed by the LHCb experiment;
- a loose preselection specifically developed for this analysis, able to separate the signal candidates from b-hadron events with similar topology;
- a selection that exploits a multivariate analysis technique, in order to reduce at a very low level the background contribution;
- a statistical subtraction of the remain background candidates that pollute the selected signal sample.

In this thesis, the data collected by the LHCb experiment during the Run 2 data taking period at the center of mass energy of 13 TeV is analysed, corresponding to 6 fb^{-1} of integrated luminosity. In addition, samples of simulated $B_s^0 \to J/\psi K^+K^-$ events are selected using the same strategy as for the signal data. These samples are used to validate the analysis steps and to correct for some experimental effects.

4.2.1 Trigger selection

The selection procedure starts searching for those *pp* collision events in which two oppositely charged muons and/or two oppositely charged kaons are produced, reconstructing the trajectories of all particles in the detector. The L0 trigger requirements are typically rather loose compared to those applied in the succeeding steps and they do not introduce any acceptance effect. For this reason, no specific L0 trigger lines have been required and all events triggering at least one L0 line are accepted. The first selection is required at the HLT1 stage. An event is accepted if it fires as TOS at least one of the following three HLT1 lines: **Hlt1DiMuonHighMass**, **Hlt1TrackMuon** or **Hlt1TwoTrackMVA**. The **Hlt1DiMuonHighMass** line selects events with two well-identified oppositely charged muons with a good quality vertex and an invariant mass larger than 2700 MeV/c^2, rejecting all the light resonances that decay into two muons. Considering that the B_s^0 meson flies on average few centimetres before decaying, its decay products do not come from the PV and have on average a high impact parameter. The **Hlt1TrackMuon** and **Hlt1TwoTrackMVA** requirements further exploit these features. The first one selects events that contain at least one muon with transverse momentum larger than 1 GeV/c and with a large impact parameter χ^2 with respect to all reconstructed PVs in the collision, while the second exploits a multivariate algorithm to identify two charged particles with a large sum of tranverse momentum and a displaced common vertex with respect to any PV. The last two trigger lines tend to select B_s^0 candidates with a large decay time, because

of the requirements on the displacement of the decay products from the PV. This introduces a non uniform experimental effect on the decay-time acceptance, not present for those events that are selected only by the first HLT1 line. For this reason the events are classified into two different categories as a function of the HLT1 selections:

- BIASED: if they are selected with al least one the two HLT1 lines **Hlt1TrackMuon** and **Hlt1TwoTrackMVA**
- UNBIASED: if they are selected only by the **Hlt1DiMuonHighMass** line.

These two categories are treated separately in the analysis, since they need different decay-time acceptance corrections.
During the HLT2 trigger stage an almost complete reconstruction of the events is made, allowing to apply selection requirements directly on the reconstructed J/ψ, significantly improving the separation between $B_s^0 \to J/\psi K^+K^-$ signal candidates and background events. The `HLT2DiMuonDetachedJPsi` is used, which requires selected events containing a muon pair with invariant mass within ± 120 MeV/c^2 of the nominal J/ψ mass value [13]. Its decay vertex is required to be significantly desplaced from any PV, to suppress J/ψ mesons that are directly produced in the primary proton-proton collision. For that reason the J/ψ flight distance (FD) significance has to be grater then 3, $FD/\sigma_{FD} > 3$, and if multiple PVs are reconstructed in the event, the PV with the minimum value of χ^2_{IP} is considered. The selection of these *detached* candidates introduces a further not uniform acceptance as a function of the decay time.

4.2.2 Offline selection

The signal candidates passing the trigger selection are saved and the entire $B_s^0 \to J/\psi K^+K^-$ decay chain is re-reconstructed offline by the BRUNEL software [78]. Considering that the trigger selection has been chosen in order to minimise the acceptance effects on the decay time, the selection at the trigger stage is not too tight and a lot of background candidates are still retained together with the signal. Thus, a further selection on the reconstructed decay chain is applied, in order to increase the signal to noise ratio. The candidates are required to pass a specific stripping selection, whose requirements are summarised in Table 4.1. The candidates passing the offline selection must have two muons with transverse momentum greater the 500 MeV/c forming a J/ψ candidate with an invariant mass that lies in a ± 80 MeV/c^2 window around the nominal mass [13]. In order to avoid the creation of a J/ψ candidates from particles wrongly identified as muons, a loose cut on the difference of the logarithmic likelihood for the muon versus the pion hypothesis is applied. The J/ψ candidate must have a good quality vertex, with a $\chi^2_{vtx}/nDof < 16$ and a small significance of the distance of closest approach of the two muon tracks, $\chi^2_{DOCA} < 20$. These requirements allow to reject fake J/ψ candidates created with random muons. A similar strategy is used to select the ϕ candidate. Then the J/ψ and ϕ candidates are combined to form a B_s^0 meson that must fulfil the following requirements: a good

Table 4.1: Stripping selection criteria used to identify $B_s^0 \to J/\psi\phi$ candidates.

	Variable	Stripping
$J/\psi \to \mu^+\mu^-$	$\Delta\ln\mathcal{L}_{\mu\pi}$ $(\mu^\pm)$	> 0
	p_T $(\mu^\pm)$	$> 500\,\mathrm{MeV/c}$
	χ^2_{DOCA}	< 20
	$\chi^2_{vtx}/nDoF$	< 16
	$m(\mu^+\mu^-)$	$\in [3016.9,\ 3176.9]\,\mathrm{MeV/c^2}$
$\phi \to K^+K^-$	χ^2_{DOCA}	< 30
	p_T (ϕ)	$> 500\,\mathrm{MeV/c}$
	$m(K^+K^-)$	$\in [980,\ 1060]\,\mathrm{MeV/c^2}$
	$\chi^2_{vtx}/nDoF$	< 25
	$\Delta\ln\mathcal{L}_{K\pi}$ (K^+)	> 0
	$\chi^2_{track}/nDoF$ $(K^\pm)$	< 5
$B_s^0 \to J/\psi\phi$	$m(J/\psi K^+K^-)$	$\in [5150,\ 5550]\mathrm{MeV/c^2}$
	$\chi^2_{vtx}/nDoF$	< 20
	t	> 0.2 ps

quality vertex, an invariant mass in the region $(5150, 5550)\ \mathrm{MeV/c^2}$. These candidates are then associated to one of the PVs of the pp collision. First, to avoid any bias on the PV position, all PVs are reconstructed again after removing the tracks that are used to build the B_s^0 candidate. Then the B_s^0 meson is associated to the PV for which it has the smallest χ^2_{IP}. In order to improve the mass, momentum and vertex resolution the decay chain is then fitted using the DTF algorithm, presented in Chap. 3, including additional constraints on the mass of the J/ψ meson, constrained to the world average [13], and on the B_s^0 meson momentum direction, required to point to the associated PV. A selection on the B_s^0 decay time, the decay-time error and the B_s^0 invariant mass obtained by the DTF fit is further used to isolate the signal candidates from the background. Indeed, in addition to the stripping selection, the B_s^0 decay time is required to be in the range $0.3 - 15$ ps. The lower bound is set in order to suppress the combinatorial background and to remove any boundary effect from the difference with the selection present in the stripping, where the PV is determined without the removal of the B_s^0 decay product tracks, while the upper limit is set because no significant amount of true B_s^0 meson decays are expected beyond this value. The decay-time error δ_t is required to be less than 0.15 ps, removing candidates with very poor determination of the decay time. Additional cuts on the invariant masses of (K^+K^-) and B_s^0 are finally applied, in order to further reject the combinatorial background: $5200 < m_{DTF}(B_s^0) < 5550\,\mathrm{MeV/c^2}$ and $990 < m(K^+K^-) < 1050\,\mathrm{MeV/c^2}$. The requirement made on the K^+K^- invariant mass, that has to lie around the nominal ϕ mass [13], does not guarantee the absence of a not resonant contribution that, as already explained in Chap. 1, has to be taken into account in the decay rate description. Finally, the momenta obtained by the DTF algorithm are used to calculate the helicity angles, improving the angular resolution.

4.2.3 Simulated sample

As anticipated in the introduction to this chapter, the simulated Monte Carlo (MC) samples have a fundamental role in the analysis because they are used in the selection, as it will be shown in the next section, in the validation of the analysis procedure and to extract the corrections to some experimental effects that distort the decay-time distribution.
Different types of MC samples are produced, using the LHCb software framework described in Section 2.4.1. Namely, samples of $B_s^0 \to J/\psi\phi$ with $\Delta\Gamma_s$ fixed to the nominal value or $\Delta\Gamma_s = 0$ and samples of $B_s^0 \to J/\psi K^+K^-$ events which contain K^+K^- resonant contribution from both the ϕ and f_0 mesons. The last sample is used to study the S-wave contribution in data. Each simulated sample is produced per year and per magnet polarity, using specific data-taking conditions that mimic as much as possible the real conditions. The decay model parameter values used in the simulation are summarised in Table 4.2.

Table 4.2: Decay model parameters for $B_s^0 \to J/\psi K^+K^-$ MC samples.

Parameter	EventType			
	$B_s^0 \to J/\psi\phi$	$\Delta\Gamma_s = 0$	$\Delta\Gamma_s = 0$ updated	$B_s^0 \to J/\psi K^+K^-$
Δm_s []	17.8	17.8	17.8	17.8
$\Delta\Gamma_s$ [ps^{-1}]	0.08543	0	0	0.08543
Γ_s [ps^{-1}]	0.6614	0.6614	0.6614	0.6614
ϕ_s [rad]	-0.03	0.07	-0.03	-0.03
$\lvert A_0(0)\rvert^2$	0.5242	0.5213	0.5242	0.5242
$\lvert A_\parallel(0)\rvert^2$	0.2256	0.2304	0.2256	0.2256
$\lvert A_\perp(0)\rvert^2$	0.2500	0.2490	0.2500	0.2500
$\lvert A_S(0)\rvert^2$	-	-	-	0.07
$\delta_\parallel - \delta_0$ [rad]	3.26	3.30	3.26	3.26
$\delta_\perp - \delta_0$ [rad]	3.08	3.07	3.08	3.08

The MC samples are selected with the same trigger and offline selection developed for the data samples. One of the advantages of MC samples is the possibility to identify the particles whose tracks are only partially reconstructed, or so-called ghosts, i.e. tracks formed by a random combination of hits in the detector. A dedicated procedure, so-called matching, was developed in LHCb in order to associate, at decay chain level, the reconstructed particles and the particles at the generator level. The hits that are in common between a reconstructed track and those produced by the generated particle are compared and the information on the matching is propagated from the final state tracks to the intermediate reconstructed particles. The BKGCAT variable classifies the particles based on the results of the matching, allowing to study the different types of tracks and to select only the B_s^0 MC candidates for which the reconstructed particles match the generated ones at a good level. This requirement is called truth matching and in this analysis only candidates that belong to the following categories are selected:

- BKGCAT = 0: the decay chain in generation is identical to the reconstructed one, meaning

that all the daughters and intermediate particles are correctly matched;

- BKGCAT = 50: it corresponds to BKGCAT=0 candidates with in addition the emission of photons due to radiative energy loss.

There are other two categories used during the analysis:

- BKGCAT = 10: the final state particles are properly matched but there are different intermediate states between the generated and reconstructed decay chain;
- BKGCAT = 60: that is populated from candidates formed by at least one final-state ghost particle.

It happens that sometimes signal candidates are erroneously categorised as ghosts, because of a non perfect matching performance. This is visible due to the fact that some of the candidates with BKGCAT=60 peaks at the nominal B_s^0 mass. Typically this happens for tracks that are badly reconstructed and thus also the invariant mass distribution shows a worse resolution with respect to the BKGCAT=0 or BKGCAT=50 events. So, to proper estimate the acceptance effects, also the BKGCAT=60 MC candidates are considered.
The usage of the MC in the analysis is made under the assumption that it provides a good representation of the signal candidates in the data samples collected by the experiment. However, this is not always true, in particular for some kinematic, multiplicity and particle identification (PID) variables. In order to rely on the simulated samples, they have to be corrected to match the data in the variables of interest for the analysis.
The correction is divided into two steps. Initially the PID variables are corrected using a dedicated tool called `PIDCalib` package, where the PID responses in MC are corrected based on PID calibration data samples. For more details about the package and the procedure see Ref. [113]. Then the simulated samples are reweighted in order to match, in the variables showing the worst data-MC agreement, the background-subtracted data samples distributions. The correction is performed year per year, and it allows to obtain a good matching between MC and data samples.

4.2.4 Selection based on a multivariate analysis

Figure 4.4 shows the reconstructed B_s^0 invariant mass distribution in the larger data sample collected in 2018 after the trigger and offline selection. A well-defined peaking structure is visible, due to the selected signal candidates, almost Gaussian distributed around the nominal value of the mass, equal to $m(B_s^0) = 5366.88 \pm 0.14$ MeV/c^2 [13]. The signal mass shape is due to the detector resolution. Therefore, it takes the form of a Gaussian-like shape, with additional tails on the left and on the right side of the resolution core, due to poorly reconstructed events. In addition, together with the visible final states, photons can be radiated in the b-meson decay, leading to a larger tail on the left side of the mass peak. In the mass distribution, a

large background component is also present, causing the flat contribution. It is created from random combinations of muons and kaons not really coming from a real B_s^0 decay, resulting in random values of the reconstructed mass. A mass model composed by a double Gaussian function for the signal and an exponential function for the background is used in the fit to the mass distribution. In order to further reduce the combinatorial background and preserve a

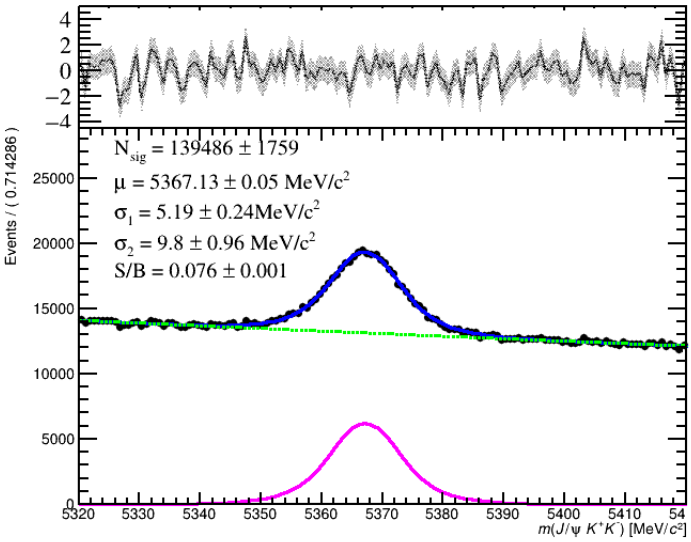

FIGURE 4.4. Distribution of reconstructed $m(J/\psi K^+K^-)$ after the stripping, trigger and offline selection in 2018. The blue solid line shows the total fit with the signal and combinatorial background described by the magenta and green lines, respectively.

high signal efficiency, the analysis makes use of a multivariate technique to further discriminate between signal and background. A gradient-boosted decision tree (BDT) with a gradient boosting has been used (see Appendix A for more details) that exploits all the correlations between kinematical and geometrical variables. The classifier is available in the Toolkit for Multivariate Analysis (TMVA) package [114]. The BDT is trained separately for each year, using the selected, truth-matched and corrected simulated sample as signal proxy, and the selected data right sideband as background proxy. The right sideband is defined by the candidates with $5450 < m(B_s^0) < 5550$ MeV/c^2, i.e. a mass region populated only by combinatorial background events. The signal and background samples have been separated into two subsamples, one for training and one for testing. Several variables that allow a good separation between signal and background distributions are used in the training procedure, namely the χ^2 of the track fit of the final-state particles, their identification probability, the quality of the J/ψ and B_s^0 decay vertices, the transverse momentum and pseudorapidity of the B_s^0 candidate, its DTF χ^2 and its IP with respect to the origin vertex and transverse momentum of the K^+K^- system. Variables such as the flight distance of the B_s^0 meson or the impact parameter of its final state particles

show a very high separation between signal and background but are strongly correlated to the decay time and would cause strong decay-time acceptance effects. Therefore, these variables are not used. The BDT is then applied to the testing sample to check for the presence of possible *overtraining**, that has been not found, and is finally applied to the whole data sample. The cut on the BDT output is chosen to maximise the figure of merit (FOM),

$$\text{FOM} = \frac{(\sum_i \omega_i)^2}{\sum_i \omega_i^2}, \tag{4.7}$$

where w_i is the weight assigned to each event by the *sPlot* [115] technique. This method is based on an extended maximum likelihood fit on the B_s^0 invariant mass distribution, called discriminating variable, to separate the signal and background contributions. From the fit, the *sPlot* method calculates per-candidate weights for each component, called sWeights. They can be used to build a weighted pure signal sample. For more details about the *sPlot* technique and the extended maximum likelihood fit see Appendix A. This figure of merit describes the number of effective signal candidates available for the analysis at the end of the selection procedure. So, the larger the FOM is, the smaller will be the uncertainties on the measured parameters.
In Fig 4.5, the mass distributions of the 2018 data sample are shown after applying the selection on the BDT response. Similar results are obtained for the other years.

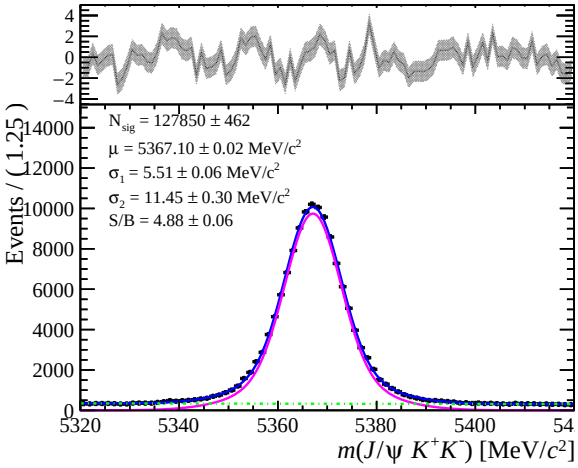

FIGURE 4.5. Distribution of reconstructed $m(J/\psi K^+K^-)$ after the BDT selection for 2018. The blue solid line shows the total fit with the signal and combinatorial background described by the red and green lines, respectively.

4.2.5 Misidentified background

The combinatorial background is not the only source of fake B_s^0 candidates that populate the selected data samples. Other b-mesons or b-hadrons that decay in a J/ψ plus two light particles

*The *overtraining* is explained in Appendix A

can be wrongly identified as a B_s^0 meson if their light decay products are misidentified as two kaons during the reconstruction. These fake B_s^0 candidates can populate the selected B_s^0 mass window and have to be removed in order not to affect the measurement of the parameters of interest. There are two main sources of misidentified background: $\Lambda_b^0 \to J/\psi pK$ decays and $B_d^0 \to J/\psi K^+\pi^-$ decays.

The procedure to estimate the level of peaking backgrounds in the $B_s^0 \to J/\psi K^+K^-$ sample is described in the following. The invariant mass of the selected candidates is recalculated giving to the two kaons different mass hypothesis. Thus, the mass of the light particles of the possible misidentified background are assigned to the kaon tracks in the B_s^0 sample, recomputing then the B_s^0 invariant mass. This procedure can be easily performed by the DTF algorithm, that allows to fit the entire decay chain with the wrong mass hypothesis assigned to the kaon tracks. If events of the supposed misidentified background are present in the sample, they will compare in the recalculated mass as a peak around the nominal mass of the misidentified b-meson or b-hadron.

The decay $B_d^0 \to J/\psi K^+\pi^-$ can be wrongly identified as a $B_s^0 \to J/\psi K^+K^-$ decay if the pion is reconstructed as a kaon. In this case the lower mass of the B_d^0 meson with respect to the B_s^0 meson is partially compensated and it can populate the selected mass region. In order to

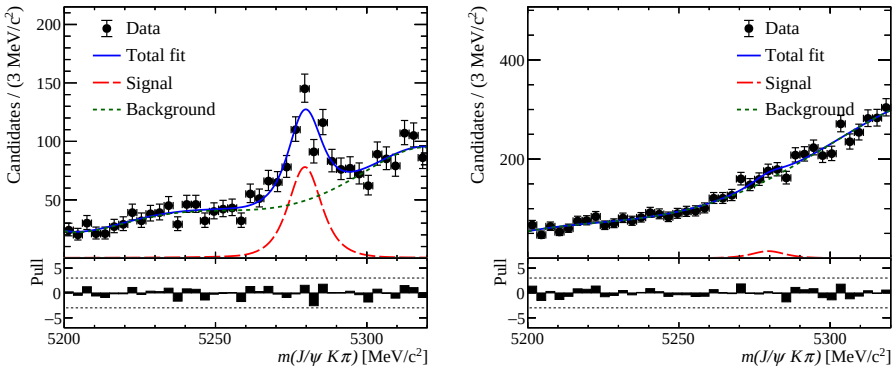

FIGURE 4.6. Distribution of recalculated $m(J/\psi K^+\pi^-)$ events after the BDT selection for 2018. The events removed by the vetoes and those that survive after it are shown on the left and on the right, respectively. The blue solid line shows the total fit with the signal and combinatorial background described by the red and green lines, respectively.

suppress the $B_d^0 \to J/\psi K^+\pi^-$ peaking background, the B_s^0 candidates satisfying the following criteria are rejected:

- the kaon with the larger ProbNNπ has ProbNN$\pi > 0.7$ or ProbNN$K < 0.35$, where the ProbNN variable has been introduced in Sec. 2.4.3.2;
- $m(J/\psi K^+\pi^-)$ is within ± 15 MeV/c^2 of the nominal B_d^0 mass [13].

The efficiency of the veto on the signal B_s^0 candidates is evaluated on MC and it is equal to $97.6 \pm 0.1\%$. The candidates in the $m(J/\psi K^+K^-)$ sideband region, $5387 - 5550$ MeV/c^2, provide a clean sample to estimate the size of this background, without the pollution due to the B_s^0 signal. The B_d^0 recalculated mass distribution in this mass region is shown in 4.6. The events removed by the vetoes and those that survive after it are shown on the left and on the right, respectively. The amount of remaining B_d^0 events after the veto is very small and it is considered to be negligible.
The $\Lambda_b \to J/\psi Kp$ background is treated with a similar strategy. It occurs when the proton is wrongly identified as a kaon. The misidentified candidates are rejected by the following vetoes:

- the kaon with the larger ProbNNp has ProbNN$p > 0.7$;
- interpreting the kaon with the higher ProbNNp as a proton, $m(J/\psi Kp)$ is within ± 15 MeV/c^2 of the Λ_b^0 nominal mass [13].

The efficiency of the veto on the B_s^0 signal candidates evaluated on MC is $98.1 \pm 0.1\%$. The yield of Λ_b^0 background events is higher compared to the B_d^0 one. In order to estimate the amount of Λ_b^0 candidates surviving the vetoes, they are divided into two categories accordingly to the the PID veto. Fig 4.7 shows the two categories and projections of fits from which the yields of $\Lambda_b^0 \to J/\psi pK$ candidates are determined. Only the $m(J/\psi K^+K^-)$ sideband regions are considered, $5200 - 5347$ MeV/c^2 and $5387 - 5550$ MeV/c^2, in order to limit the pollution due to the B_s^0 signal candidates. The background component is modelled by a fourth order

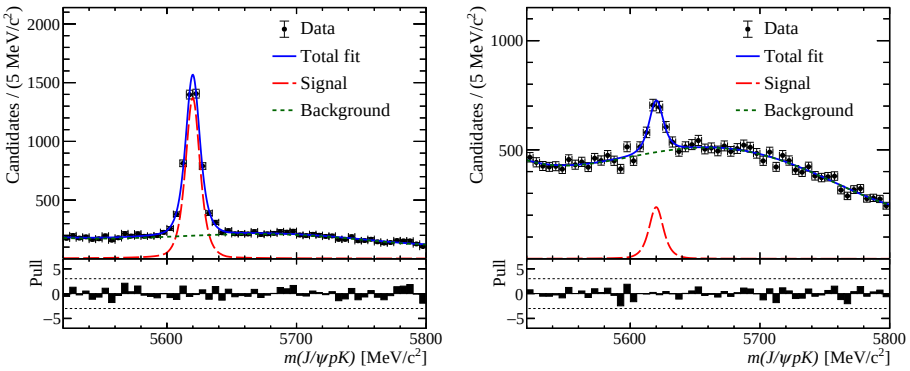

FIGURE 4.7. Distribution of recalculated $m(J/\psi Kp)$ events after the BDT selection for 2018. The sample is split according to this ProbNNp value being larger (left) or smaller (right) than 0.7. The blue solid line shows the total fit with the signal and combinatorial background described by the red and green lines, respectively.

polynomial and the signal is described by an Ipatia function [116] (discussed in more details in Appendix A), whose parameters are fixed from simulation. To extract the number of remaining Λ_b^0 candidates in the whole B_s^0 sample, the signal shape, the mass window of the veto and a simulation driven correction are used, where the latter accounts for the fact that only the B_s^0

mass sidebands are considered. The estimated number of Λ_b^0 events which survive the veto is summarised in Table 4.3. Considering the big amount of the remaining events and that the Λ_b^0

Table 4.3: The estimated number of Λ_b^0 events which survive veto for different year. The certainty is statistical uncertainty.

Year	The number of Λ_b^0
2015	289 ± 40
2016	1447 ± 93
2017	1309 ± 82
2018	1647 ± 92

has a very different decay time distribution with respect to the B_s^0, this residual background cannot be considered negligible and has to be further removed. The strategy will be presented in the next section.

4.2.6 Subtraction of the remaining background

After the BDT selection and the misidentified background suppression, the B_s^0 sample is still polluted by a small combinatorial background contribution and residual Λ_b^0 events. Both backgrounds must be totally removed, since their very different decay-time distributions with respect to the signal one can considerably distort the ϕ_s, Γ_s and $\Delta\Gamma_s$ measured values. First, the $\Lambda_b^0 \to J/\psi Kp$ background is removed. Two main strategies can be used. The first exploits the *sPlot* technique, using the B_s^0 invariant mass as a discriminating variable, to remove the Λ_b^0 events by looking at the sWeighted signal sample. The downside of this strategy is that the correlations between the invariant mass and the angular variables are not taken into account. To avoid this, a second method is used in which simulated Λ_b^0 candidates, properly corrected to match the Λ_b^0 events in data, are injected into the data sample with negative weights, such that the sum of the negative weights is equal to minus the number of the estimated Λ_b^0 events, shown in Table 4.3.

The *sPlot* technique is then used in order to subtract the combinatorial background. This allows to obtain background subtracted signal distributions of the helicity angles, the decay time and the tagging information that will be used to perform the final fit and measure the parameters of interest. The *sPlot* method requires the discriminating variable, $m(J/\psi K^+K^-)$, and the variables of interest to be uncorrelated. However, the mass resolution depends on the transverse muon momentum, which results in a correlation between $\cos\theta_\mu$ and the B_s^0 mass shape. In order to take into account this correlation, the following mass model has been used. The DTF invariant mass is used as a discriminating variable. The DTF allows to determine also a per-event invariant mass uncertainty, δm, that can be used to take into account the correlation with $\cos\theta_\mu$. The combinatorial background is modelled with an exponential function

and the signal distribution with a double-side *Crystal Ball* (CB) function [117] (discussed in more details in Appendix A), where the σ parameter is written as a function of δm. The CB function is chosen to account for the two tails originating from radiative processes (the tail on the left) and from the constraint on the J/ψ mass required by the DTF (the tail on the right). The tail parameters of the signal distribution, α_1, α_2, n_1, n_2, are fixed in the fit to the values obtained from the fit to the simulated samples. In this case, to avoid any difference in the mass shape of data and MC, the (BKGCAT = 0 or 50 or 60) truth-matching requirement is used. An additional background source is considered in the fit, coming from the $B_d^0 \to J/\psi K^+K^-$ decay. It has the same final state as the signal decay but it is distributed around the B_d^0 nominal mass value [13]. This contribution is described with a Gaussian function with resolution, $\sigma(B_d^0)$, fixed to the resolution determined by fitting the $B_d^0 \to J/\psi K^+\pi^-$ control channel, procedure that will be explained in Sec. 4.5.1.
To account for a difference between the measured and the true per-event mass error, the σ parameter is parametrised to have a polynomial dependence on the DTF per-event mass error. The order of the polynomial is defined through MC studies. The polynomial has the form

$$\sigma = s_1 \times (\delta m) + s_2 \times (\delta m)^2, \tag{4.8}$$

where s_1 and s_2 are free parameters in the fit. This model is chosen as a baseline, but other polynomials could be considered to estimate the magnitude of a possible source of systematic uncertainty. The probability density function (PDF) used to model the B_s^0 invariant mass can be written as

$$PDF(m|\sigma) = f_{sig}\, CB(m;\mu,\alpha_1,\alpha_2,n_1,n_2,s_1,s_2|\sigma) + f_{bkg}\,((1-f_{B_d^0})e^{-\gamma m} + f_{B_d^0} G(m;\mu_{B_d^0},\sigma_{B_d^0})), \tag{4.9}$$

where f_{sig} and f_{bkg} are the fractions of signal and background candidates in the sample, γ is the coefficient of the exponential function that describes the combinatorial background, and $f_{B_d^0}$ is the fraction of $B_d^0 \to J/\psi K^+K^-$ background events with respect to the total background yield. As it will be shown in Sec. 4.7, the final decay-time and angular fit is performed simultaneously for the four data-taking years, for the two trigger categories (Unbiased and Biased), and in six bins of the invariant mass of the K^+K^- system, namely [990, 1008, 1016, 1020, 1024, 1032, 1050] MeV/c^2. Then, the *sPlot* technique is used independently for each subsample. The total number of $B_s^0 \to J/\psi K^+K^-$ candidates is 16096 ± 134, 102595 ± 340, 104908 ± 342 and 124329 ± 373 in the mass fit range, respectively for 2015, 2016, 2017 and 2018 data samples. The distributions together with the resulting fit curves in the six mass bins of the (K^+K^-) system and in the two different trigger categories for 2018 data are shown in Fig. 4.8 and Fig. 4.9.

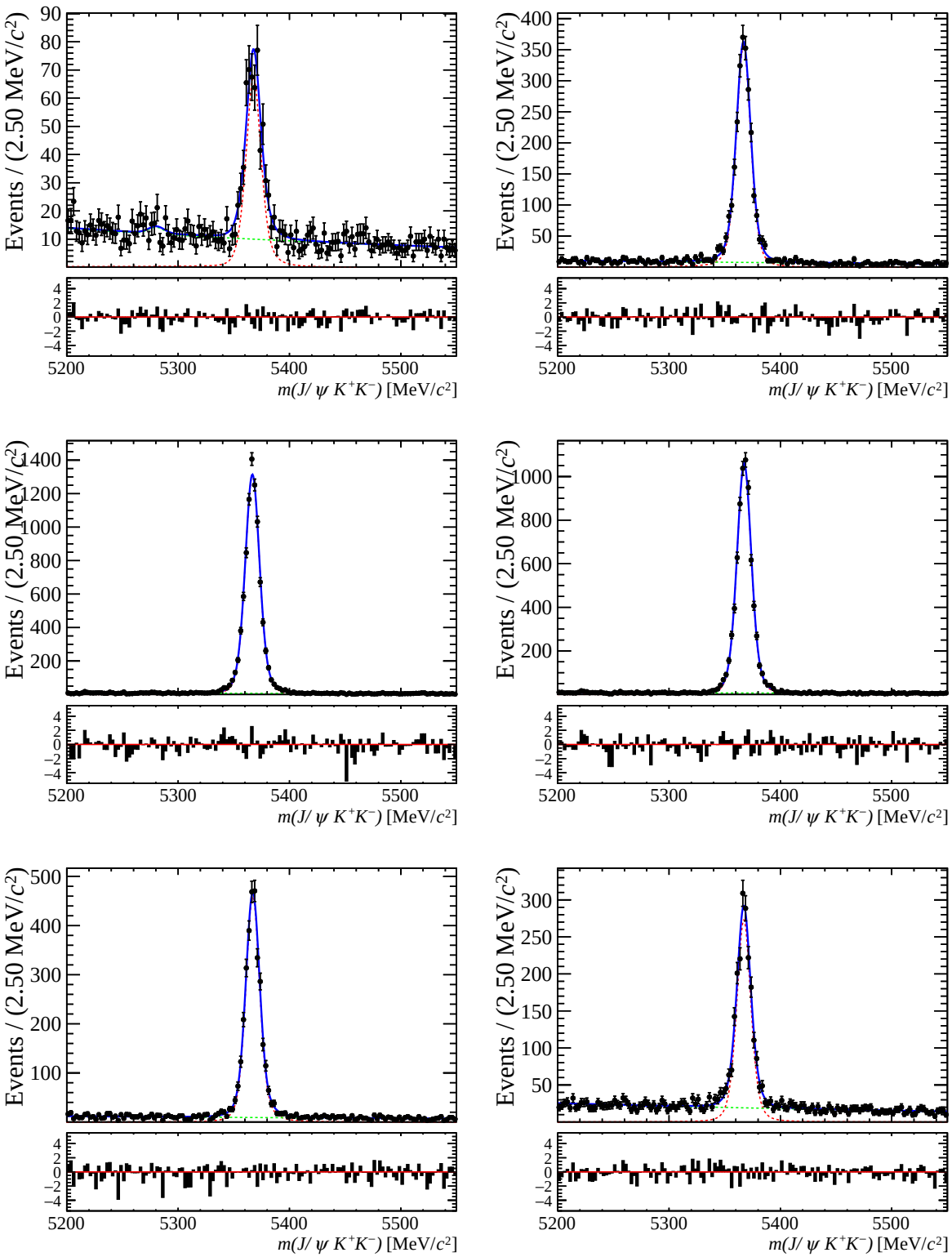

FIGURE 4.8. Fit to the $m(J/\psi K^+K^-)$ distribution in biased 2018 sample (solid blue line) using a CB shape for the signal (dashed blue line) and an exponential for the background (dashed green line). The associated fit residuals in bin uncertainty units (pulls) are shown in the bottom box of each plot. The $m(J/\psi K^+K^-)$ distribution is divided into six bins in $m(K^+K^-)$.

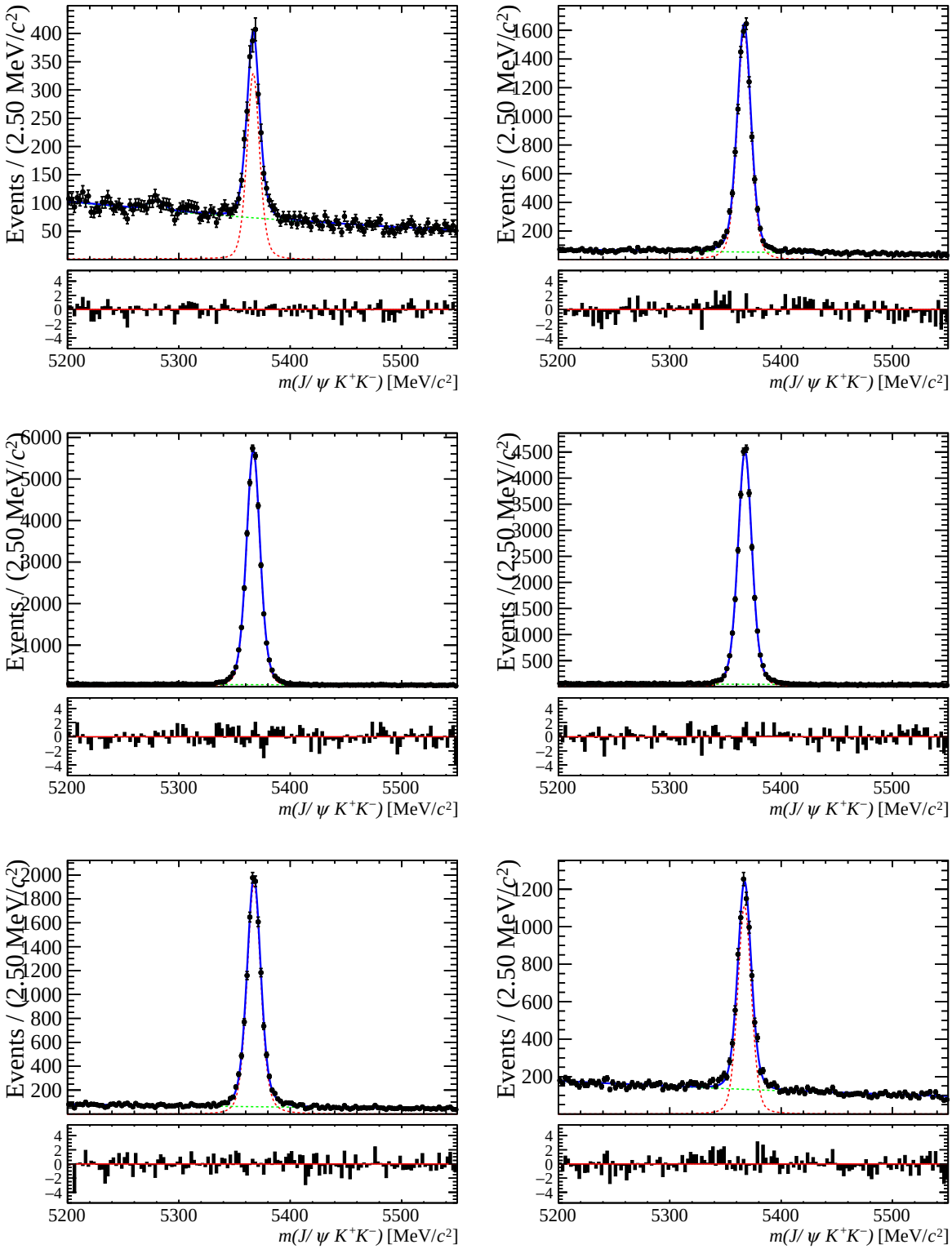

FIGURE 4.9. Fit to the $m(J/\psi K^+K^-)$ distribution in unbiased 2018 sample (solid blue line) using a CB shape for the signal (dashed blue line) and an exponential for the background (dashed green line). The associated fit residuals in bin uncertainty units (pulls) are shown in the bottom box of each plot. The $m(J/\psi K^+K^-)$ distribution is divided into six bins in $m(K^+K^-)$.

4.3 Selection of the $B_d^0 \to J/\psi K^+\pi^-$ control mode channel

The control mode $B_d^0 \to J/\psi K^*(892)^0$, with the $K^*(892)^0 \to K^+\pi^-$, has a decay topology very similar to the signal channel $B_s^0 \to J/\psi K^+K^-$, allowing to apply a nearly identical selection strategy. The K^{*0} meson, as the ϕ in the signal decay, is a vector particle, therefore the final state has the same polarization structure of the signal. For this reason, it is a good candidate for an angular efficiency cross-check, as it will be discussed in Section 4.5.4. In addition, it represents the decay of a B_d^0 meson into a flavour specific final state so the $\overline{B_d^0}$ meson can contribute to the same final state only via the mixing phenomena, that however in the B_d^0 system is quite suppressed, hence its final state is a direct probe of the flavour of the B_d^0 meson at production, avoiding the usage of the flavour tagging algorithms. Moreover, since $\Delta\Gamma_d$ is almost equal to zero, the resulting decay-time distribution is almost a pure single-exponential distribution, with a very well-known decay parameter: $\tau_{B_d^0} = 1.519 \pm 0.004$ps [2]. These features makes this channel the best decay to be used as a control mode in the determination of the decay-time acceptance, since the only distortion effects to the decay rate, taking into account the decay-time resolution, are due to the reconstruction and selection procedures that are in common with the signal decay channel.

4.3.1 Trigger, offline and multivariate selections

The same trigger selection as for the signal decay is used and each data sample is divided into two subsamples accordingly to the biased and unbiased trigger categories. Considering that most of the trigger selection is based on muons and that in both the signal and the control mode the muons come from the J/ψ, this selection causes a very similar decay-time acceptance in the two channels.
Then a dedicated stripping selection is applied, that presents small differences compared to that required for the signal channel. The B_d^0 stripping requirements are summarised in Table 4.4. In addition to the stripping selection, other offline requirements on the pion transverse momentum and goodness of particle identification algorithm are applied,

- $\pi^-_{PT} > 250\,\mathrm{MeV/c}$,
- $\pi^-_{PIDK} \neq -1000$.

While the differences on the K^{*0} and B_d^0 mass intervals are simply due to the different particles involved in the decay, the tighter requirements on the transverse momentum of the pion and the K^{*0} resonance are used to further suppress the larger amount of combinatorial background with respect to the signal channel, due to the presence of a pion in the final state. The cuts on the pion particle identification, as well as ensuring to consider only the events for which the PID algorithms work correctly ($\pi^-_{PIDK} \neq -1000$), allow to suppress most of the misidentified background sources, mainly due to the wrong identification of kaon tracks as pions. Then the

Table 4.4: Stripping selection criteria used to identify $B_d^0 \to J/\psi K^*$ candidates. The differences with respect to the signal stripping selection are shown in bold.

	Variable	Stripping
$J/\psi \to \mu^+\mu^-$	$\Delta\ln\mathcal{L}_{\mu\pi}$ $(\mu^\pm)$	> 0
	p_T $(\mu^\pm)$	> 500 MeV/c
	$\chi^2_{\rm DOCA}$	< 20
	$\chi^2_{\rm vtx}$/nDoF	< 16
	$m(\mu^+\mu^-)$	$\in [3016.9, 3176.9]$ MeV/c^2
$K^* \to K^+\pi^-$	$\boldsymbol{p_T}$ $\boldsymbol{(K^*)}$	$\boldsymbol{> 1300\,\mathrm{MeV/c}}$
	$\boldsymbol{m(K^+\pi^-)}$	$\boldsymbol{\in [826, 966]\,\mathrm{MeV/c^2}}$
	$\chi^2_{\rm vtx}$/nDoF	< 25
	$\chi^2_{\rm DOCA}$	< 30
	$\chi^2_{\rm track}$/nDoF (K^+)	< 5
	$\chi^2_{\rm track}$/nDoF (π^-)	< 5
	$\Delta\ln\mathcal{L}_{K\pi}$ (K^+)	> 0
	$\Delta\ln\mathcal{L}_{K\pi}$ (π^-)	< 0
$B_d^0 \to J/\psi K^*$	$\boldsymbol{m(J/\psi K^+\pi^-)}$	$\boldsymbol{\in [5150, 5450]\mathrm{MeV/c^2}}$
	$\chi^2_{\rm vtx}$/nDoF	< 20
	t	> 0.2 ps

same BDT classifier trained for the signal decay is applied to the $B_d^0 \to J/\psi{*}0$ sample, where the transverse momentum of the ϕ meson has been replaced by the one of the K^{*0} resonance, and the same optimised cut on the BDT output is used. This selection step allows to further reduce the combinatorial background contribution without adding additional decay-time acceptance effects with respect to the $B_s^0 \to J/\psi K^+K^-$ decay.

4.3.2 Correction of MC samples

In order to be able to rely on simulation for the determination of the decay-time acceptance, for studying the invariant mass shape after the BDT selection and to evaluate the efficiency of the vetoes applied to suppress misidentified backgrounds, the MC sample has to match the distributions in data. This correction is performed before the BDT selection.

As briefly explained for the signal decay, a known difference between data and MC samples is present for the distributions of some PID variables, especially for the muons. Figure 4.10 shows the PID variables comparison between 2018 truth-matched MC and sWeighted data sample before the correction. Major differences are visible and, therefore, the distributions in simulation of ProbNN variables have to be corrected. The correction is performed independently per year and per magnet polarity and the standard `PIDCalib` package is employed for this. In order to correct the muon and hadron PID variables, two different algorithms available in the `PIDCalib` package are used, called respectively `PIDGen` and `PIDCorr`. `PIDCorr` performs a transformation of the existing MC PID variables in such a way that they are distributed as in the calibration

data samples. In this way the correlations between PID variables for the same particle are maintained. The `PIDGen` approach performs a resampling of the PID variables, where the PID response is completely replaced by the one randomly generated from calibration PDFs, loosing all the information about the correlation between same particle PID variables. Both packages account for the dependence of the PID variables on the momentum and transverse momentum of the tracks, and the number of tracks in the event (N_{tracks}), that ensures to take always into account correlations between PID variables of different particles of the final state. The choice of using the `PIDGen` algorithm instead of the `PIDcorr` one for the muons is due to the absence of the proper calibrated samples for the last method.

The following step is to correct the MC for significant differences in the B_d^0 production kinematics and number of tracks. Additionally, some of the variables used in the BDT application that show different distributions are also considered in the correction procedure. The correction is obtained thanks to a weighting of the simulated sample in order to become closer to a statistically background-subtracted signal sample. The fit model used to apply the *sPlot* technique is composed of a double-side Ipatia [116] function for the description of the signal peak and an exponential function to describe the combinatorial background. The Ipatia function is used instead of a CB since it provides a better description of the signal peak for a higher number of events. The core of the peak is parametrised as a function of five shape parameters: ζ, β, λ, σ, and μ. The first two are set to zero, since ζ is determined to be very small and $\beta = 0$ implies that the core is symmetric with respect to the mean. The mean and width are left floating in the fit to data, while the tails parameters are fixed from studies performed on the MC samples. The fit in data is performed in the mass region $5230 < m(B_d^0) < 5330\,\text{MeV/c}^2$ and it is shown in Fig 4.11 for the 2018 data taking period.

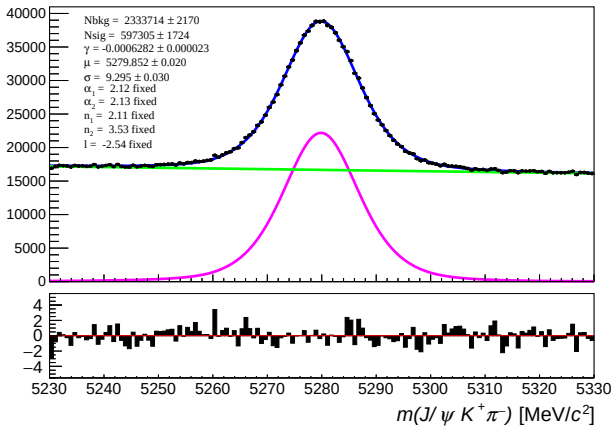

FIGURE 4.11. Distribution of $m(J/\psi K^+\pi^-)$ after the stripping and trigger selection in 2018. The blue solid line shows the total fit with the signal and combinatorial background described by the pink and green lines, respectively. The tail parameters of the fit are constrained from MC studies. The associated fit residuals in bin uncertainty units (pulls) are shown in the bottom box.

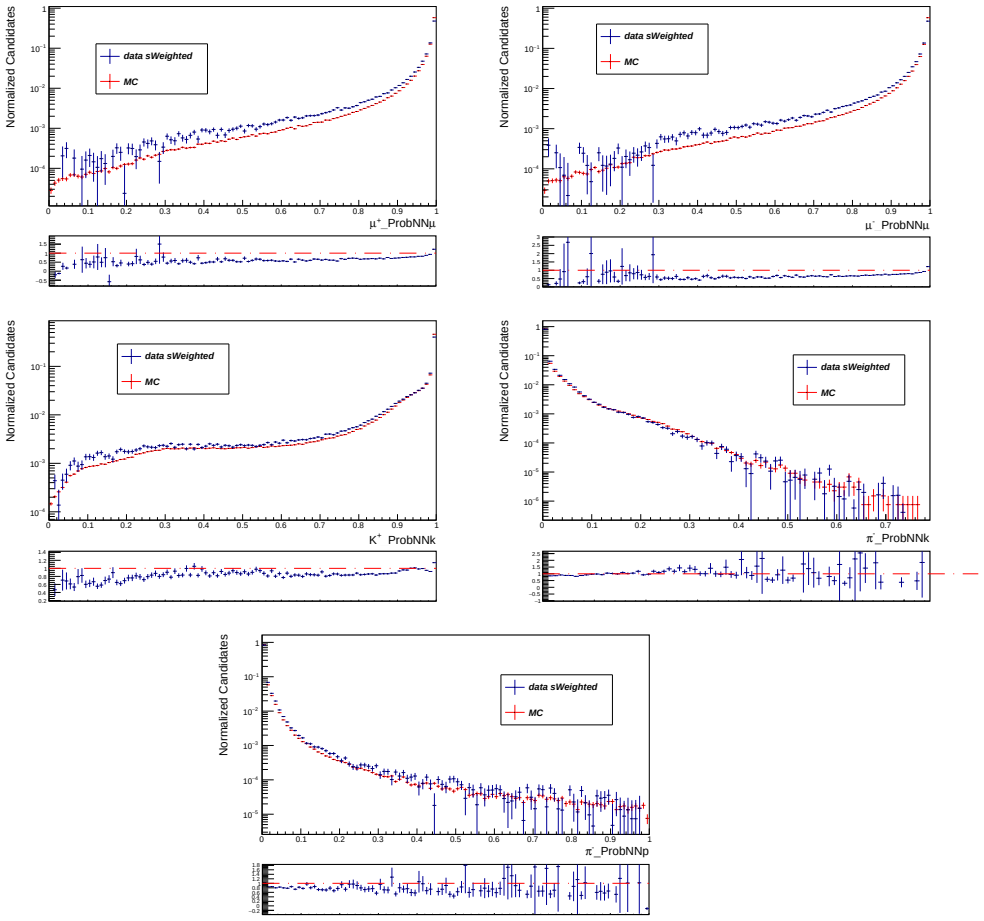

FIGURE 4.10. Distributions of PID variables for the sWeighted data (black dots) and simulated MC (red circles) in 2018 before the PID correction for $B_d^0 \to J/\psi K^{*0}$. The ratio between data and MC distributions is shown in the bottom box of each plot.

Typically, the weighting procedure is applied splitting the variable space into bins, and then for each bin a weight is defined to compensate the difference between the target (data) and the original (simulation) distributions. This approach works fine only if few variables are considered. Otherwise, it could lead to empty bins due to low statistics, obtaining divergent or null weights. To avoid this problem, the method of Gradient Boosting (GB) Reweighting [118] is used. It exploits BDT classifiers to find the optimal regions suitable for reweighting with a minimum number of bins. The following variables are used in the GB reweighting,

- χ^2_{track}/nDoF of K^+, π^- and muons;
- the B_d^0 transverse momentum $B^0_{d_{PT}}$;
- the B_d^0 pseudorapiduty $B^0_{d_\eta}$;
- the number of Long tracks `nLongTracks`.

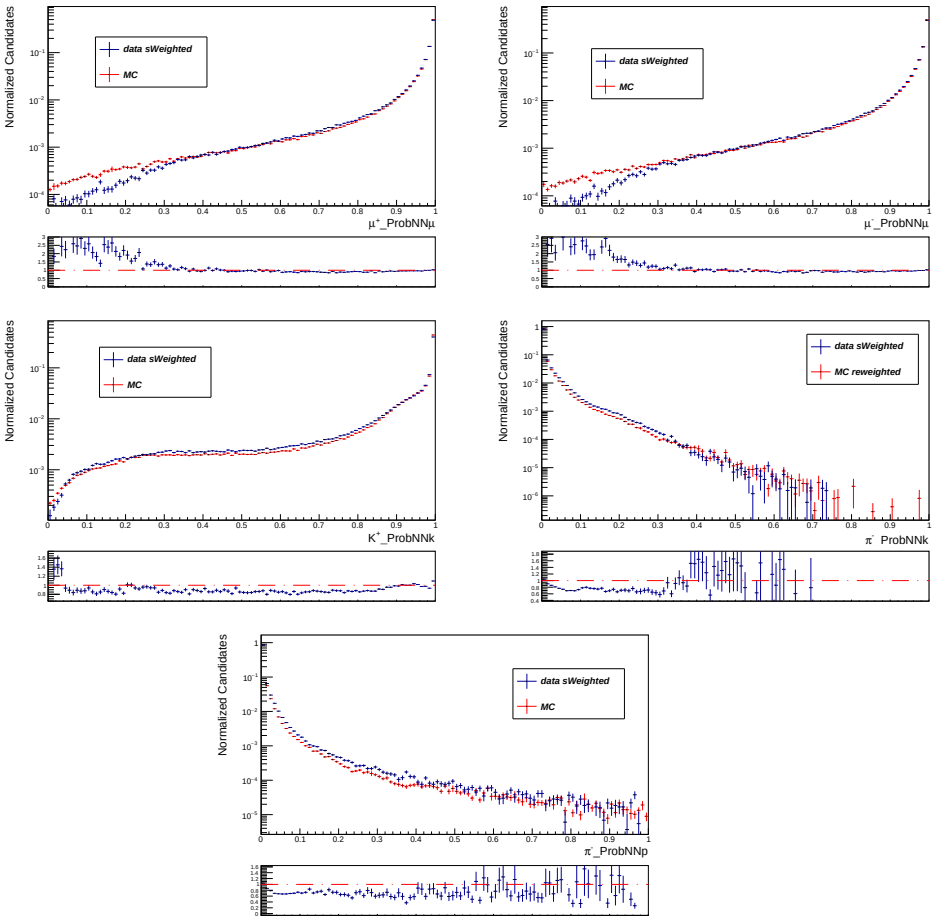

FIGURE 4.12. Distributions of PID variables for the sWeighted data (black dots) and GB weighted MC (red circles) in 2018 after the BDT selection for $B_d^0 \to J/\psi *0$. The ratio between data and MC distributions is shown in the bottom box of each plot.

The GB procedure is trained on truth-matched MC and then applied to the full MC sample. For each year a separate training is performed, except for 2015 MC for which the training result of 2016 is used, to avoid the limiting factor of low number of signal events in 2015 data. Figs. 4.13 and 4.12 show the distributions of the BDT variables and of the PID variables for the sWeighted data and GB weighted MC in 2018 after the BDT selection. All the variables show a good agreement.

4.3.3 Misindentified background

After the BDT selection, a small contribution due to misidentified background decays is present. Background from $\Lambda_b^0 \to J/\psi\, K^- p^+$ arises when the proton is misidentified as a pion. Figure 4.14 shows in blue the recalculated $J/\psi K^- p^+$ invariant mass distribution obtained from the B^0 data sample when interpreting the pion as a proton. The solid black line shows the Λ_b^0 mass value

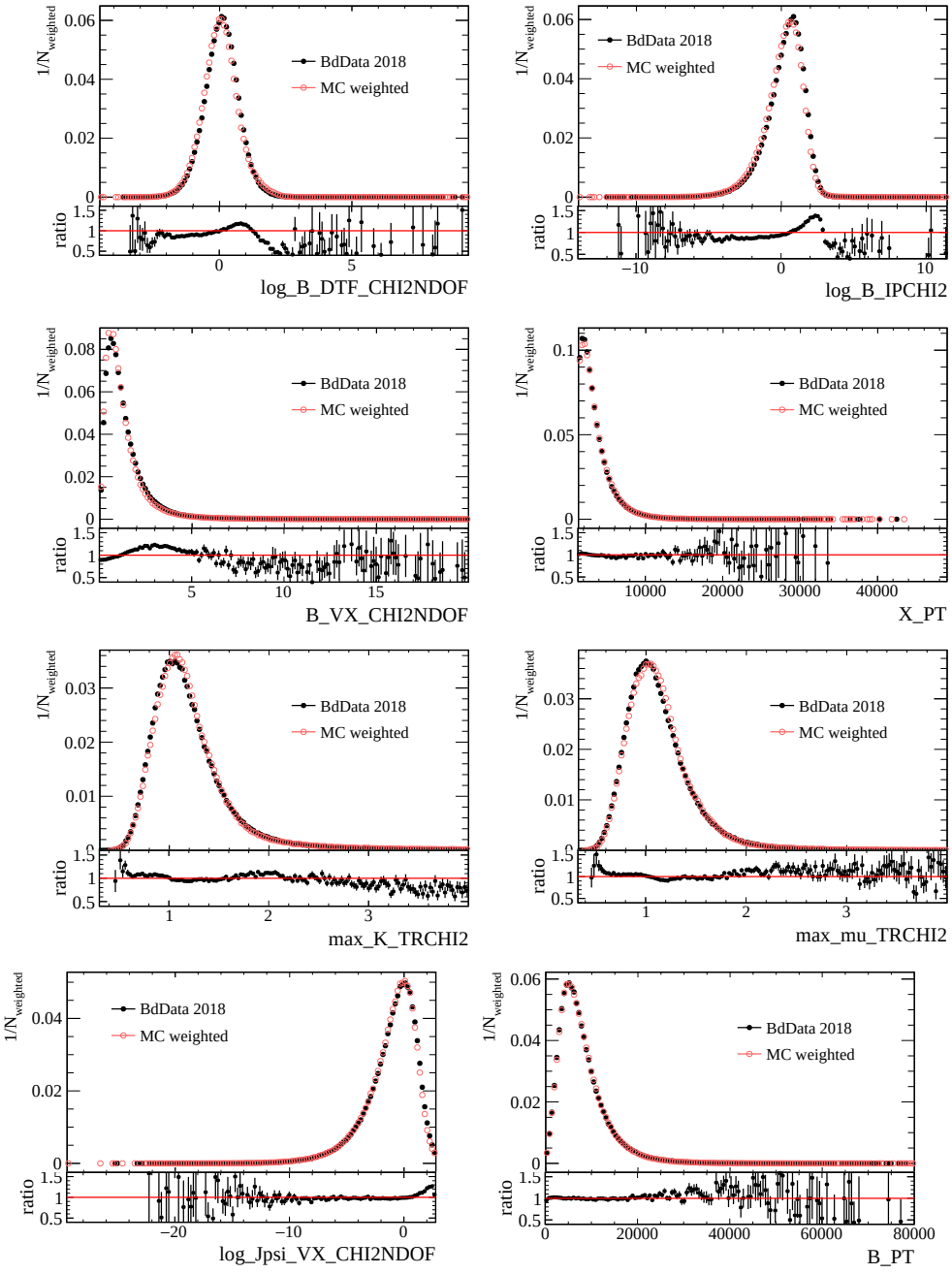

FIGURE 4.13. Distributions of some BDT variables for the sWeighted data (black dots) and GB weighted MC (red circles) in 2018 after the BDT selection for $B_d^0 \to J/\psi*0$. The ratio between data and MC distributions is shown in the bottom box of each plot.

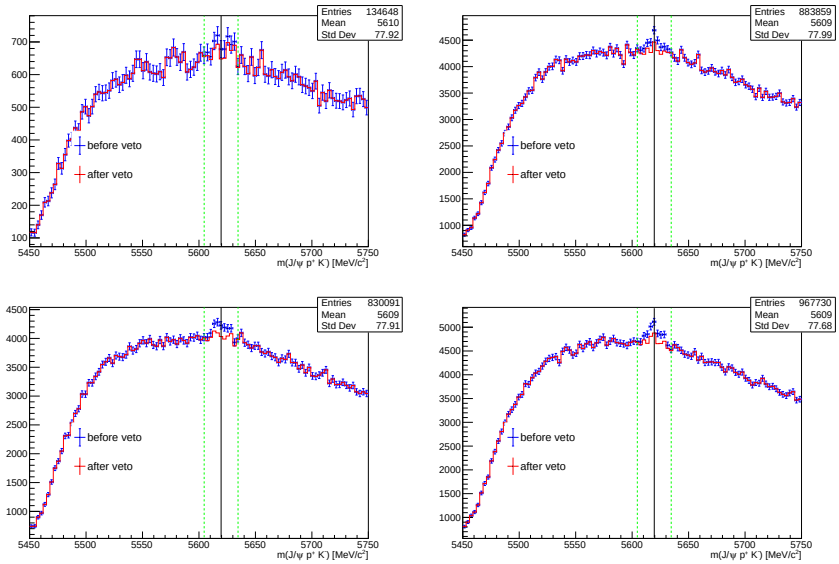

FIGURE 4.14. Invariant mass distributions for each year, namely 2015 (top left), 2016 (top right), 2017 (bottom left), 2018 (bottom right), of $J/\psi K^-p^+$ in data events after the BDT has been applied. The distribution m($J/\psi K^-p^+$) is obtained from the $B_d^0 \to J/\psi K^+\pi^-$ data sample when interpreting the pion as a proton. The black solid line highlights the Λ_b^0 mass value [13] at which the peaking background is expected, while the dotted green lines define a region of 15 MeV/c^2 around it.

[13] around which the peaking background is expected, while the green dotted ones delimit an interval of 15 MeV/c^2 around it. A small peaking structure is present in all data samples and it is removed requiring the following veto. All the candidates with a good identification probability of a pion as a proton, ProbNNp > 0.7, and $|(m(J/\psi K^-p^+)$ - $m(\Lambda_b^0)| < 15\,\text{MeV/c}^2$ are removed from the sample. The efficiency in the signal of this requirement is taken from simulation and is greater than 99.99%. The $m(J/\psi K^-p^+)$ distribution after the veto is shown in Fig. 4.14 by the red distribution.

Another possible peaking background source is due to $\Lambda_b^0 \to J/\psi\pi^-p^+$ decays, that could be misidentified as signal decays if the proton is wrongly identified as a kaon. However, the pion mode is Cabbibo suppressed compared to the kaon mode, therefore it is not expected to contribute, considering that the $\Lambda_b^0 \to J/\psi K^-p^+$ contribution has been already found to be very small. Misidentified candidates of $B_s^0 \to J/\psi K^+K^-$ could arise if one of the kaons is misidentified as a pion. Figure 4.15 shows in blue the recalculated $J/\psi K^+K^-$ invariant mass distribution obtained from the B_d^0 data sample when interpreting the pion as a kaon. The black solid line highlights the B_s^0 mass value [13] while the green dotted lines define a region of 30 MeV/c^2 around it. For each year no peaking structure is visible. Just to confirm the absence of misidentified decays, all candidates with the pion ProbNN$K > 0.7$ and $|(m(J/\psi K^-K^+) - m(B_s^0)| < 30\,\text{MeV/c}^2$ are removed from the sample. The efficiency of this requirement on the signal candidates is

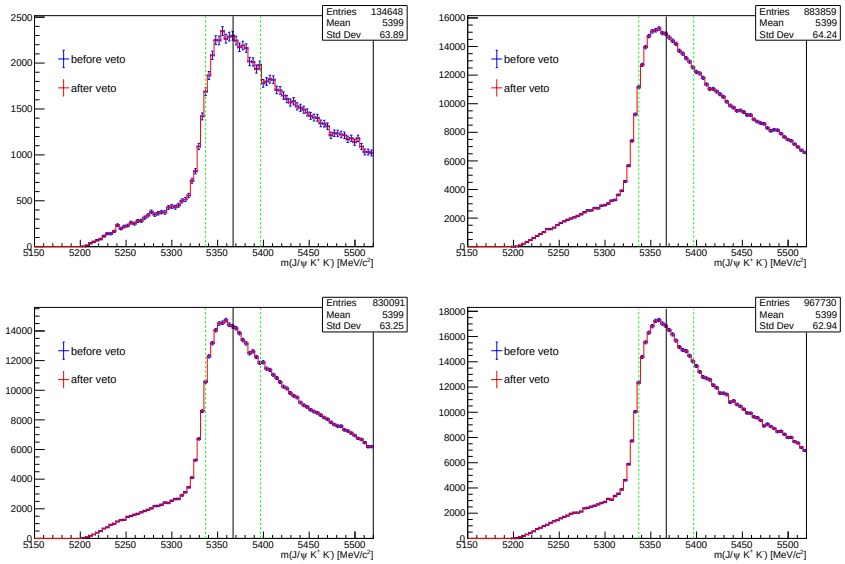

FIGURE 4.15. Invariant mass distributions for each year, namely 2015 (top left), 2016 (top right), 2017 (bottom left), 2018 (bottom right), of $J/\psi K^+K^-$ in data events after the BDT has been applied. The distribution m($J/\psi K^-K^-$) is obtained from the $B_d^0 \to J/\psi K^+\pi^-$ data sample when interpreting the pion as a kaon. The black solid line highlight the B_s^0 mass value [13] at around which the peaking background is expected, while the green dotted lines define a region of 30 MeV/c^2 around it.

evaluated in simulations and it is basically equal to one.
The background due to $B_s^0 \to J/\psi K^+\pi^-$ decays has the same final state as the B_d^0 decay and is shifted with respect to the nominal B_d^0 mass value by $\Delta m(B_s^0 - B_d^0) \sim 87$ MeV/c^2 [13]. Thus, in order to remove this contribution, it is enough to remove all the events above 5350 MeV/c^2. The last possible source of background comes from the decay of a B_s^0 meson into $J/\psi\pi^+\pi^-$. It could arise if one of the pions is misidentified as a kaon. However, since its branching ratio is only 26% with respect to the kaon mode [13], it is considered negligible.

4.3.4 Subtraction of the remaining background

In order to remove the small combinatorial background contribution that survives the BDT selection, for each year and trigger category the data sample is fitted to determine the final sWeights. The fit model is composed of a double-sided Ipatia function to describe the signal and an exponential function for the background. The tail parameters a_1, a_2, n_1 and n_2 are fixed to those obtained from a fit to the related truth-matched signal MC sample. The other parameters are left floating in the data fit. The invariant mass distribution and the fit projections are shown in Fig. 4.16. The fitted number of B_d^0 candidates is equal to 74299 ± 704, 476056 ± 1895, 471813 ± 1894 and 564007 ± 2037 in 2015, 2016, 2017 and 2018 data samples, respectively.

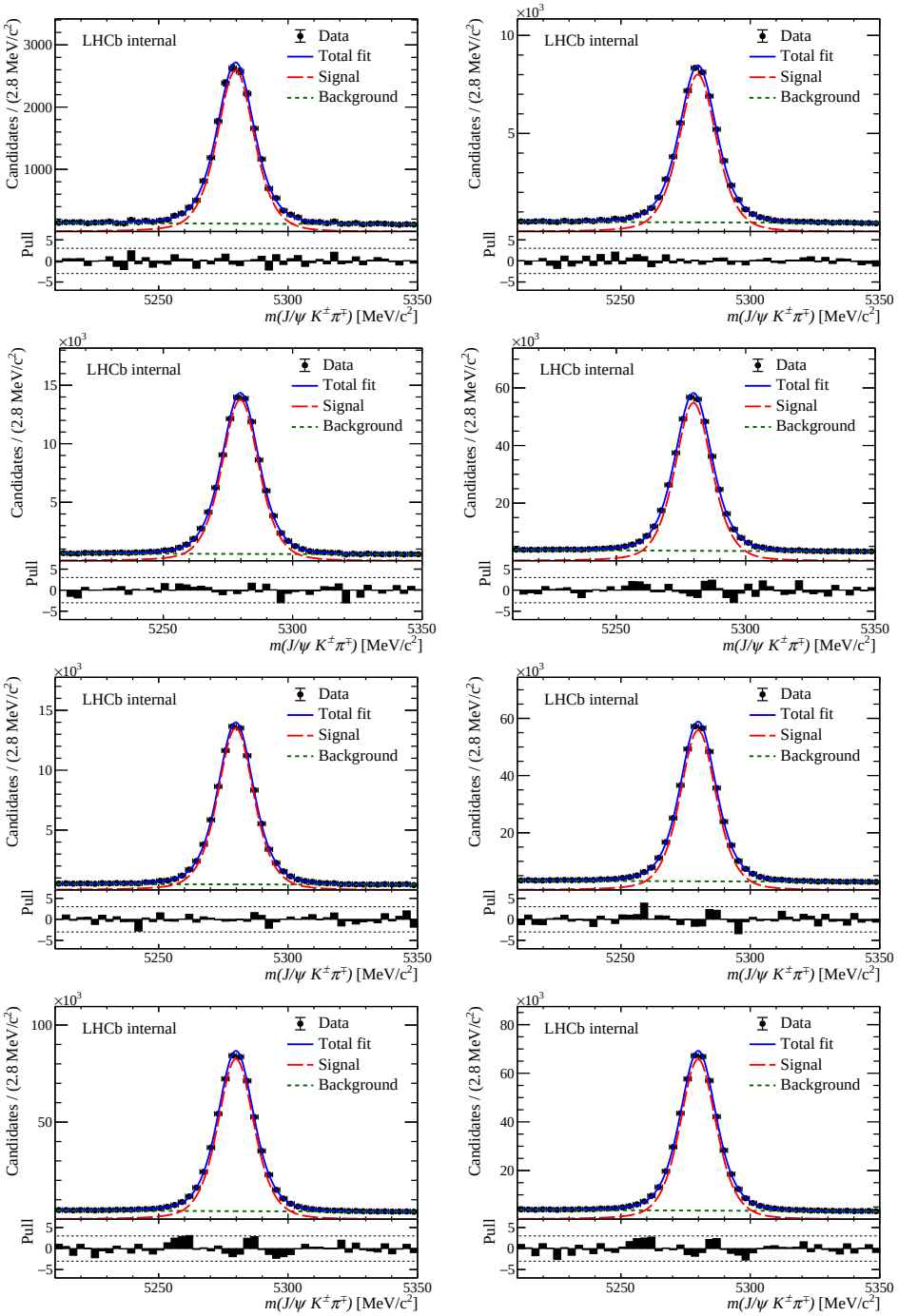

FIGURE 4.16. Fit to the $m(J/\psi K^+\pi^-)$ distribution in 2015-2018 data (black dots), from top to bottom. The total fit (solid blue line) sums up the signal shape corresponding to a Ipatia function (dashed red line) and a polynomial distribution for the background (dashed green line). Left plots are for the biased category and right plots are for the unbiased one. The associated fit residuals in bin uncertainty units (pulls) are shown in the bottom box of each plot.

4.4 Selection of the $B^+ \to J/\psi K^+$ control mode channel

The $B^+ \to J/\psi K^+$ channel is a "*self-tagged*" decay mode, since the flavour of the B^+ meson can be determined simply by looking at the charge of the kaon in the final state. For this reason it is used to calibrate the OS tagging algorithm.

4.4.1 Trigger and offline selections and correction of simulated samples

Considering the similarity with the signal decay mode, the same trigger selection is applied. Then a dedicated stripping line is used to further suppress the combinatorial background and similar b-decays, whose requirements are shown in Table 4.5.
In addition, the same requirements on the DTF decay time and decay-time uncertainty applied

Table 4.5: Stripping selection criteria used to identify $B^+ \to J/\psi K^+$ candidates.

	Variable	Stripping
$J/\psi \to \mu^+\mu^-$	$\Delta\ln\mathcal{L}_{\mu\pi}$ $(\mu^\pm)$	> 0
	p_T $(\mu^\pm)$	$> 500\,\mathrm{MeV/c}$
	χ^2_{DOCA}	< 20
	$\chi^2_{\mathrm{vtx}}/\mathrm{nDoF}$	< 16
	$m(\mu^+\mu^-)$	$\in [3016.9, 3176.9]\mathrm{MeV/c^2}$
K^+	p_T (ϕ)	$> 500\,\mathrm{MeV/c}$
	$\chi^2_{\mathrm{track}}/\mathrm{nDoF}$	< 5
	$\Delta\ln\mathcal{L}_{K\pi}$	> 0
$B^+ \to J/\psi K^+$	$m(J/\psi K^+)$	$\in [5150, 5450]\mathrm{MeV/c^2}$
	$\chi^2_{\mathrm{vtx}}/\mathrm{nDoF}$	< 10
	t	> 0.2 ps

for the signal decay are used, namely

- $0.3\,\mathrm{ps} < t < 15\,\mathrm{ps}$,
- $\sigma_t < 0.15\,\mathrm{ps}$.

Figure 4.17 shows the B^+ invariant mass distribution of 2018 data sample, after trigger and offline selections. A high contribution of combinatorial background events is visible and an additional selection procedure is required. To maximise the separation power between signal and background, a BDT classifier is used, in order to take into account also the correlations between the discriminating variables. Differently from what has been done for the $B^0_d \to J/\psi K^*(892)^0$ control mode, in this case a new BDT has to be trained, because of the quite different final state of the B^+ decay. As for the signal decay, MC simulated samples are taken as a proxy for the signal and the invariant mass sideband in data samples, $5400\,\mathrm{MeV/c^2} < m(B^+) < 5500\,\mathrm{MeV/c^2}$

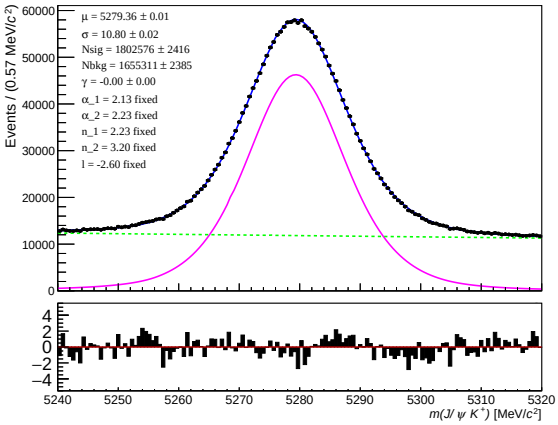

Figure 4.17: Distribution of $m(J/\psi K^+)$ after the trigger and offline selections in 2018. The blue solid line shows the total fit with the signal and combinatorial background described by the pink and green lines, respectively. The parameters of the fit are constrained from MC studies. The associated fit residuals in bin uncertainty units (pulls) are shown in the bottom box.

as a proxy for the background. Therefore, the MC samples have to be corrected to be as similar as possible to data. In order to have a description of the signal distributions directly from data to be used as a target for the correction procedure, the *sPlot* technique is exploited. The B^+ invariant mass distribution is fitted to assign an event-by-event signal $sWeight$. The model used for the fit is composed by a double-sided Ipatia function to describe the signal and an exponential for the background, and the fit projections for the 2018 sample are shown in Fig. 4.17. The signal tails are constrained to the values determined by fitting truth-matched simulated samples. The comparison between MC and data samples for PID and BDT variables before the correction are presented in Figs 4.18 and 4.19 To correct for the discrepancies in the PID variables existing between data and MC samples, the `PIDCalib` package is used, with the same strategy described for the $B_d^0 \to J/\psi K^*(892)^0$ decay mode. Then, accordingly to what has been explained in Sec. 4.3.2, the MC samples are reweighted. Differently from the corrections of the B_s^0 and B_d^0 decays, the number of tracks is used instead of the number of long tracks. This is chosen to optimise the agreement in the PID variables, considering their strong dependence on the number of tracks in the collision event. Figs. 4.20 and 4.21 show the PID and BDT distributions for 2018 sWeighted data, in blue, and GB weighted simulation, in red. After the correction the agreement between data and MC is very good. Then, MC samples can be used to train the BDT classifier.

4.4.2 Multivariate selection

To further remove combinatorial background events, a BDT classifier for each year is trained, except for 2015. Indeed, due to its limited statistics, the training result of 2016 is applied also to 2015 data directly. The variables used to train the BDT are:

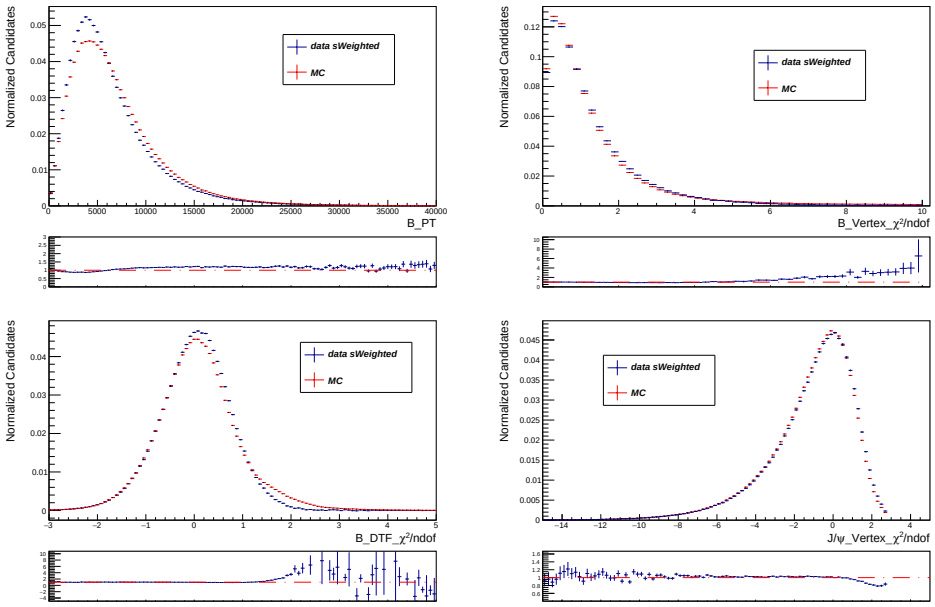

Figure 4.18: Distribution of some BDT variables for the sWeighted data (blue) and simulated MC (red) in 2018, before the GB weighting. The ratio between MC and data distributions is shown in the bottom box of each plot.

- the kaon track χ^2;
- the maximum of the muon track χ^2;
- the logarithm of the kaon ProbNNk, where the corrected variable is used for the signal description;
- the minimum of the logarithm of the ProbNNμ of the muons, where the corrected variable is used for the signal description;
- the logarithm of the J/ψ decay vertex χ^2/nDoF;
- the B^+ transverse momentum $p_{\mathrm{T}}^{\mathrm{B}}$;
- the B^+ decay vertex χ^2/nDoF;
- the logarithm of the B^+ impact parameter χ^2_{IP};
- the logarithm of the B^+ decay tree fit χ^2_{DTF}.

As for the other decay modes, in the DTF algorithm the J/ψ mass is constrained to the known value [13] and a constraint on the origin vertex position of the B^+ meson is applied, requiring that it comes from the PV. The signal and background distributions of the variables used for the BDT training and the output of the BDT classifier, for 2018, are shown in Figs 4.23 and

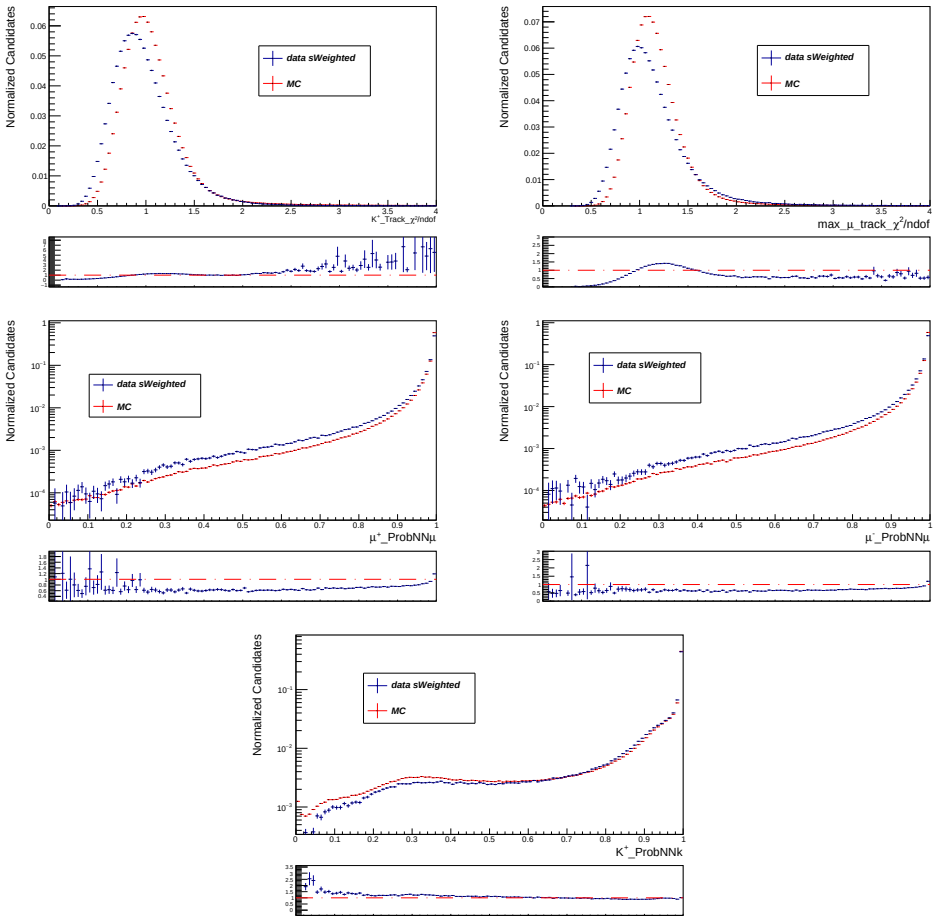

Figure 4.19: Distribution of some BDT and PID variables for the sWeighted data (blue) and simulated MC (red) in 2018, before the GB weighting and the PID correction. The ratio between MC and data distributions is shown in the bottom box of each plot.

4.22, respectively. A good separation is visible between background and signal. The figure of merit used to optimise the BDT response is the same of the signal decay and it is given by

$$\text{FOM} = \frac{(\sum_i w_i)^2}{\sum_i w_i^2}, \tag{4.10}$$

where the index i runs over all candidates in the sample and w_i are per-candidate weights that are determined with the *sPlot* technique from the invariant mass fit performed at each point in the scan over the BDT response. The fit is performed using an Ipatia function for the signal and an exponential function for the background. The tail parameters of the Ipatia are fixed to the values obtained from a fit to simulated signal events. The performance of the BDT classifier in 2018 is presented in Fig. 4.22. In the plot on the left the BDT outputs for the training and test samples are compared, showing the absence of overtraining, while the ROC curve presented in the right plot shows very good performances in terms of background rejection and signal

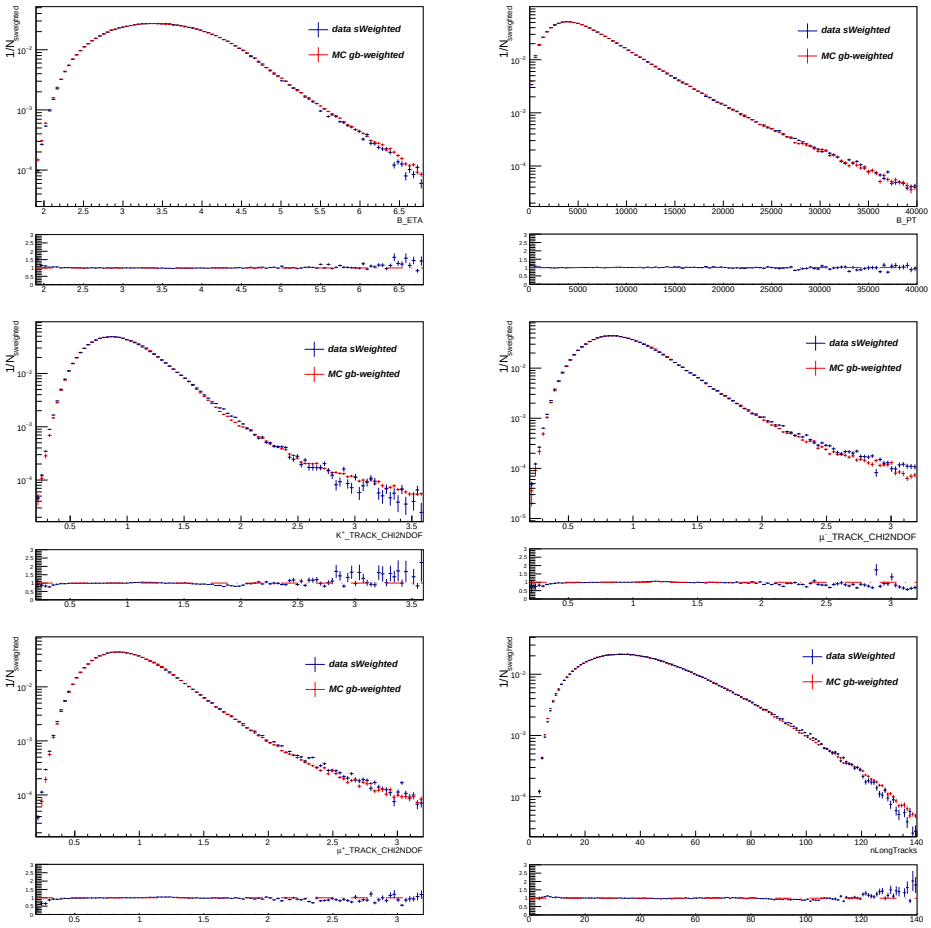

Figure 4.20: Distribution of GB weighting variables for the sWeighted data (blue) and GB weighted MC (red) in 2018. The ratio between MC and data distributions is shown in the bottom box of each plot.

efficiency. Figure 4.24 shows the FOM values as a function of the different BDT cuts, while the blue arrow indicates the optimisation point. All the events with a BDT output greater then -0.16, -0.22 and -0.24, respectively for 2015+2016, 2017 and 2018, are selected as B^+ signal decays.

4.4.3 Subtraction of the remaining background

The rejection of the remaining combinatorial background events is performed by sWeighting the $B^+ \to J/\psi K^+$ data samples to statistically remove the residual background contribution. The invariant mass fit is performed separately for each year and trigger category, according to the definitions given in Sec. 4.2.1. The fit model is composed of an Ipatia function for the signal and of an exponential function for the background description. The tail parameters α_1, α_2, n_1, n_2 are fixed from a fit to the related signal MC sample, with (BKGCAT = 0 or 50). The

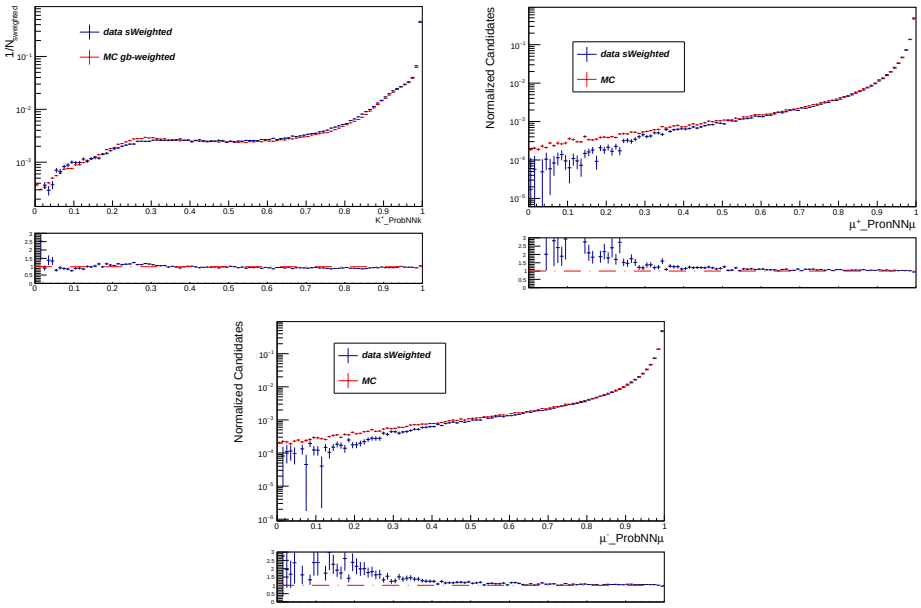

Figure 4.21: Distribution of PID variables for the sWeighted data (blue) and GB weighted and PID corrected MC (red) in 2018. The ratio between MC and data distributions is shown in the bottom box of each plot.

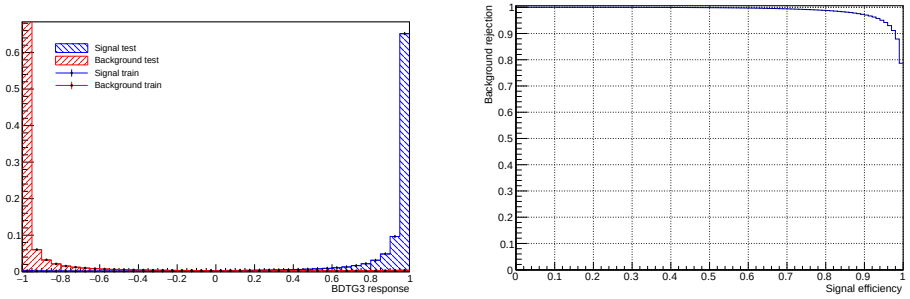

Figure 4.22: BDT output from the training and test (left) and ROC curve (right) for 2018.

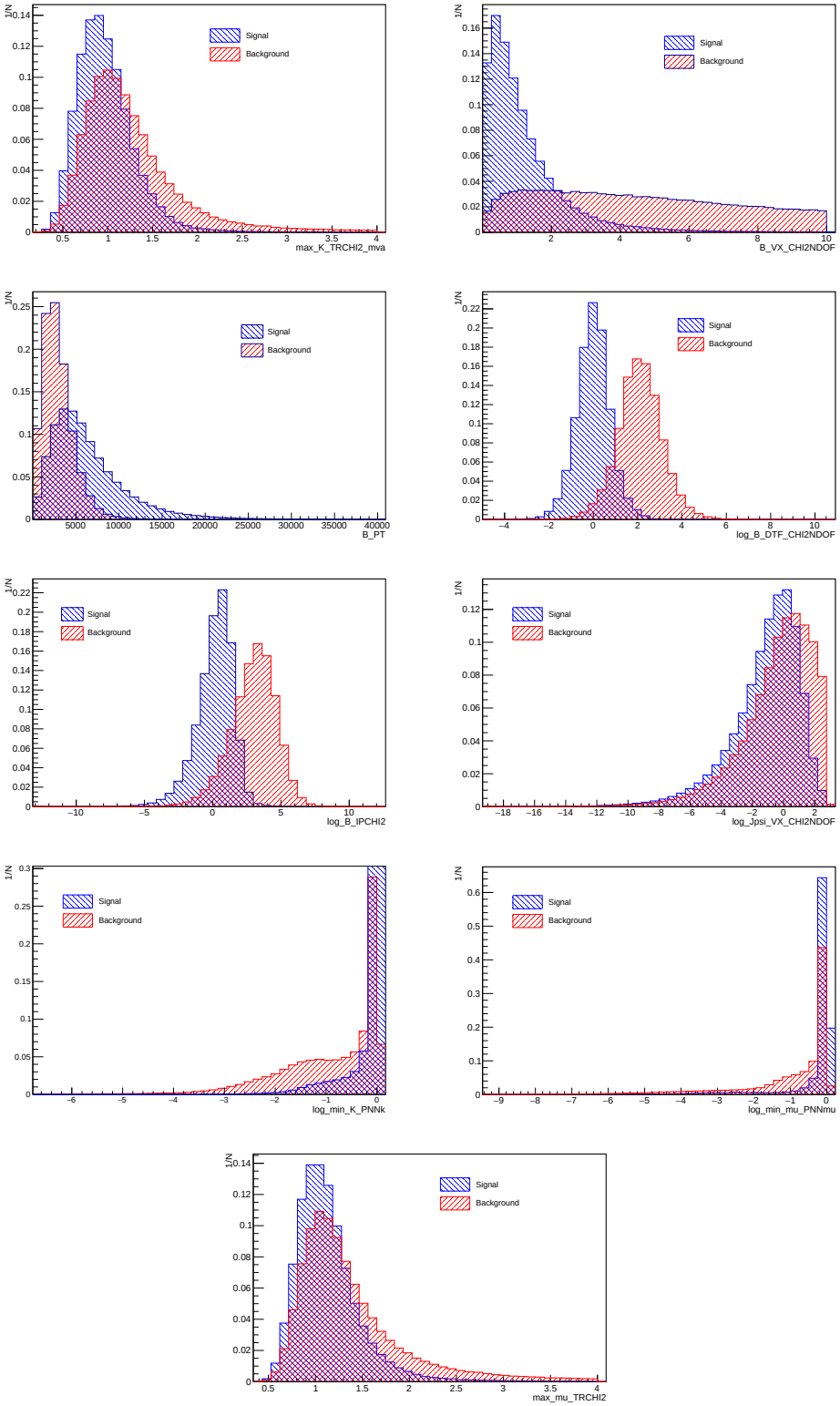

Figure 4.23: Distributions of variables in the signal (blue) and background (red) samples that are used for the BDT training of $B^+ \to J/\psi K^+$ for 2018.

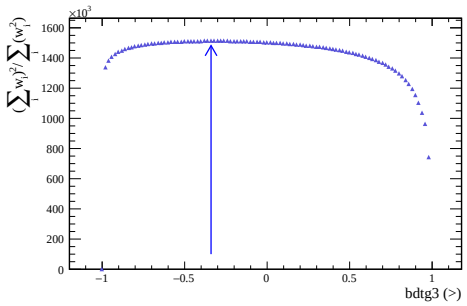

Figure 4.24: BDT cut output for 2018 evaluated to optimise the BDT response for the figure of merit. The blue arrow shows the optimisation point.

BKGCAT = 60 is not used in order to avoid the introduction of a bias due to the presence of real ghost events. The other parameters are left floating in the data fit. The fit to the selected $m(J/\psi K^+)$ invariant mass distributions is shown in Fig. 4.25, for 2015, 2016, 2017 and 2018 data samples, separately for the two trigger categories. The fits find 228590 ± 556 signal events in 2015 data, 1443272 ± 1397 in 2016, 1444088 ± 1365 in 2017 and 1732206 ± 1506 in 2018.

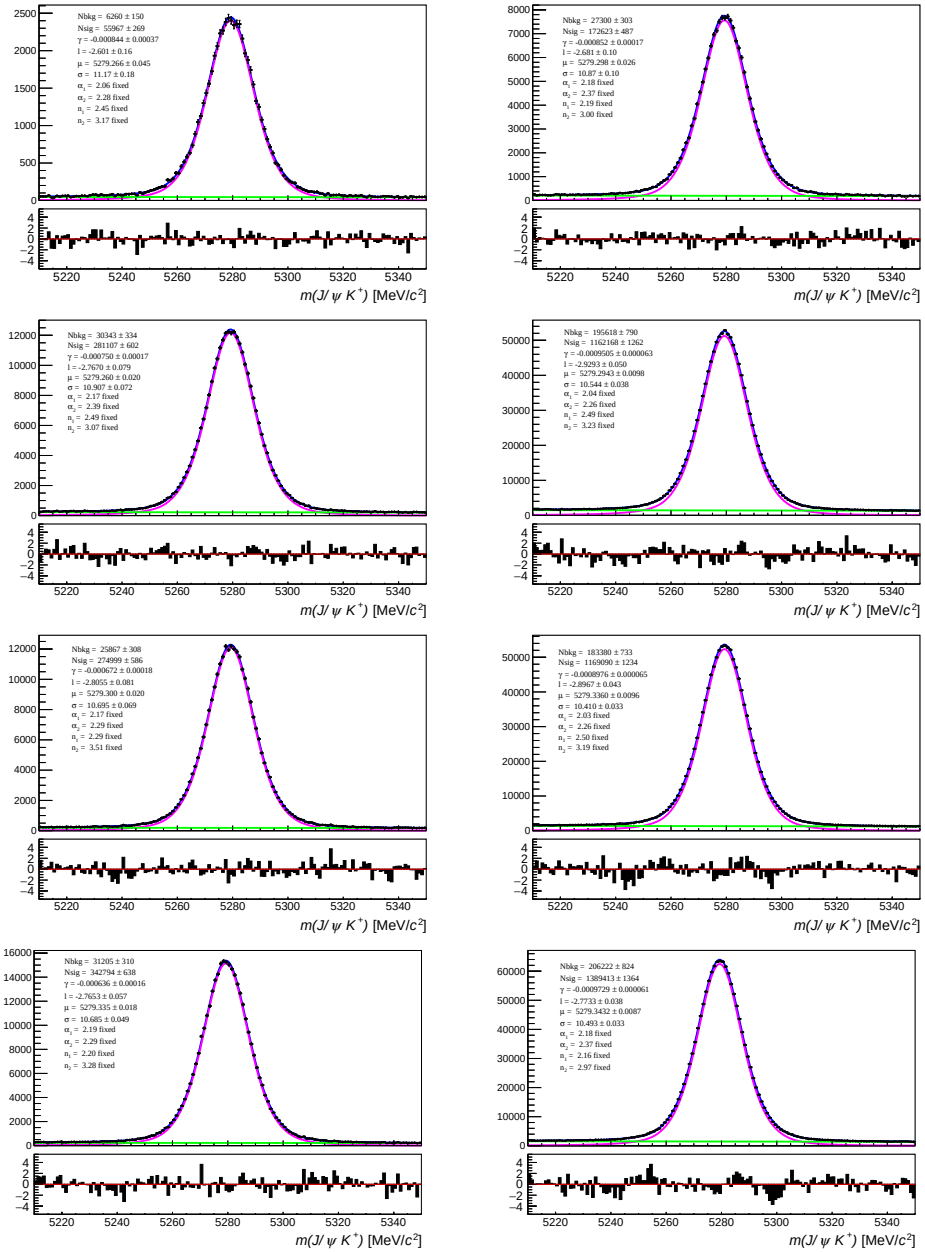

Figure 4.25: Fit to the selected $m(J/\psi K^+)$ invariant mass distribution in 2015, 2016, 2017 and 2018, respectively from top to bottom. Left plots show the fit on data samples selected with biased trigger category, right plots show the fit on data samples selected with unbiased trigger category. The blue solid line shows the total fit, with the signal and combinatorial background modelled by the pink and green lines, respectively. The associated fit residuals in bin uncertainty units (pulls) are shown in the bottom box of each plot.

4.5 Experimental effects

In this section the most important experimental effects are discussed. After the explanation of their sources, the strategies used to correct for them are presented.
Given that the correction for the decay-time bias, described in Chap. 3, is still under study, it will not be further discussed in this thesis.

4.5.1 Decay-time resolution

As already mentioned in Sec. 3.1.1, a first estimation of the per-event decay-time resolution is given by the DTF decay-time uncertainty. In order to obtain the real resolution value, this estimation has to be calibrated. Instead of relying on the simulated sample, a data driven method is used. A sample of fake $B_s^0 \to J/\psi K^+K^-$ candidates is employed, where all the tracks in the final state come from the PV. Such a sample is reconstructed by combining a prompt J/ψ particle with two random opposite kaons coming from the interaction point. These candidates have, by definition, a true decay time equal to zero. Then, any spread of the decay-time distribution is due to the decay-time resolution effect, that can be therefore measured. By performing this measurement in bins of the decay-time uncertainty, the calibration of the per-event estimated decay-time resolution can be carried out.

4.5.1.1 Selection of prompt $J/\psi K^+K^-$ candidates

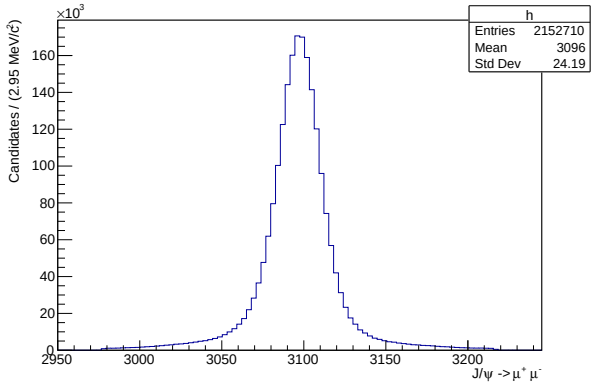

Figure 4.26: Distribution of prompt $J/\psi \to \mu^+\mu^-$ invariant mass in 2018, with all the selection steps applied.

The data sample of prompt $J/\psi K^+K^-$ events is selected using similar trigger and offline selections as it has been done for the signal $B_s^0 \to J/\psi K^+K^-$ decay. However, any requirement on the decay time or on the separation between the PV and the B_s^0 decay vertex is omitted, since they would reject mainly prompt events. Considering the loose requirements applied, an extremely high amount of events can pass the selection and have to be written to disk. In order

to limit this number, the HLT2 line is prescaled by a factor 0.2, i.e. only 20% of the events passing the trigger selection is stored. Then, the same multivariate selection and misidentified backgrounds rejection developed for the signal decay mode is applied. In order to evaluate the remaining background pollution of the prompt selected data sample, we can look at the invariant mass of the $J/\psi \to \mu^+\mu^-$. However, as it is shown in Fig. 4.26, after the selection procedure the data sample is populated mainly by real prompt J/ψ events, so no further selection steps are needed.

4.5.1.2 Decay-time resolution measurement

Figure 4.27 shows the decay-time distribution of the prompt $J/\psi K^+K^-$ sample of the 2018 dataset.

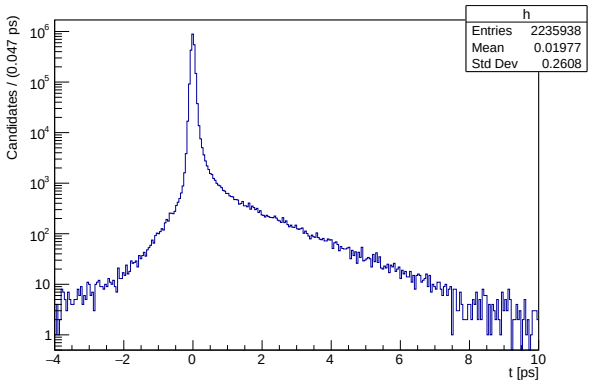

Figure 4.27: Decay-time distribution of prompt $J/\psi K^+K^-$ dataset in 2018, with all the selection steps applied.

The distribution is populated by three different components:

- the contribution from real prompt $J/\psi K^+K^-$ events, that constitutes the most important component and creates the peak around zero;
- the contribution from long-lived b-hadron decays, that populate the long tail at positive values of decay time, due to real $B_s^0 \to J/\psi K^+K^-$ decays or due to candidates combined from a prompt kaon pair and a J/ψ coming from the true B_s^0 decay;
- the contribution from the wrong PV determination, due to the fact that the finite momentum resolution results in the non-negligible probability of picking up a wrong primary vertex, with respect to which the decay time is determined. This small contribution is called the wrong primary vertex component and causes a broad symmetric distribution in the decay time.

The prompt component is not centred at zero. This is due to the presence of a decay-time bias due to the VELO misalignment. This effect will be taken into consideration during the extraction of the resolution effect.
The prompt decay-time distribution complexity leads to the development of a model-independent strategy to evaluate the decay-time resolution. Typically, in the previous ϕ_s measurements at LHCb, the resolution was determined by a maximum likelihood fit to the observed decay-time distribution of the prompt sample. The related PDF consisted of a component describing the real prompt candidates, $P(t)$, a component that accounted for candidates from the decay of long-living particles, $L(t)$, both convoluted with a resolution model $\mathcal{R}(t)$, and the wrong PV component, $W(t)$, in order to consider also the events in which the PV is wrongly associated. The PDF is then given by

$$PDF(t) = (1 - f_{wpv})\mathcal{R}(t) \otimes [f_{prompt}P(t) + (1 - f_{prompt})L(t)] + f_{wpv}W(t). \tag{4.11}$$

This approach requires a very good description of the decay-time distribution tails. In addition it does not take into account the contribution to the resolution due to the wrong PV component. Indeed, in this way this component is treated as a background contribution that is not included in the resolution model $\mathcal{R}(t)$. Thus, considering that its contribution to the total yield is very small, it is not convoluted with the resolution function in order to obtain a simpler and stable fit.
To avoid these problems, in this analysis a numerical and model-independent approach is used. In Chap. 3 it has been explained that the dilution D corresponds to the real part of the Fourier transform of the resolution function, evaluated at the frequency of Δm_s. The expression of the resolution function is unknown but the prompt sample allows to derive its distribution. In fact, the negative values of the prompt decay-time distribution are directly caused by the resolution effect and represent its distribution. The dilution D can be, therefore, numerically calculated by performing a Discrete Fourier transform (DFT) of the decay-time distribution of the prompt sample for $t < 0$. The usual equation of the DFT of a sequence of values can be adapted to compute the DFT of a distribution of events for a specific observable. In particular, the DFT of a distribution of N values, $t_0 ... t_{N-1}$, leads to a vector T, whose k-th component can be written as

$$T_k = \frac{1}{N} \sum_{n=0}^{N-1} e^{-i\nu_k t_n}, \tag{4.12}$$

where ν_k is the corresponding k-th frequency. The real part of the expression is given by

$$T_k = \frac{1}{N} \sum_{n=0}^{N-1} \cos(\nu_k t_n). \tag{4.13}$$

Therefore, the dilution can be computed numerically considering only the contribution due to the Δm_s frequency:

$$D = \frac{1}{N}\sum_{i=1}^{N}\cos(\Delta m_s(t - t_0)) \quad \text{if } t < 0, \tag{4.14}$$

where $\Delta m_s = 17.74\ \text{ps}^{-1}$. The value t_0 corresponds to the mean of the peak resulting from a fit of the decay-time distribution, described in more details in Sec. 4.5.1.4. This accounts for the presence of a decay-time bias. However, the correction performed is approximative, since the VELO is aligned at each injection of protons, and the decay-time bias depends on the specific event and it is therefore not equal for each candidate. New studies on the decay-time bias are ongoing and not presented in this thesis.
Eq. 4.14 represents a model independent way to determine the decay-time resolution automatically accounts also for the contribution of the wrong-PV events, that also populate the negative time distribution in the prompt sample. In order to calibrate the DTF decay-time uncertainty σ_t, for all the years, the dilution is computed in 10 bins of σ_t: (0.01, 0.021, 0.026, 0.032, 0.038, 0.044, 0.049, 0.054, 0.059, 0.064, 0.08). Due to the limited statistics in the 2015 samples, the first two and the last two bins are merged.
Because of the non-zero resolution, the long-lived candidates can also pollute the negative decay-time values. In order to take into account a small contribution of these non true prompt decays, the long-lived component shape is extracted from the fit described in Sec. 4.5.1.4. In order to correct the dilution for it, we can express the total dilution as the sum of the contribution due to non prompt and wrong PV events, D_{corr}, and the contribution of the long-lived component, D_{ll}, namely

$$D = \sum_i f_i D_i, \tag{4.15}$$

where f_i is the fraction corresponding to each contribution D_i to the total dilution. Then, the dilution due only to the real prompt events, D_{corr}, can be retrieved as

$$D_{corr} = \frac{1}{1 - I_{ll}} D - \frac{I_{ll}}{1 - I_{ll}} D_{ll}, \tag{4.16}$$

where D is the experimental dilution determined from the prompt sample, computed using Eq. (4.14), I_{ll} is the fraction of the long-lived component in the negative decay-time range and D_{ll} is its contribution to D. The latter is computed randomly generating event that follow the long-lived shape obtained by the fit of the decay-time distribution, see Fig. 4.28. The long-lived component counts only for about $2 - 5\%$ of the negative side of the decay-time distribution, as a function of the specific σ_t bin.

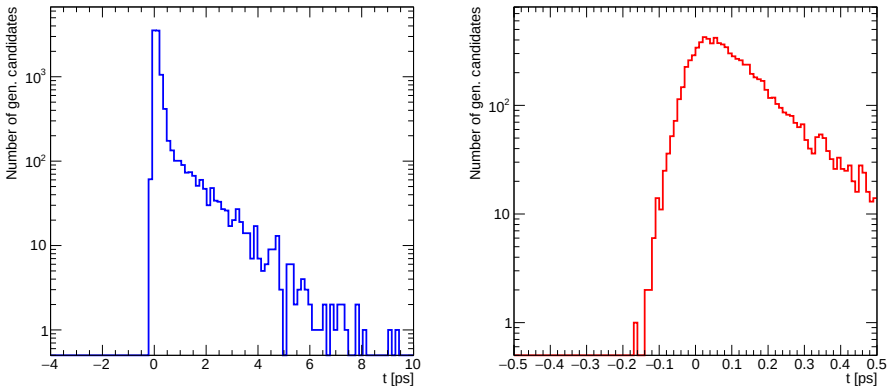

FIGURE 4.28. Long-lived component shape, sampled from the 2018 prompt $J/\psi K^+K^-$ real data sample in $\sigma_t \in (0.038, 0.044)$ps bin. The total generated statistics is 10000 events. In range $t \in (-4, 10)$ps (left) and $t \in (-0.5, 0.5)$ps (right).

In order to compare the dilution between the different data taking years, an effective dilution is defined

$$D_{eff} = \sqrt{\frac{1}{N}\sum_i n_i D_i}\,, \tag{4.17}$$

where N is the total number of events, and D_i and n_i are the dilution and the number of events in the i-th bin of σ_t, respectively. The effective dilution is 0.6938 ± 0.0008, 0.6919 ± 0.0003, 0.7188 ± 0.0003 and 0.7208 ± 0.0003 for 2015, 2016, 2017 and 2018 respectively.
From the corrected dilution, the effective width of the corresponding single-Gaussian resolution model, σ_{eff}, in each bin of σ_t can be extrapolated as

$$\sigma_{eff} = \sqrt{-2\ln(D_{corr})/\Delta m_s^2}. \tag{4.18}$$

The average effective resolutions are equal to 48.20 ± 0.12 fs for 2015, 48.38 ± 0.04 fs for 2016, 45.81 ± 0.04 fs for 2017 and 45.61 ± 0.04 fs for 2018. It is possible to note that the effective resolution in data has improved in 2017 and 2018, compared to 2015 and 2016.

4.5.1.3 Calibration of the estimated decay-time resolution

The effective width is used to calibrate the per-event decay-time error σ_t as

$$\sigma_{eff} = p_0 + p_1(\sigma_t - \sigma_0), \tag{4.19}$$

where σ_0 is an average offset that reduces correlation between p_0 and p_1. Figure 4.29 shows the calibration curve for the inclusive prompt $J/\psi K^+K^-$ sample for 2018. In order to evaluate a possible systematic uncertainty due to the calibration model, studies are ongoing, using for

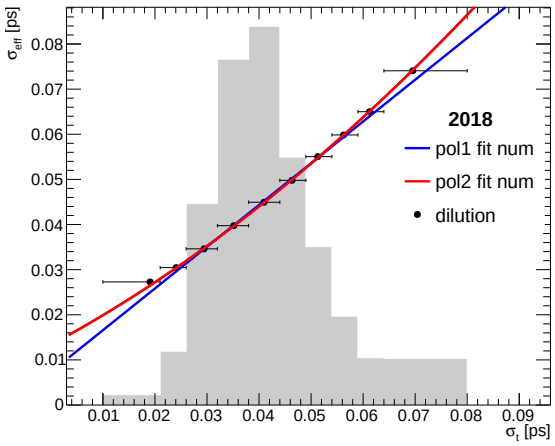

Figure 4.29: Calibration curve of prompt $J/\psi K^+K^-$ sample for 2018. The grey area shows the normalised distribution of σ_t.

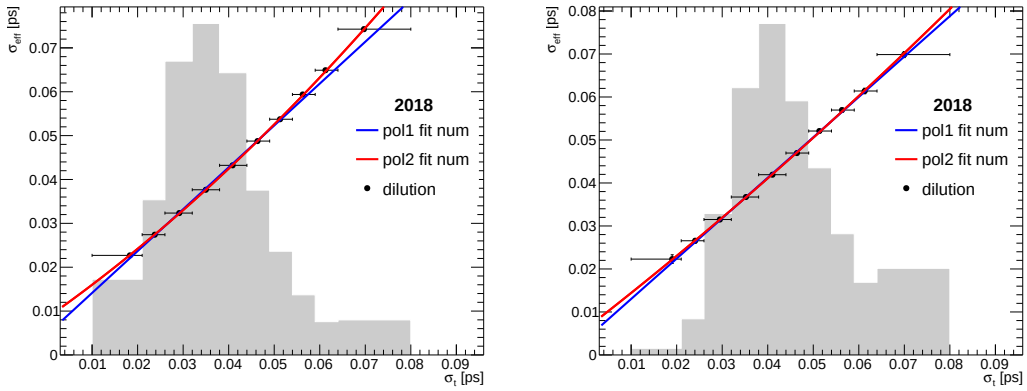

Figure 4.30: Calibration curve of MC prompt $J/\psi K^+K^-$ sample (at right) and MC signal $B_s^0 \to J/\psi K^+K^-$ sample (at left) for 2018. The grey area shows normalised distribution of σ_t.

example a second order polynomial function. Any difference in the final result with respect the nominal one obtained with the linear model will be set as systematic uncertainty.

Because of the different phase space between signal and prompt samples, see Fig. 4.32, the calibration can not be directly applied to the signal $B_s^0 \to J/\psi K^+K^-$ channel. The differences between prompt and signal decay modes are corrected by using MC simulation samples of prompt and signal decays. For both samples, a calibration curve for each year is determined, using the same procedure performed in the prompt data sample. The curves obtained for 2018 are shown in Fig. 4.30. For the signal MC sample, the difference between the reconstructed and the generated decay time, $t_{reco} - t_{true}$, is used to compute numerically the dilution in bins of σ_t. There is no long-lived component in this case, so the dilution does not have to be corrected. The

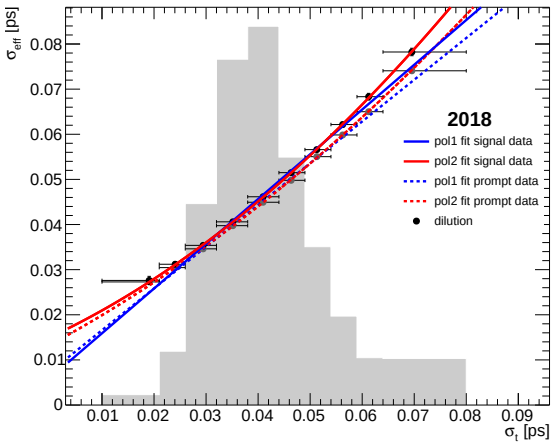

Figure 4.31: Calibration curve of data signal $B_s^0 \to J/\psi K^+K^-$ sample for 2018, in comparison with data prompt curve. The grey area shows normalised distribution of σ_t.

ratio between the dilutions obtained from these two MC samples, corrected for the long-lived component in the case of the prompt sample, is used to translate the dilution D_{corr} of prompt data into a signal data dilution:

$$D_{signal,\,data} = \frac{D_{signal,\,MC}}{D_{prompt,\,MC}} D_{prompt,\,data}. \tag{4.20}$$

The corresponding uncertainty is also propagated to the signal sample. This procedure has been performed in each bin of σ_t. The resulting effective dilutions are 0.7539 ± 0.0008, 0.7571 ± 0.0003, 0.7640 ± 0.0003 and 0.7695 ± 0.0003 for 2015, 2016, 2017 and 2018, respectively.
Each signal data dilution is translated into the corresponding width of the single effective Gaussian resolution and the signal data calibration curve is obtained. The result for 2018 is shown in Fig. 4.31, in comparison with that obtained from the prompt data sample. The average effective resolutions for signal $B_s^0 \to J/\psi K^+K^-$ decays are equal to 42.37 ± 0.12 fs for 2015, 42.05 ± 0.05 fs for 2016, 41.36 ± 0.05 fs for 2017 and 40.80 ± 0.06 fs for 2018.

4.5.1.4 Decay-time resolution fit

In order to extract the long-lived component shape, the $\delta t = t_{reco} - t_{true}$ distribution is fitted with the PDF described in Eq. 4.21. It is modelled as the sum of two exponential functions with different lifetimes τ_1 and τ_2, that are free to vary in the fit

$$\mathcal{L}(t) = f_{sl} e^{-t/\tau_1} + (1 - f_{sl}) e^{-t/\tau_2}, \tag{4.21}$$

where f_{sl} is the fraction fraction of the exponential function.
The resolution $\mathcal{R}(t)$ is modelled as the sum of two Gaussians with shared mean and different

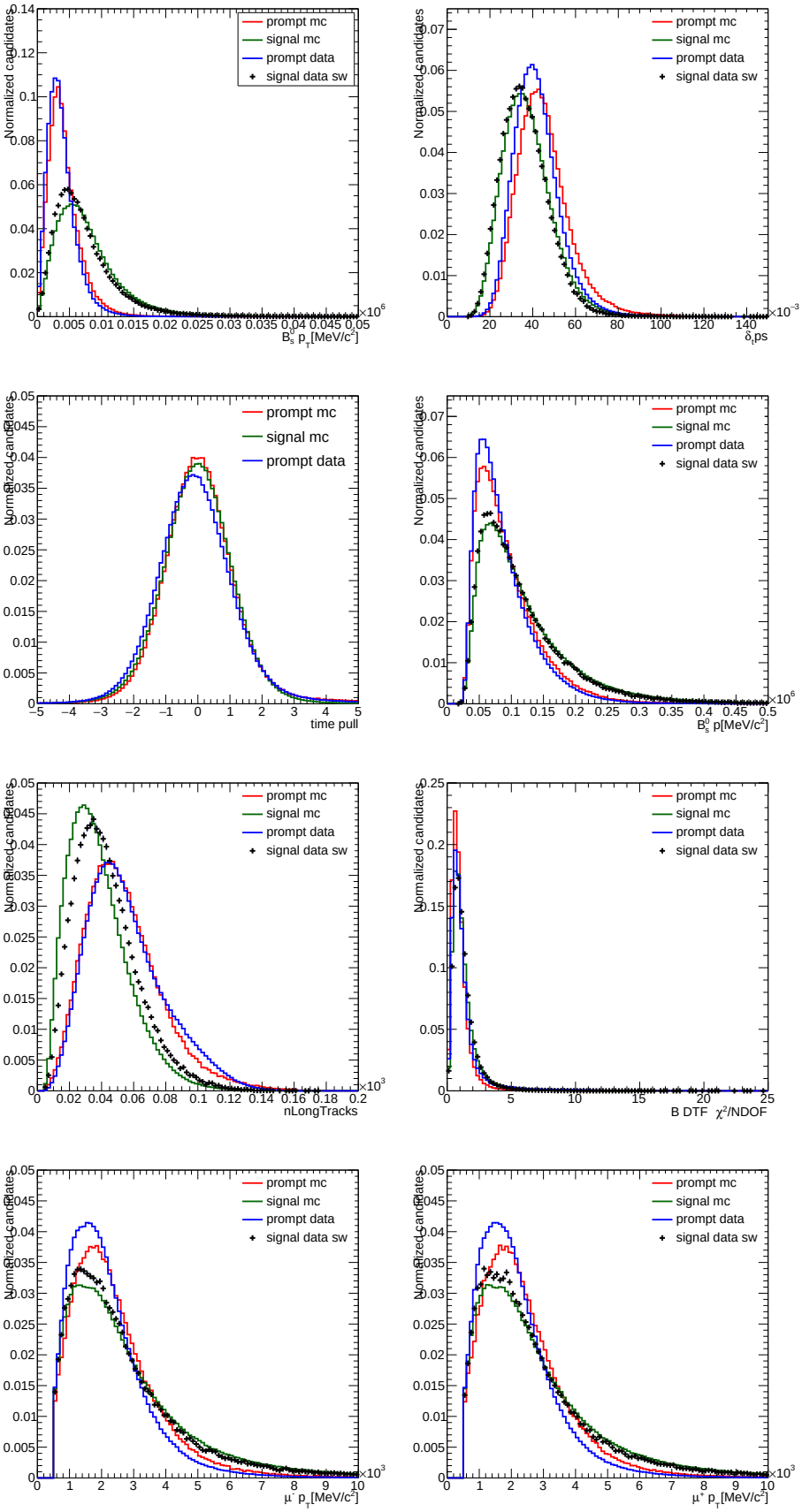

Figure 4.32: Distributions of B_s^0 p_T, σ_t, time pull $t_{reco} - t_{true}$, B_s^0 p, nLongTracks, DTF χ^2/nDoF and muons p_T for the prompt real (blue) and simulated $J/\psi K^+K^-$ samples (red) and B_s^0 signal sWeighted real sample (black markers) and signal simulated sample (green) for 2018 year.

widths. In order to reduce the correlation between widths and fractions of Gaussians, the widths are reparametrised as

$$\sigma_1 = s_1 - \sqrt{\frac{f}{1-f}}\, s_2, \tag{4.22}$$

$$\sigma_2 = s_1 + \sqrt{\frac{1-f}{f}}\, s_2, \tag{4.23}$$

where s_1, s_2 are the coefficients of the parametrisation and f is the fraction between the two Gaussians.

The shape of the wrong primary vertex component is fixed in the fit and is obtained from studies performed using dedicated samples, produced using the `PVMixer` tool. For each reconstructed event, this tool allows to add vertices from another event. After adding an extra primary vertex, a search for the best primary vertex is performed again. If the added vertex is picked as a new best primary vertex, the event is flagged. This contribution is modelled with the sum of two double-sided exponential functions, with

$$W(t) = f e^{-|t|/\tau_1^{wpv}} + (1-f) e^{-|t|/\tau_2^{wpv}}, \tag{4.24}$$

where f is the fraction of the first function. The resulting wrong-PV shape for the 2018 prompt data sample, together with the fit result, is shown in Fig. 4.33.

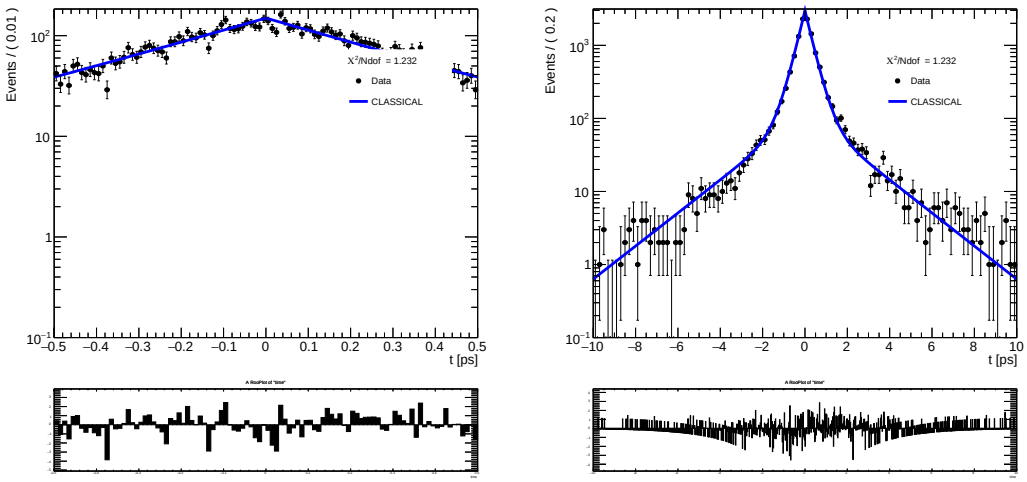

Figure 4.33: The wrong primary vertex shapes (indicated by the solid line and the name classical) obtained from the `PVMixer` tool for the 2018 prompt $J/\psi K^+K^-$ data sample.

The model described above is used to fit both real and simulated prompt samples. For the signal simulated sample, the long-lived component is left out.

4.5.2 Flavour Tagging calibration

In order to study the time-dependent decay rate of B_s^0 mesons, it is important to determine their initial flavour at production. In Chap. 3, two different tagging algorithms are presented. Namely, the same side (SS) and the opposite side (OS) taggers. Both the algorithms provide the tag $q = (+1, -1, 0)$, for B_s^0, $\overline{B}_s^0$ and untagged candidates, respectively, and an estimated mistag probability $\eta \in (0, 0.5)$, that quantifies the probability to give the wrong tag. In order to correctly take into account the possibility of wrongly tagged candidates, the previously mentioned estimated mistag probability needs to be calibrated. For this purpose, two different flavour specific decays are used, one for each tagging algorithm: the $B^+ \to J/\psi K^+$ for the OS and the $B_s^0 \to D_s^- \pi^+$ for the SS. Given the estimated mistag probability η for each of the tagging algorithms, the respective calibrated version of it, $\omega(\eta)$, is parametrised as

$$\omega(\eta) = \Big(p_0 + \frac{\Delta p_0}{2}\Big) + \Big(p_1 + \frac{\Delta p_1}{2}\Big)\Big(\eta - \langle\eta\rangle\Big), \tag{4.25a}$$

$$\overline{\omega}(\eta) = \Big(p_0 - \frac{\Delta p_0}{2}\Big) + \Big(p_1 - \frac{\Delta p_1}{2}\Big)\Big(\eta - \langle\eta\rangle\Big), \tag{4.25b}$$

where $\langle\eta\rangle$ is the average estimated mistag probability of the sample used in the calibration and p_0 and p_1 are the main calibration parameters. Δp_0 and Δp_1 allow for different calibrated mistag probabilities ω and $\overline{\omega}$ for initial Bs^0 and $\overline{B}_s^0$ mesons, respectively.

4.5.2.1 Calibration of the opposite-side algorithm

The OS tagger is calibrated using the selected and background-subtracted $B^\pm \to J/\psi K^\pm$ decay, for which the charge of the kaon determines the initial flavour of the b-meson. In order to use the calibration here discussed for the B_s^0 signal decay, the $B^\pm$ sample is weighted to match the distributions of P_T, rapidity, nTracks and number of PVs of the selected signal sample, using the already mentioned gradient boosting (GB) algorithm. The $B^\pm$ sample is then divided in bins of η. For each bin, the true mistag probability is directly accessible simply comparing the charge of the kaon to the tag obtained from the tagging algorithm. Then the true mistag probability is compared to the estimated one, as shown in Fig. 4.34 for the different years. An unbinned maximum likelihood fit is then performed on the true mistag probability rate as a function of η. Depending on the $B^\pm$ flavour, the PDF defined is

$$PDF(a|\eta) = (1-a)\omega(\eta) + a(1-\omega(\eta)), \tag{4.26a}$$

$$PDF(a|\eta) = (1-a)\overline{\omega}(\eta) + a(1-\overline{\omega}(\eta)), \tag{4.26b}$$

where the first equation if for the B^+ and the second for the B^-. The parameter a is equal to 1 if the candidate is tagged correctly, otherwise is 0. In order to evaluate a possible systematic effect due to the fit procedure, a χ^2 binned fit is also performed. In addition, the evaluation of a possible systematic effect on the background-subtraction procedure is also consider. A

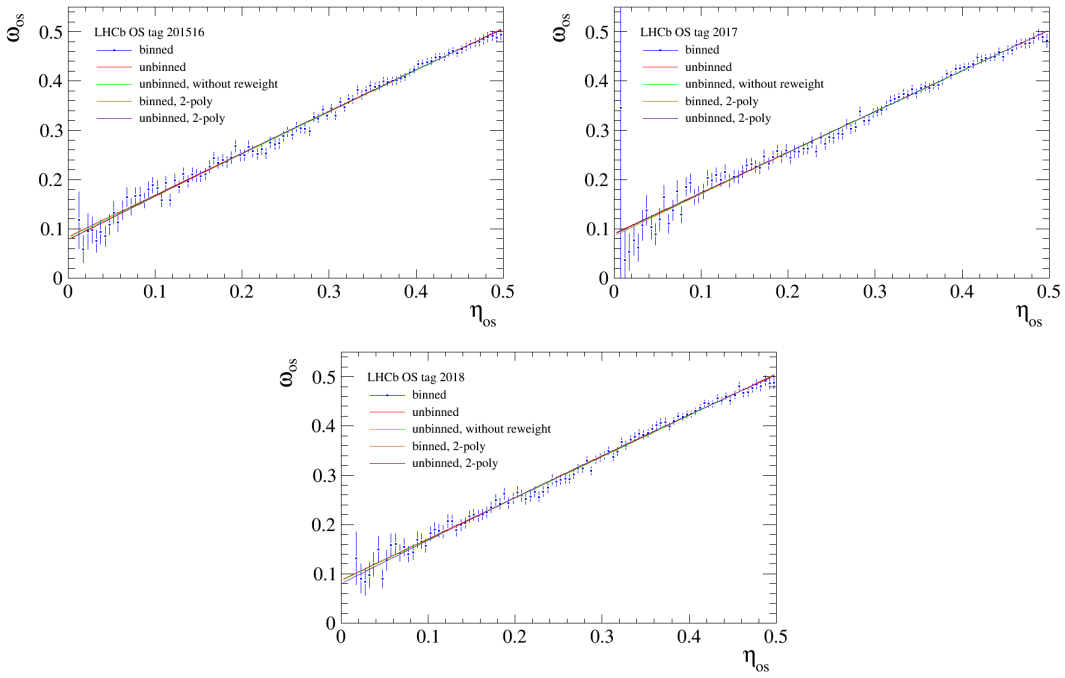

Figure 4.34: Calibration of the mistag probability η^{OS} of the opposite side algorithm. The ω^{OS} is fitted as a function of η^{OS}, using the binned and unbinned approaches for the different years.

two-order polynomial function, instead of the nominal exponential one, is used to subtract the combinatorial background using the *sPlot* technique. The calibration is then repeated and the difference in the parameters is set as systematic uncertainty, see Sec. 4.8. In order to check the portability of the calibration from the $B^+ \to J/\psi K^+$ data sample to the $B_s^0 \to J/\psi K^+K^-$ one, the calibration curves of truth-matched simulated samples of these two channels are obtained, using the *true* information about the initial flavour of the B_s^0 meson. Good consistency is observed and the differences in calibration parameters are treated as systematic uncertainties.

4.5.2.2 Calibration of the same-side algorithm

The calibration of the SS tagger has to be performed using a B_s^0 decay. This causes the development of a more complex strategy compared to the OS case, since it is subject to $B_s^0 - \overline{B}_s^0$ oscillations. The most used channel in LHCb is the $B_s^0 \to D_s^+\pi^-$ decay. The calibration is performed by resolving the $B_s^0\overline{B}_s^0$ oscillations with a time-dependent maximum likelihood fit of the $B_s^0 \to D_s^+\pi^-$ decay rate $\Gamma(t)$. After being properly selected, in order to subtract all the background contributions, the $B_s^0 \to D_s^+\pi^-$ data sample is weighted to match the distributions of P_T, rapidity, number of tracks and number of PVs of the selected $B_s^0 \to J/\psi K^+K^-$ signal sample. This ensures the portability of the calibration from the $D_s^+\pi^-$ mode to the signal decay. Then the sample is divided into eight bins of the estimated mistag probability η and a

simultaneous unbinned maximum likelihood fit is performed to the different bins. The PDF is defined as

$$PDF(t|q_{mix}) \propto \epsilon(t)\Big(\Gamma(t|q_{mix}) \otimes R(t)\Big), \tag{4.27}$$

where

$$\Gamma(t|q_{mix}) = e^{-\Gamma_s t}\Big(\cosh(\Delta\Gamma_s t/2) + q_{mix}(1-2\omega(\eta))\cos(\Delta m_s t)\Big). \tag{4.28}$$

The parameter q_{mix} takes the values $= +1$ (-1) if B_s^0 meson has (has not) changed flavour between its production and decay, determined by comparing the charge of the final state pion and the tag decision of the SS algorithm. The relation between the true mistag probability and the estimated one is the same used for the OS tagger calibration, defined in Eq. 4.25. In the fit, the $\epsilon(t)$ and $R(t)$ functions account for the decay-time acceptance and resolution effects, respectively. The acceptance is modelled with the empirical function $\epsilon(t) = 1 - 1/(1+(at)^n + b)$, while the resolution is modelled as a single Gaussian function, where the width is defined event-by-event as the calibrated DTF decay-time uncertainty. The decay-time uncertainty calibration is taken from the Δm_s analysis, see Ref. [119]. The results of the fit are shown in Fig.4.35 for the 2018 data sample. A binned fit is also performed. Furthermore, the same procedure as for the OS tagging calibration is performed to estimate the systematic uncertainty associated to the portability from the calibration to the signal sample. A good consistency is observed.

4.5.2.3 Tagging combination and its performance on the $B_s^0 \to J/\psi K^+K^-$ sample

The calibration curves just described are used to correct the estimated mistag probability of the $B_s^0 \to J/\psi K^+K^-$ candidates. Indeed, the probability that an initial $B_s^0(\overline{B}_s^0)$ meson is correctly tagged is

$$p(B_s^0) = \frac{1}{N}(1 + q^{OS}(1-2\omega^{OS})(1 + q^{SS}(1-2\omega^{SS}), \tag{4.29a}$$

$$p(\overline{B}_s^0) = \frac{1}{N}(1 - q^{OS}(1-2\overline{\omega}^{OS})(1 - q^{SS}(1-2\overline{\omega}^{SS}), \tag{4.29b}$$

where $q^{OS/SS}$, $\omega^{OS/SS}$ are the tag decisions and the calibrated mistag probabilities of the two tagging algorithms and $N = p(B_s^0) + p(\overline{B}_s^0)$ is a normalisation factor.
As it has been shown in Chap. 3, the effect of a non perfect tagging performance is the reduction, of a factor D^2, of the size of a sample. So the performance of the tagging algorithm directly acts on the final precision achievable in the measurement.
The effective tagging power ϵD^2 quantifies the tagging performance and, for a sample with a single tagger per-event calibrated mistag probability ω_i, it can be calculated as

$$\epsilon D^2 = \frac{\sum_{tagged} v_i(1-2\omega_i)^2}{\sum_{all} v_i}, \tag{4.30}$$

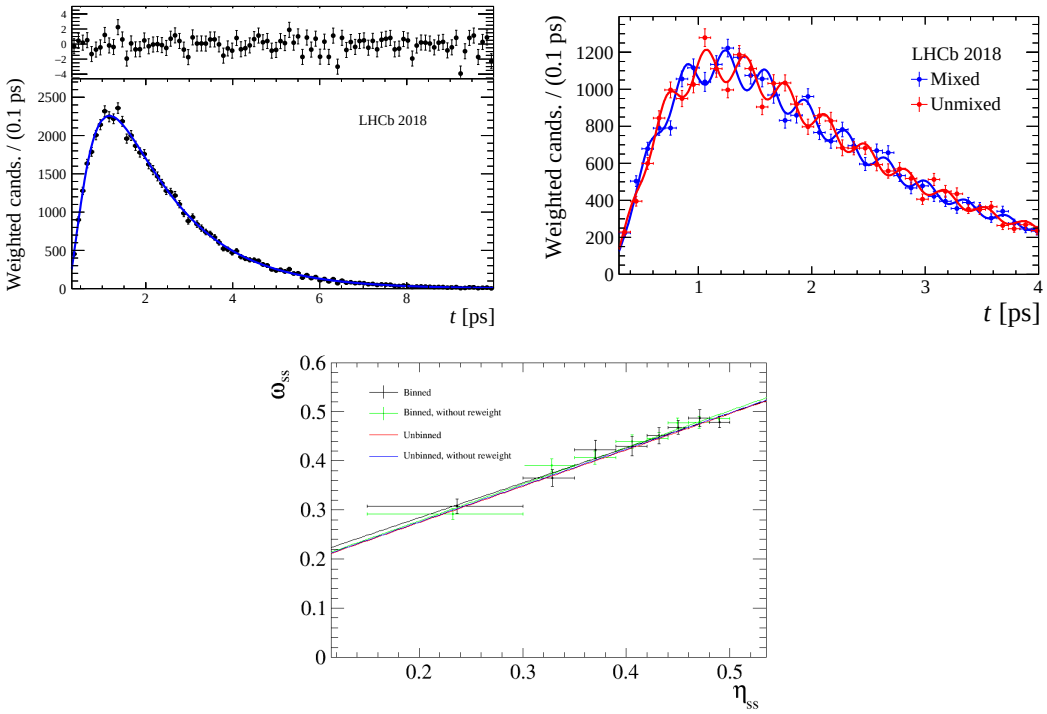

Figure 4.35: Top: the $B_s^0 \to D_s^+\pi^-$ decays time distribution overlaid with the fitted curve on the left, and the distribution of the mixed and unmixed decays shown separately on the right. Bottom: calibration curve for the same side algorithm. All the plots refer to the 2018 year.

where v_i is a generic per-event weight (due to a reweighting procedure or to the application of the *sPlot* technique), and the sums are extended to the tagged events or all events. To take into account both the tagging algorithms we can consider that the dilution factor can be expressed in terms of the probabilities $p(B_s^0)$ and $p(\overline{B}_s^0)$ as $D = (1 - 2\omega_i) = |p(B_s^0) - p(\overline{B}_s^0)|$. So, the effective tagging power of the combined tagged events can be calculate as

$$\epsilon D^2 = \frac{\sum_{tagged} v_i \Big(p(B_s^0) - p(\overline{B}_s^0)\Big)^2}{\sum_{all} v_i}. \tag{4.31}$$

The tagging performances for the $B_s^0 \to J/\psi K^+K^-$ decay mode are summarised in Tab. 4.6. Significantly higher combined tagging performances are obtained with respect to Run 1, which was of $(3.73 \pm 0.15)\%$ [100].

4.5.3 Decay-time acceptance

The displacement requirements asked in the trigger selection on the final state particles or on the decay vertex of the B_s^0 introduce a non-trivial acceptance shape for low values of decay time (typically < 5 ps). In addition, a small acceptance effect is present also at high values

Table 4.6: Tagging performances, where with the flag $SS/OS-only$ is meant all the events tagged by the SS/OS algorithm and with $OS+SS$ is meant all the events tagged by both the algorithms.

Category(201516)	Faction(%)	$\varepsilon(\%)$	D^2	$\varepsilon D^2(\%)$
$OS-only$		11.34	0.0778 ± 0.0044	0.88 ± 0.02
$SS-only$		42.58	0.0234 ± 0.0044	1.00 ± 0.18
$OS+SS$		23.93	0.0967 ± 0.0058	2.31 ± 0.14
$Total$		77.85	0.0513	4.19 ± 0.23
Category(2017)	Faction(%)	$\varepsilon(\%)$	D^2	$\varepsilon D^2(\%)$
$OS-only$		11.08	0.0773	0.86 ± 0.01
$SS-only$		43.16	0.0224	0.97 ± 0.19
$OS+SS$		24.43	0.0975	2.38 ± 0.09
$Total$		78.67	0.0535	4.21 ± 0.22
Category(2018)	Faction(%)	$\varepsilon(\%)$	D^2	$\varepsilon D^2(\%)$
$OS-only$		11.17	0.0798	0.89 ± 0.01
$SS-only$		42.98	0.0244	1.05 ± 0.20
$OS+SS$		24.77	0.0983	2.43 ± 0.09
$Total$		78.92	0.0535	4.37 ± 0.22

of decay time, where tracks with large impact parameter with respect to the beam line are not efficiently reconstructed by the VELO reconstruction algorithm. Instead of relying on simulations, the decay-time acceptance is studied by using a data-driven method. A b-meson decay very similar to the signal channel is used as a control mode, in order to extract the shape of the acceptance. As it has been explained in Sec. 4.3, the well-known decay $B^0_d \to J/\psi K^*(892)^0$, with $K^*(892)^0 \to K^+\pi^-$, is used for this purpose.

4.5.3.1 Decay-time acceptance measurement

In order to determine the decay-time acceptance shape $\epsilon(t)$ of a given sample of events, the fully selected decay-time distribution and its corresponding *true* decay-time distribution are compared. The *true* distribution corresponds to the generated one, without any effect caused by the reconstruction or selection procedures. Since $\Delta\Gamma_d$ is compatible with zero, the *true* decay-time distribution for the B^0_d control mode is obtained from a single exponential function with the time constant equal to the well-know lifetime of $B^0_d \to J/\psi K^*(892)^0$ decay. To be able to use the same simple model also for the B^0_s simulation samples MC sample with decay-width difference $\Delta\Gamma_s = 0$ ps has been generated, and a systematic uncertainty will be evaluated to account for this approximation. Using this strategy, the decay-time acceptances in B^0_s MC and B^0_d MC and data samples can be obtained. Then the acceptance for B^0_s signal decay, $\epsilon^{data}_{B^0_s}$, can

be determined as

$$\epsilon_{B_s^0}^{data} = \epsilon_{B_d^0}^{data} \times \frac{\epsilon_{B_s^0}^{MC}}{\epsilon_{B_d^0}^{MC}}. \tag{4.32}$$

The term $\epsilon_{B_d^0}^{data}$, is the decay-time acceptance in data of the fully triggered, selected, sWeighted and kinematically weighted events in $B_d^0 \to J/\psi K^{*0}$ control channel. The ratio $\epsilon_{B_s^0}^{MC}/\epsilon_{B_d^0}^{MC}$, is the ratio of acceptances of the simulated signal and control channels after being fully triggered, selected, sWeighted and kinematically weighted. This term accounts for the small differences in the lifetime and kinematics between the B_s^0 and B_d^0 modes. The sWeight procedure in MC samples is necessary because also the MC events with BKGCAT=60 are included, as explained in Sec. 4.2.3, and the sample contains a small fraction of background. The reweighting procedure mentioned above is described in Sec. 4.5.3.2.

In order to be less sensitive to statistical fluctuations that are present when using directly the acceptance histograms, an analytical parametrisation is used. This allows a direct implementation of the decay-time acceptance into the final PDF that will be used for the measurement of the *CP*-violating and mixing parameters. A cubic spline function is chosen for this purpose. It is defined on a set of intervals, whose extremes are called knots. In each interval, the cubic spline is given by a third order polynomial and, at every knot position, the spline is required to be continuously differentiable. If we consider a cubic spline with N knots, it can be parametrised with N + 2 parameters. In this analysis the cubic spline $s(t)$ is written as a linear combination of N + 2 independent basic cubic splines $b(t)$, called b-splines

$$s(t) = \sum_{i=0}^{N+2} c_i b_i(t), \tag{4.33}$$

where c_i are the parameters that uniquely define the cubic spline $s(t)$. The knots are positioned at [0.3, 0.91, 1.96, 9.00] ps and this choice accounts for the exponential nature of the distribution. In fact, the first three time bins are almost equally populated and in the last one, 9-15 ps, the spline is extrapolated with a linear function, in order to be less dominated by the statistical fluctuations of the less populated bin.

The determination of the decay-time acceptance in the signal data is achieved by a simultaneous fit to the decay-time distributions of the $B_d^0 \to J/\psi K^{*0}$ data sample and the simulated samples of $B_d^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi K^+K^-$ decays. Each distribution is described by the *true* exponential function convoluted with a Gaussian resolution function and multiplied with an acceptance. The Gaussian resolution model is centred at zero and in MC samples the resolution widths are determined following the method discussed in Sec. 4.5.1. The ratio between the B_s^0 and B_d^0 widths observed in the simulation samples is then used to scale the B_s^0 data resolution widths, in order to obtain the B_d^0 one in data samples. The acceptances of the three samples are

parametrised as a function of the respective cubic splines as

$$\epsilon^{MC}_{B^0_s} = s^{MC}_{B^0_s}(t), \tag{4.34a}$$

$$\epsilon^{MC}_{B^0_d} = s^{MC}_{B^0_s}(t) \times s^{MC}_{B^0_d/B^0_s}(t), \tag{4.34b}$$

$$\epsilon^{data}_{B^0_d} = s^{data}_{B^0_s}(t) \times s^{MC}_{B^0_d/B^0_s}(t). \tag{4.34c}$$

Thanks to the aforementioned modellisation of the acceptances, the simultaneous fit allows to extrapolate directly the B^0_s data spline coefficients and their uncertainties, which allows an easy control of the associated systematic uncertainty.

The decay-time acceptance is obtained separately for each year of data taking and for the two different trigger categories, corresponding to `unbiased` and `biased`. The spline components for the different years and trigger categories are shown in Fig. 4.36. In addition, for comparison, the plots show the B^0_s data acceptance histograms, obtained using Eq. 4.32 to the histogram acceptances of B^0_s MC and B^0_d MC and data samples. In this case, the decay-time resolution effect is ignored.

The measured biased acceptances show the typical drop at low decay time due to the trigger selection, while, for both the trigger categories, the effect due to the VELO track reconstruction at high values of the decay time is visible. An additional small effect is present in both the subsamples, around $2-3$ ps, although it is not visible in the biased one because of the dominant acceptance effect already present at low decay time values. The dip is caused by events in which the B^0_s secondary vertex is reconstructed as an additional PV. If this happens, the reconstructed $B^0_s \to J/\psi K^+K^-$ decay does not pass the vertex separation requirements made by the trigger selection. Since the minimum number of tracks used to form a possible PV is five, in order to be reconstructed as a PV, the B^0_s vertex has to gain another track. This is more likely for B^0_s decays with low decay time. However, if the B^0_s has a very low decay time, its final state tracks are likely used in the reconstruction of the true PV, explaining why the dip is not visible at very low values of the decay time. Furthermore, in the LHCb reconstruction software, a vertex cannot be separated more than a certain maximal distance from the beam axis in order to be classified as a PV. For low multiplicity vertices this distance is 0.2 mm, requirement that is easily exceeded by B^0_s mesons with high decay time. It is interesting to note that this effect strongly depends on the transverse momentum of the B^0_s meson. If the B^0_s has a low transverse momentum, it is more likely that its decay vertex fulfil the requirement of the maximal allowed distance to the beam axis, being therefore more likely reconstructed as an additional PV.

The procedure to extract the decay-time acceptance is validated using two b-meson decays with a well-known lifetime. A first test is performed using the $B^0_d \to J/\psi K^{*0}$ decay. Both MC and data samples are split into two independent subsamples, one used as signal and one used as a control sample. Initially, a random splitting is used. Then additional requirements are made, in order to better mimic the differences between the B^0_s signal and the B^0_d control decays existing in the standard procedure, namely

- a cut on the per-event decay-time uncertainty σ_t: the sample used as signal mode has $\sigma_t < 0.04$ ps,

- a cut on the K^{*0} invariant mass: the sample used as signal mode has $m(K^{*0}) > 980$ MeV/c^2.

For each set of cuts, the decay-time acceptance is determined from the control sample and a maximum likelihood fit to the decay-time distribution of the B_d^0 background-subtracted data sample treated as signal is performed, and the lifetime is measured. The test is made with and without the application of the correction procedure, consisting in the reweighiting of the MC samples, as it will be discussed in Sec. 4.5.3.2. The lifetimes obtained are reported in Tab 4.7 and compared to the world average, $\tau(B_d^0) = (1.520 \pm 0.004)$ ps [13]. The deviations of the obtained values with respect to the world average one are shown in brackets. Good agreement in these tests gives confidence that the decay-time acceptance for $B_s^0 \to J/\psi K^+K^-$ is correctly determined. However, the large deviations seen in the test when splitting based on σ_t are currently under investigation.
A second test is performed using the $B^+ \to J/\psi K^+$ decay as a signal channel. The obtained decay-time acceptance and a fixed decay-time resolution obtained in the same way as for the B_s^0 decay are used in the maximum likelihood fit to the decay-time distribution of the background-subtracted $B^+ \to J/\psi K^+$ data sample. In the fit, the decay-width difference between the B^+ and B_d^0, $\Delta\Gamma_d^u$, is a free parameter, while the decay-width of B_d^0 is fixed to the world-average value. The measured $\Delta\Gamma_d^u$ is then translated into the ratio $\tau(B^+)/\tau(B_d^0)$. The obtained values are summarised in Tab. 4.8 and compared to the current world-average of $\tau(B^+)/\tau(B_d^0) = 1.076 \pm 0.004$ [13].

Table 4.7: Values of $\tau(B_d^0)$ obtained with (C) and without (U) a reweighiting of the MC samples for the validation of the time acceptance in 2015-2018.

Year	Event	$m(K^{*0})$	σ_t
2015 C	1.515 ± 0.014 (0.3σ)	1.520 ± 0.016 (0.0σ)	1.493 ± 0.016 (1.7σ)
2015 U	1.514 ± 0.014 (0.4σ)	1.515 ± 0.016 (0.3σ)	1.477 ± 0.015 (2.8σ)
2016 C	1.5292 ± 0.0055 (1.7σ)	1.5304 ± 0.0063 (1.6σ)	1.4791 ± 0.0066 (6.2σ)
2016 U	1.5273 ± 0.0054 (1.4σ)	1.5288 ± 0.0062 (1.4σ)	1.4896 ± 0.0065 (4.6σ)
2017 C	1.5218 ± 0.0064 (0.3σ)	1.5217 ± 0.0059 (0.3σ)	1.5211 ± 0.0072 (0.2σ)
2017 U	1.5199 ± 0.0063 (0.0σ)	1.5199 ± 0.0058 (0.0σ)	1.5333 ± 0.0072 (1.9σ)
2018 C	1.5245 ± 0.0056 (0.8σ)	1.5213 ± 0.0067 (0.2σ)	1.5248 ± 0.0070 (0.7σ)
2018 U	1.5222 ± 0.0055 (0.4σ)	1.5186 ± 0.0065 (0.2σ)	1.5366 ± 0.0069 (2.4σ)

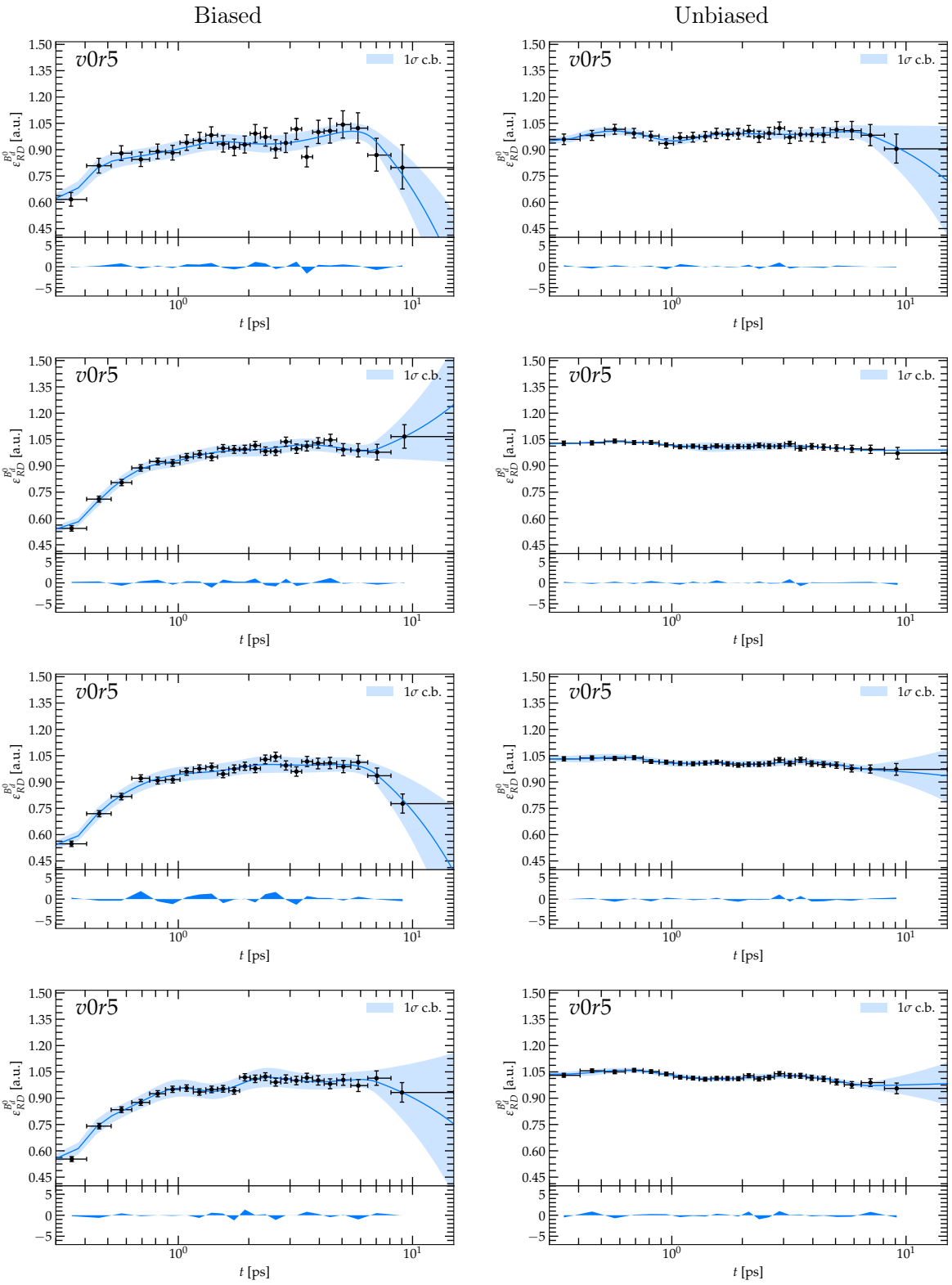

Figure 4.36: B_s^0 data decay-time acceptance normalised plots for each trigger category and year (from top to bottom 2015 to 2018), with the fit of a cubic spline function superimposed, as well as the 1σ confidence band.

Table 4.8: Values of $\tau(B^+)/\tau(B_d^0)$ (world average = 1.076 ± 0.004) obtained for the validation of the time acceptance, with and without the reweighiting procedure to correct the MC samples.

Year	With corrections	Without corrections
2015	1.0750 ± 0.0074 (0.1σ)	1.0763 ± 0.0069 (0.0σ)
2016	1.0824 ± 0.0039 (1.6σ)	1.0817 ± 0.0038 (1.5σ)
2017	1.0863 ± 0.0041 (2.5σ)	1.0852 ± 0.0040 (2.3σ)
2018	1.0866 ± 0.0044 (2.4σ)	1.0888 ± 0.0038 (3.3σ)

The lifetime ratios seem to increase with data taken years, leading to a larger discrepancy with the expected value. This behaviour is not completely understood yet but it is under investigation.

4.5.3.2 Reweighting procedure for the decay-time acceptance determination

Before extracting the B_s^0 decay-time acceptance, the B_d^0 simulated and data samples, as well as the B_s^0 simulated sample, are weighted. Initially, the B_d^0 data sample is weighted to match the respective B_s^0 data distributions of momentum and transverse momentum of the B meson. The reweighting in the momentum is performed in order to ensure that the relation between decay time and flight distance is the same for both channels, ignoring the small difference between the two masses. The weighting in the transverse momentum is motivated by the strong dependence of the decay-time acceptance on this quantity, as discussed in the previous section. Then, the simulated B_s^0 and B_d^0 samples are corrected to be more similar to the respective data samples. Initially they are weighted to have the same fraction of the two magnet polarities as the corresponding data samples. Then, the fact that the S-wave component is not present in the simulated samples is addressed. This represents the most important difference to be corrected for. In addition, the relative phases and fractions of the different polarization states do not necessarily agree with the ones observed in real data. In order to correct for these differences, the simulated samples are weighted to match the S-wave and polarization fractions and phases measured in earlier LHCb analyses [100, 120]. The reweighiting procedure is divided into two steps. Initially, a per-event weight is determined by evaluating the decay-time-dependent angular PDF, in one case with the physics parameters set to the ones used to produce the simulation sample, and in the other case with the parameters observed in the previous analyses. The ratio of these two values is then used as a per-event weight ω

$$\omega_i = \frac{PDF(t_i, \theta_{K_i}, \theta_{l_i}, \varphi_i/data)}{PDF(t_i, \theta_{K_i}, \theta_{l_i}, \varphi_i/sim)}. \tag{4.35}$$

As it will be explained in Sec. 4.7, in this way also the fraction of the S-wave component is corrected, since it is taken into account in the PDF. However, in order to ensure a optimal correction for the S-wave discrepancies, also the mass distribution of the hadron system is weighted in simulation to match the data distribution. Furthermore, the simulated samples

are also weighted in order to remove the differences with data in the transverse momentum distributions.
The reweighting procedure is summarised in the scheme shown in Fig. 4.37.

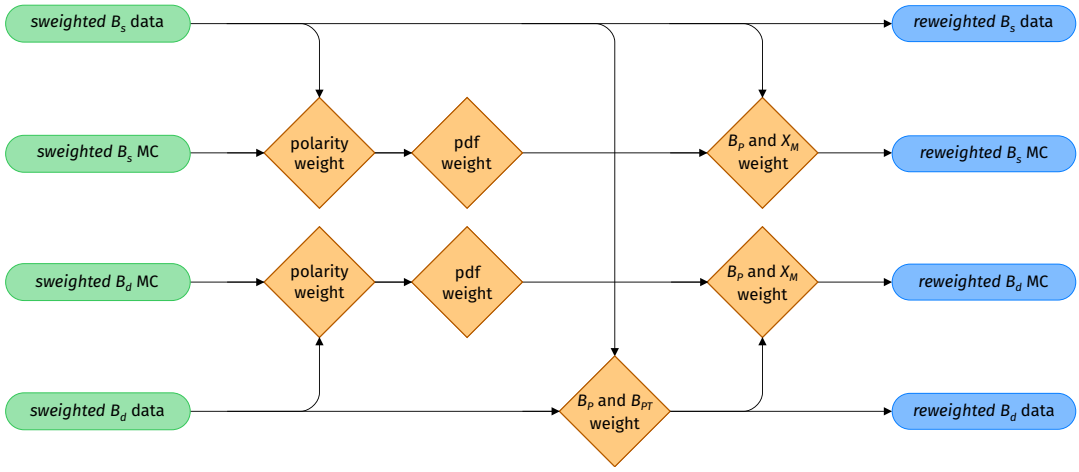

Figure 4.37: Reweighting procedure scheme used to correct the samples with the determination of the decay-time acceptance.

4.5.4 Angular acceptance

The principal sources of the angular acceptance are the geometrical acceptance of the detector, the selections in the momentum performed by the magnet and by the offline and online reconstruction of the tracks in the final state. In order to correct for this effect, the angular acceptances are studied by using simulated signal samples. A common strategy used by LHCb collaboration is to weight the selected sample in order to match the angular distributions of a non biased target dataset. The simpler method is to compute the weights extracting the angular acceptance from the comparison between the angular distributions of samples produced with (observed sample) and without (expected sample) the biasing selection applied. For this purpose the GB reweighting tool, discussed in Sec. 4.3.2, can be used. Otherwise, the analytical shape of the acceptance can be studied, by fitting the ratio between biased and unbiased angular distributions

$$\epsilon(\theta_\mu, \theta_K, \varphi) = \frac{N^{obs}(\theta_\mu, \theta_K, \varphi)}{N^{exp}(\theta_\mu, \theta_K, \varphi)}. \tag{4.36}$$

Again, the important aspect to correct for is the shape and not the overall normalisation, that can be neglected. The per-event weight is then defined as

$$\omega_i(\theta_\mu, \theta_K, \varphi) = \frac{1}{\epsilon(\theta_\mu, \theta_K, \varphi)}. \tag{4.37}$$

This method is not simple, because a multi-dimensional angular function should be determined, since the total acceptance is not factorisable in the product of three single-angle acceptances.

Both methods have the disadvantage that depend on the specific setting used: the GB reweighting settings, for the first method, and the model used to fit the acceptance shape for the second. For this reason, in this analysis a complete model-independent approach has been employed, that is not subject to any model-related systematic uncertainty. All the details on the method are available in Ref. [121]. It is based on the assumption that the angular and decay-time acceptances can be factorised, and it consists on assigning to each event the same set of average weights, one for each of the ten terms of the differential decay rate. Namely, the single weight is computed as described in the following

$$\omega_k = \frac{1}{N_{exp}} \sum_{i}^{observed} \frac{f_k(\Omega_i)}{PDF^{exp}(\Omega_i|t_i)}, \tag{4.38}$$

where N_{exp} and $PDF^{exp}(\Omega_i|t_i)$ are, respectively, the number of the expected events and their theoretical physical decay rate distribution evaluated at their decay time, see Eq. 1.75. The function f_k are the angular functions, described in Tab. 1.3, evaluated using the observed angles. This weights will be introduced as normalisation weights of the final PDF used to perform the measurement of the *CP*-violating and mixing parameters.

Since the angular acceptance is determined from simulation, it is important to check that the simulated sample correctly describes the data. Therefore, the MC selected sample is weighted to have the same fraction of the two magnet polarities as the corresponding data sample. Then the MC sample is corrected, using the GB reweighting tool, for its differences in kinematic variables with respect to the data sample, to which the angular acceptance is strongly sensitive. Namely, the variables are the momentum and the transverse momentum of the B_s^0 and the dikaon invariant mass. The last correction takes into account the absence of the S-wave component in simulations. After this procedure, some minor differences are still visible, mainly due to a mismodel of the angular distributions in simulation at generator level caused by the non correct generation of the dikaon structure. To correct for this, an iterative weighting procedure is used, which gradually matches the generated differential decay rate PDF with that obtained using the parameters extracted fitting the data. During this iterative procedure, the MC sample is also weighted to match the momentum and transverse momentum distributions of kaons in data. The iterative procedure runs till a convergence is reached, and the obtained angular weights are the ones used in the final fit. The procedure is performed separately for each trigger category, so two sets of angular weights are derived. Figure 4.38 shows a schematic illustration of the iterative procedure used to obtain the angular acceptance from simulated $B_s^0 \to J/\psi K^+K^-$ samples.

The procedure of extracting the angular acceptance weights is validated using the $B_d^0 \to J/\psi K^*(890)^0$ data and simulation samples. Indeed, this decay mode has the same angular structure of the signal channel, being the decay of a pseudo scalar particle to two vector particles. After applying the entire procedure described above, the polarisation amplitudes of the $J/\psi K^{*0}$ system have been measured by fitting simultaneously the distribution of helicity angles for

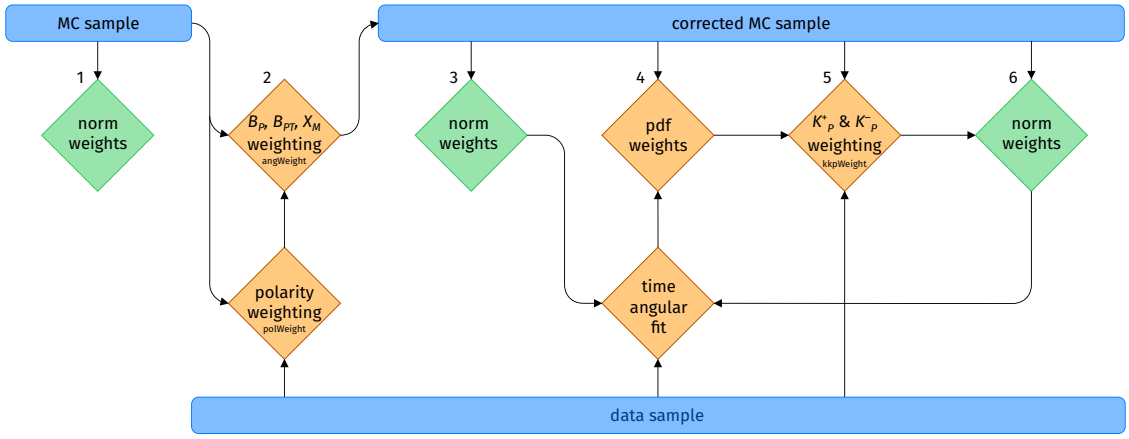

Figure 4.38: Schematic illustration of the iterative procedure used to obtain the angular acceptance from simulated $B^0_s \to J/\psi K^+K^-$ samples.

all the years, where in the PDFs the amplitudes are treated as common parameters. The result is shown in Tab. 4.9 and it is compared to the previous measurement done inside the LHCb experiment, see Ref. [120]. A good agreement is seen, confirming the validity of the method.

Table 4.9: Parameters of the angular fits to the $B^0_d \to J/\psi K^{*0}$ data samples, compared to the values obtained in Ref. [120]. The number in brackets shows the difference between the result obtained in this work and those in the reference, in number of standard deviations.

Parameters	This analysis (stat. only)	Ref [120]
$\|A_\parallel\|^2$	0.22240 ± 0.00026 (-0.39σ)	0.227±0.004 ±0.011
$\|A_\perp\|^2$	0.21013 ± 0.00083 (+1.02σ)	0.201±0.004 ±0.008
$\delta_\parallel - \delta_0$	-2.94895 ± 0.00532 (-0.25σ)	-2.94 ±0.02±0.03
$\delta_\perp - \delta_0$	2.91339 ± 0.00404 (-0.94σ)	2.94±0.02 ±0.02
Γ_d	0.65514 ± 0.00058	...

Finally, the factorisation of the angular and the decay-time acceptances is investigated. For this check, the angular acceptance weights are recalculated in equally populated bins of the B^0_s decay time. Fig 4.39 shows the comparison between the projections of the time-dependent angular efficiencies and the nominal one in the different helicity angles, in the 2018 dataset and for each trigger category. A dependence on the time is observed. The impact of this effect is under study, but preliminary results show a low impact in the determination of the parameters of interest and a related small systematic uncertainty will be assign.

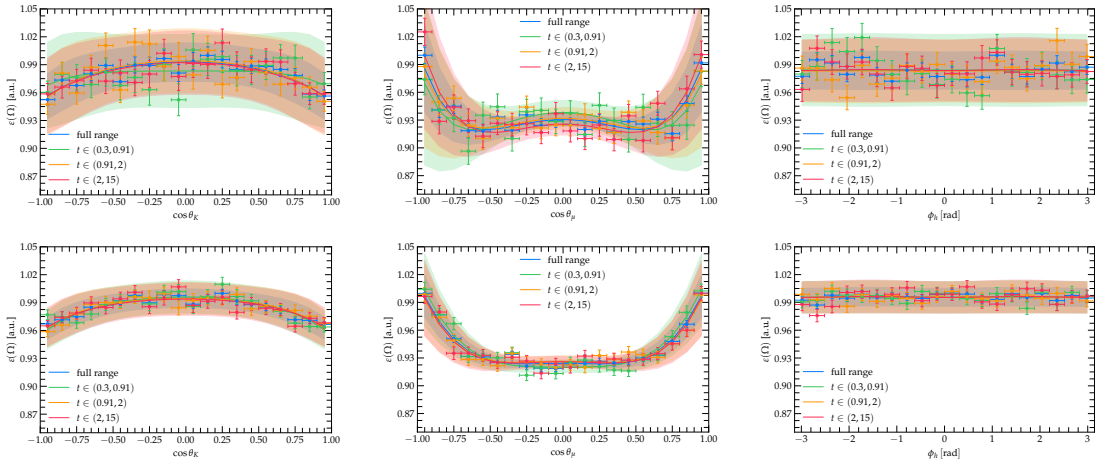

Figure 4.39: Time dependence of angular efficiency for 2018, in biased (top) and unbiased (bottom) trigger categories. The plots show the 1σ confidence band of the nominal angular acceptance and of the angular acceptance in bins of decay time.

4.6 The fit procedure

The physics parameters ϕ_s, $\Delta\Gamma_s$ and Γ_s are determined by fitting the four-dimensional distribution of the selected signal events, where the different dimensions are the decay time and the three helicity angles. The model used to describe the distribution is given by the theoretical differential decay rate, presented in Sec. 1.3.2, properly modified to take into account the experimental effects treated in the previous sections. This chapter covers the explanation and construction of the final fit. At the end, the preliminary results obtained by performing the fit to the full Run 2 data sample of $B_s^0 \to J/\psi K^+K^-$ signal candidates are presented.

4.6.1 Inclusion of the $S-wave$ component

The inclusion of the S-wave component in the description of the differential decay rate, allows mainly to resolve the twofold ambiguity in the measurement of ϕ_s and Γ_s, as described in Sec. 1.3.2. For this reason, the final fit is performed simultaneously in six bins of the K^+K^- invariant mass, taking into account the different contributions of the S-wave and P-wave amplitudes. The chosen binning scheme is: $[990, 1008, 1016, 1020, 1024, 1032, 1050]\text{MeV/c}^2$. The dependence of the four amplitudes, A_j, on the dikaon invariant mass can be expressed as

$$A_{0,\perp,\parallel}(m_{KK}) = A_{0,\perp,\parallel}\; p(m_{KK}), \tag{4.39}$$

for the P-wave amplitudes, with $j = 0, \perp, \parallel$, and

$$A_S(m_{KK}) = A_S\; s(m_{KK}) \tag{4.40}$$

for the S-wave one. The line shapes of the P- and S-wave are denoted as $p(m_{KK})$ and $s(m_{KK})$, respectively. Both are normalised to unity over the dikaon mass range. In this way, the dependence on the dikaon mass is introduced only in the interference terms between the P and the S-wave amplitudes

$$\int_{m_{KK}^L}^{m_{KK}^H} A_j A_S^* \, p(m_{KK}) s^*(m_{KK}) dm_{KK} = |A_j A_s| C_{SP} \, e^{-i\theta_{SP}}. \tag{4.41}$$

While the phase θ_{SP} is absorbed by the strong phase δ_S , the real factor C_{SP} has to be calculated in order to correctly determine the fraction of the S-wave component as

$$C_{SP} e^{-i\theta_{SP}} = \frac{\int_{m_{KK}^L}^{m_{KK}^H} p(m_{KK}) \times s^*(m_{KK}) \; dm_{KK}}{\sqrt{\int_{m_{KK}^L}^{m_{KK}^H} |p(m_{KK})|^2 \; dm_{KK} \int_{m_{KK}^L}^{m_{KK}^H} |s(m_{KK})|^2 \; dm_{KK}}}, \tag{4.42}$$

Here, the P-wave is described by a Breit-Wigner function with the mean and width fixed to those of the ϕ meson. The S-wave is assumed to be composed entirely by the f_0 resonance, which is described by a Flatté distribution. While Eq. 4.42 is defined in terms of the generated dikaon invariant mass, in the implementation the reconstructed one is used, causing the possible migration of candidates between bins. Then, the mass resolution has to be taken into account and it is incorporated as an efficiency correction, $\epsilon_i(m_{KK})$, of the C_{SP} factors according to

$$C_{SP} e^{-i\theta_{SP}} = \frac{\int_{m_{KK}^L}^{m_{KK}^H} p(m_{KK}) \times s(m_{KK})^* \times \epsilon(m_{KK}) \; dm_{KK}}{\sqrt{\int_{m_{KK}^L}^{m_{KK}^H} |p(m_{KK})|^2 \times \epsilon(m_{KK}) \; dm_{KK} \int_{m_{KK}^L}^{m_{KK}^H} |s(m_{KK})|^2 \times \epsilon(m_{KK}) \; dm_{KK}}}. \tag{4.43}$$

Commons efficiencies for all data-taking years are used. Figure 4.40 shows the common efficiency distributions as a function of the dikaon mass bin.

4.6.2 The probability density function

The theoretical signal PDF is obtained by normalising the time- and angular-dependent decay rate defined in Eq. 1.75

$$PDF^{theo}(t, \theta_K, \theta_\mu, \varphi_h) \;\; = \;\; \frac{\sum_{k=1}^{10} A_k h_k(t) f_k(\theta_K, \theta_\mu, \varphi_h)}{\int\int dt d\Omega \sum_{k=1}^{10} A_k h_{k,q}(t) f_k(\theta_K, \theta_\mu, \varphi_h)}, \tag{4.44}$$

where $\Omega = (\theta_K, \theta_\mu, \varphi_h)$ and q is the flavour at production of the b-meson. The functions A_k and $f_k(\Omega)$ are defined in Tab. 1.3, while $h_{k,q}(t)$ is defined in Eq. 1.76 and the related coefficients in Tab. 1.2. The parameters of interest are contained in the time-dependent functions. Considering that the studied data samples are affected from the experimental effects discussed in the previous section, the theoretical PDF is modified into an experimental one, as explained in the following. The experimental effects added to the theoretical PDF depend on the data taking year. This is

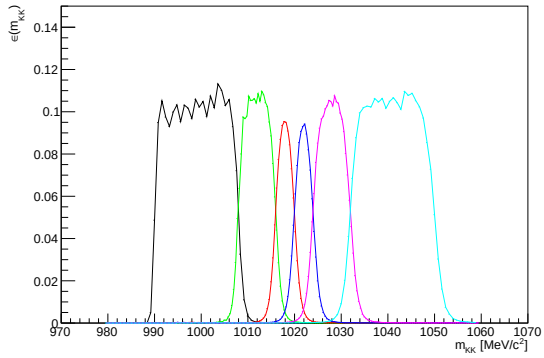

Figure 4.40: Efficiency of the selection in each m_{KK} bin as a function of MC true m_{KK} for the combined simulation sample for all years used for the calculation of the baseline C_{SP} factors.

considered in the following formulas by introducing the index $y = (2015, ..., 2018)$.
Including the tags and mistag probabilities resulting from the application of the tagging algorithms, the experimental PDF can be written as

$$\begin{aligned}\mathrm{PDF}^y(t,\Omega|q^{OS,SS},\eta^{OS,SS}) = \frac{1}{N}\sum_{k=1}^{10} A_k f_k(\theta_K,\theta_{mu},\varphi_h)\times \\ \Big[h_{k,B_s^0}(t)[(1+q^{OS}(1-2\omega(\eta^{OS}))(1+q^{SS}(1-2\omega(\eta^{SS})))] \\ + h_{k,\overline{B}_s^0}(t)[(1+q^{OS}(1-2\overline{\omega}(\eta)^{OS})(1+q^{SS}(1-2\overline{\omega}(\eta^{SS})))]\Big].\end{aligned}$$

Here, N is a normalisation factor and it is equal to the integral of the PDF in the four domensional space $N = \int\int \mathrm{PDF}^{exp}(t,\Omega|q^{OS,SS},\eta^{OS,SS})dtd\Omega$. If the decay-time resolution effect is taken into account, the PDF has to be convoluted with the effective resolution function

$$\begin{aligned}PDF^y(t,\Omega|q^{OS,SS},\eta^{OS,SS},\sigma_t) = \frac{1}{N}\sum_{k=1}^{10} A_k f_k(\theta_K,\theta_{mu},\varphi_h)\times \\ \Big[h_{k,B_s^0}(t)[1+q^{OS}(1-2\omega(\eta^{OS})(1+q^{SS}(1-2\omega(\eta^{SS})))] \\ + h_{k,\overline{B}_s^0}(t)[1+q^{OS}(1-2\overline{\omega}(\eta)^{OS})(1+q^{SS}(1-2\overline{\omega}(\eta^{SS})))]\Big] \\ \otimes R(t,\sigma_{eff}(\sigma_t)),\end{aligned}$$

and the normalisation changes accordingly. Moreover, the acceptances effects have to be considered. Additionally to the data taking year, both depend on the trigger categories, explicitly considered with the index c. The PDF is multiplied to the spline function of the decay-time acceptance, $\epsilon^c(t)$, and the per-event angular acceptance weights, ω_k^c, are considered in the

normalisation factor. The final PDF is therefore

$$PDF^{y,c}(t,\Omega|q^{OS,SS},\eta^{OS,SS},\sigma_t) = \frac{1}{N}\sum_{k=1}^{10} A_k f_k(\theta_K,\theta_{mu},\varphi_h)\epsilon^c(t)\times$$
$$\Big[h_{k,B_s^0}(t)[1+q^{OS}(1-2\omega(\eta^{OS})(1+q^{SS}(1-2\omega(\eta^{SS})))]$$
$$+ h_{k,\overline{B}_s^0}(t)[1+q^{OS}(1-2\overline{\omega}(\eta)^{OS})(1+q^{SS}(1-2\overline{\omega}(\eta^{SS})))]\Big]$$
$$\otimes R(t,\sigma_{eff}(\sigma_t)),$$

and

$$N^{y,c} = \int dt \int d\Omega \sum_{k=1}^{10} A_k\omega_k^c\ f_k(\theta_K,\theta_{mu},\varphi_h)\epsilon^c(t)\times$$
$$\Big[h_{k,B_s^0}(t)[1+q^{OS}(1-2\omega(\eta^{OS})(1+q^{SS}(1-2\omega(\eta^{SS})))]$$
$$+ h_{k,\overline{B}_s^0}(t)[1+q^{OS}(1-2\overline{\omega}(\eta)^{OS})(1+q^{SS}(1-2\overline{\omega}(\eta^{SS})))]\Big]$$
$$\otimes R(t,\sigma_{eff}(\sigma_t)).$$

In order to take into account the S-wave contribution and therefore resolve the twofold ambiguity in the differential decay rate, the amplitudes A_k, for $k \in (8,9,10)$ are multiplied by the C_{SP} factor.

4.7 Results

In this section, the preliminary results of a maximum likelihood fit of the background-subtracted data samples is presented. The fit is performed using the signal PDF described in the previous section, simultaneously in each trigger category, each year of data taking and in six bins of the dikaon invariant mass. The measured parameters are the CP violating parameters ϕ_s and λ, the difference in decay width and mass between the two mass eigenstates of the $B_s^0 - \overline{B}_s^0$ system, $\Delta\Gamma_s$ and Δm_s and the polarisation amplitudes and their strong phases. Instead of measuring the average decay width for the B_s^0 meson, Γ_s , the difference between the average decay width for the B_s^0 and B_d^0 mesons is measured; $\Gamma_s - \Gamma_d$. This is done because in the determination of the decay-time acceptance, the data sample of the control channel $B_d^0 \to J/\psi K^{*0}$ is used. Instead of including a systematic for the uncertainty on the Γ_d world average, this strategy allows to extract a value that is independent on the particular value of Γ_d used in the acceptance computation. Moreover, from a theoretical point of view, lifetime differences are significantly more precisely determined thanks to a cancellation of many systematic effects, making this comparison more meaningful. The sum of the squared P-wave amplitudes is defined to be equal to unity, so the parallel amplitude can be determine as

$$|A_\parallel|^2 = 1 - |A_0|^2 - |A_\perp|^2. \tag{4.49}$$

Instead of fitting for the S-wave amplitude $|A_S|^2$, the fraction of the S-wave component with respect to the total amplitude, F_S, is determined for each dikaon bin:

$$F_s = \frac{|A_S|^2}{1 + |A_S|^2}. \tag{4.50}$$

Since the analysis described in this thesis is not finished yet by the time of writing, the value of the two main parameters of interest, ϕ_s and $\Delta\Gamma_s$, are still "blinded", meaning that they are shifted during the measurement by an arbitrary quantity, in order to avoid to introduce unintentional bias in the development of the analysis procedure. One of the ingredients that is still missing in this version of the analysis is the correction for the decay-time bias effect, described in Sec. 3.1.3, that is still under investigation by the collaboration. The fit is performed assuming common ϕ_s and λ values for all the polarisation states. Currently the possibility to publish polarisation-dependent results is under study. Tab. 4.7 shows the preliminary results for all physics parameters together with their statistical uncertainties. Since the analysis is carried out by a large collaboration of scientists, different fitters have been developed. This allows to exclude the presence of a biasing effect due to the specific fit. The different fitters results are compared and an excellent agreement is visible, especially for the so-called Heidelberg and Santiago fitters. Thanks to the obtained precision for ϕ_s and Γ_s, these measurements will supersed the results of the previous analysis [5], improving the LHCb combined values shown in Sec. 4.1 and playing a key role in the precision of the future world average value.

4.8 Major systematic uncertainties

The precision of the measurement discussed in this chapter is statistically dominated, as it has been observed in the previous measurement [5]. So the impact of the systematic uncertainty due to the analysis procedure is very low. Anyway, with the improvement of the LHCb reachable luminosity foreseen in the future upgrades, the impact of the systematic uncertainties due to the analysis strategy will become always less negligible with respect to the statistical precision. For this reason, an accurate analysis of the possible systematic sources is fundamental. Since the measurement is still ongoing, the systematic uncertainties are currently under study and the evaluation has not been completed yet. Anyway, the strategy of the analysis is very similar to the previous measurement of ϕ_s in $B_s^0 \to J/\psi K^+K^-$ decay mode [5], so the sources of systematic uncertainty are largely the same. In this section, relying on the strategies developed in the previous analysis, the most important sources of systematic uncertainties and their determination are briefly discussed.

Mass factorisation The suppression of the combinatorial background with the *sPlot* technique discussed in Sec. 4.2.6, is based on the assumption of no correlation between the mass of background and signal events and the variables used to perform the final measurement.

Table 4.10: Comparison of the fit results obtained with the different fitters. The numbers in the last columns show the difference in each parameter as a percentage of the statistical error.

Parameter	Heidelberg	Nikhef	Santiago	H vs N	H vs S	N vs S
ϕ_s[rad]	0.293 ± 0.022	0.293 ± 0.021	0.293 ± 0.021	0.3%	0.9%	0.6%
$\lvert\lambda^0\rvert$	1.0063 ± 0.0096	1.0063 ± 0.0096	1.0063 ± 0.0096	0.1%	0.2%	0.0%
$\Delta\Gamma_d^s$[ps^{-1}]	-0.0086 ± 0.0013	-0.0086 ± 0.0014	-0.0087 ± 0.0013	0.2%	0.9%	0.7%
$\Delta\Gamma_s$[ps^{-1}]	0.2672 ± 0.0044	0.2671 ± 0.0044	0.2672 ± 0.0044	0.3%	0.5%	0.7%
Δm_s[ps^{-1}]	17.739 ± 0.032	17.739 ± 0.032	17.739 ± 0.032	0.9%	0.4%	1.3%
$\lvert A_\perp\rvert^2$	0.2458 ± 0.0023	0.2459 ± 0.0023	0.2458 ± 0.0023	1.4%	0.3%	1.7%
$\lvert A_0\rvert^2$	0.5189 ± 0.0017	0.5189 ± 0.0017	0.5189 ± 0.0017	0.4%	0.1%	0.5%
C_{SP_1}	0.467 ± 0.024	0.468 ± 0.024	0.467 ± 0.024	0.4%	0.5%	1.0%
C_{SP_2}	0.0421 ± 0.0049	0.0417 ± 0.0049	0.0420 ± 0.0051	7.9%	1.7%	6.2%
C_{SP_3}	0.0035 ± 0.0015	0.0033 ± 0.0015	0.0033 ± 0.0016	7.0%	8.0%	0.9%
C_{SP_4}	0.0045 ± 0.0026	0.0045 ± 0.0026	0.0044 ± 0.0025	0.4%	2.2%	1.8%
C_{SP_5}	0.0560 ± 0.0071	0.0561 ± 0.0071	0.0560 ± 0.0071	1.2%	0.9%	2.0%
C_{SP_6}	0.152 ± 0.011	0.152 ± 0.011	0.152 ± 0.011	0.5%	1.1%	1.7%
$\delta_\parallel - \delta_0$[rad]	3.130 ± 0.062	3.131 ± 0.062	3.131 ± 0.061	0.5%	1.8%	1.2%
$\delta_\perp - \delta_0$[rad]	2.767 ± 0.073	2.768 ± 0.073	2.769 ± 0.073	0.7%	1.6%	0.9%
$\delta_{S1} - \delta_\perp$[rad]	2.03 ± 0.13	2.03 ± 0.13	2.02 ± 0.13	4.8%	4.5%	9.3%
$\delta_{S2} - \delta_\perp$[rad]	1.71 ± 0.18	1.71 ± 0.19	1.70 ± 0.19	0.1%	5.0%	4.9%
$\delta_{S3} - \delta_\perp$[rad]	0.98 ± 0.31	0.96 ± 0.31	1.01 ± 0.37	5.1%	10.4%	15.3%
$\delta_{S4} - \delta_\perp$[rad]	-0.16 ± 0.12	-0.18 ± 0.12	-0.16 ± 0.11	15.4%	1.6%	14.5%
$\delta_{S5} - \delta_\perp$[rad]	-0.596 ± 0.066	-0.601 ± 0.066	-0.597 ± 0.065	8.4%	1.3%	7.2%
$\delta_{S6} - \delta_\perp$[rad]	-0.987 ± 0.076	-0.991 ± 0.076	-0.989 ± 0.077	6.0%	2.2%	3.8%
p_0^{os}	0.3912 ± 0.0027	0.3912 ± 0.0027	0.3912 ± 0.0027	0.1%	0.2%	0.1%
Δp_0^{os}	0.0090 ± 0.0014	0.0090 ± 0.0014	0.0090 ± 0.0014	0.0%	0.0%	0.0%
p_1^{os}	0.840 ± 0.025	0.840 ± 0.025	0.840 ± 0.025	0.2%	0.4%	0.2%
Δp_1^{os}	0.015 ± 0.012	0.015 ± 0.012	0.015 ± 0.012	0.0%	0.4%	0.4%
p_0^{ss}	0.4421 ± 0.0046	0.4421 ± 0.0046	0.4421 ± 0.0046	0.5%	0.5%	0.0%
Δp_0^{ss}	-0.013 ± 0.027	-0.013 ± 0.027	-0.013 ± 0.027	0.0%	0.8%	0.8%
p_1^{ss}	0.632 ± 0.057	0.632 ± 0.057	0.632 ± 0.057	0.5%	0.6%	1.1%
Δp_1^{ss}	0.002 ± 0.030	0.002 ± 0.029	0.002 ± 0.030	0.0%	0.1%	0.2%

While the correlation between the signal mass shape and cosθ_μ is already taken into account by parametrising the mass resolution as a function of the DTF mass uncertainty, correlations with the other helicity angles and decay time have to be verified. For this reason, the data sample will be split in bins of the variable under study and a new fit of the invariant mass could be made for each bin. New signal weights can be then computed using the *sPlot* technique in each bin. The final measurement will be repeated using the new weights and the difference in the parameters from this and the nominal measurement will be set as systematic uncertainty. From the previous measurement [5], this turned out to be the major source of systematic uncertainty in the determination of ϕ_s, λ, $\Delta\Gamma_s$ and Δm_s.

Multiple candidates The chance to produce two $B_s^0 \to J/\psi K^+K^-$ signal candidates within one pp collision is extremely small. However, in the data set of this analysis roughly 1.5% of the events contain more than one candidate. These candidates are mostly due to fake B_s^0

events, which have been created usually by combining the J/ψ coming from the b-meson with at least one random kaon. So these fake multiple candidates are in general distributed like combinatorial background events and, therefore, are rejected by the background-subtraction procedure. Sometimes, they can also create a peak below the true B_s^0 mass region, when the wrongly chosen kaon is built from a clone track of the true kaon track. Therefore these fake events are selected as signal events. In order to evaluate the possible systematic uncertainty due to their presence in the fitted final data sample, the following procedure will be used. If two candidates are produced in the same collision it is likely that one is the clone of the other. If this is true, the opening angle between the track of a given final state particle of one candidate and the track of the same final state particle of the other candidate will be very small. This is true for all the tracks in the final state. Then all the events with possible multiple candidates where all the pairs of final state tracks have opening angles smaller than a chosen value could be removed. The background subtraction with the *sPlot* technique, the acceptance determination and the final fit can be repeated. Any difference in the new results with respect to the nominal one will be assigned as systematic uncertainty.

Angular acceptance The angular acceptance weights are determined from simulation, therefore it is important to correct them from any difference with the data sample. This is done reweighting the MC samples, as described in Sec. 4.5.4. The specific strategy used in the analysis could lead to systematic uncertainties on the final fit parameters. In order to estimate their magnitude, alternative reweighting corrections could be used. The first test could be done including the reweighting for the momentum and transverse momentum of the B_s^0 and for the dikaon invariant mass in the iterative procedure, while nominally these are only reweighted at the beginning of the procedure. The reason is that this ensures that the agreement between data and MC for these distributions do not get disturbed by the weighting in the kaon variables used in the iterative procedure. A second test could be performed, including also the transverse momentum of the two muons in the iterative procedure, to account for possible remaining differences. In addition, another possible systematic source could be due the specific configuration of the GB reweighting. For this, several alternative more and less powerful configurations could be used. For each test, the maximal deviation, in terms of fit results, from the nominal one has to be taken as systematic uncertainty.
Another source could be due to the finite size of the MC samples used to obtain the normalisation weights, that are subject on their statistical fluctuations. The related systematic uncertainty could be accessed by repeating the final fit many times, with normalisation weights that are randomly varied according to their covariance matrix. The spread of the obtained distributions of deviations in the fit parameters can be assigned as a systematic uncertainty.
The last main source of systematic effect could be due to the assumption that the angular acceptance is independent of the decay time. This could be not true if we consider the geometrical requirements due to the impact parameter selection performed in the trigger, that can

create a dependence between the angle and the separation of the B_s^0 decay vertex from the primary vertex and therefore also on the decay time. This could be tested by calculating the normalisation weights in bins of the decay time.

Decay-time acceptance As for the angular acceptance, the decay-time acceptance suffers from the systematic uncertainty due to the finite size of the samples used to determine it. This causes an uncertainty on the spline parameters used to implement the acceptance in the fit. To translate this uncertainty to a systematic uncertainty on the parameters of the time-dependent angular fit, new spline coefficients are randomly generated according to their covariance matrices and the fit is repeated. This can be done many times and the standard deviation of the distributions of the difference between the new results and the nominal one can be used as an estimation of the systematic uncertainty.

CHAPTER 5

Measurement of the *CP*-violating phase ϕ_s with $B_s^0 \to J/\psi\pi^+\pi^-$ decays

In this chapter the ongoing measurement of the *CP*-violating phase ϕ_s using the second most important channel $B_s^0 \to J/\psi\pi^+\pi^-$ is presented. The $B_s^0 \to J/\psi\pi^+\pi^-$ decay topology, with a J/ψ decaying in two opposite charged muons, is shown in Fig. 4.3 and its total branching fraction is approximately 1×10^{-5}. The events collected by the LHCb experiment during the data taking period 2015-2018, at a center of mass energy of 13 TeV, have been analysed. The analysis strategy is almost equal to that discussed for the golden channel $B_s^0 \to J/\psi K^+K^-$ in Chapter 4. However, two major differences are present. First, the analysis of the *CP* components contributing to this decay is more complicated, since different resonances can contribute to the $\pi^+\pi^-$ system. For that reason, an amplitude analysis on the $\pi^+\pi^-$ system is needed. Second, due to the almost completely *CP*-odd nature of the decay [31], the Γ_s, $\Delta\Gamma_s$ and Δm_s variables cannot be determine with high sensitivity, while a precise measurement of Γ_H can be carried out.

The analysis is being performed in collaboration with other two scientists from a different LHCb group, who are mostly involved in the decay-time and angular acceptances determination, in the calibration of the flavour taggers and in performing the final fit. My contribution to this analysis consists into the development of selection of the signal and control modes, the correction of the related simulated samples, to look more similar to the corresponding data samples, and the fundamental measurement of the decay-time resolution. The work is being summarised in the analysis note *"CP violation in $B_s^0 \to J/\psi\pi^+\pi^-$ decays with Run 2 data"* [8].

The measurement is still ongoing, thus only the analysis steps already concluded will be

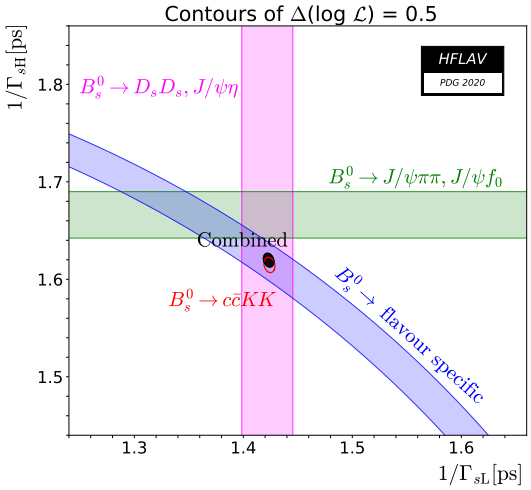

Figure 5.1: CL contours corresponding to 39% and 68% CL in the plane $(1/\Gamma_L, 1/\Gamma_H)$, for all the analyses performed on these observables up to date and their combination. Plot taken from Ref. [2].

discussed in this thesis. In particular, after a brief summary of the current state of the art of Γ_H measurements, the decay mode selection and the study of some selected experimental effects are explained in detail. Most of the concepts used in this chapter have been already met and explained in details in Chap. 4.

5.1 Current status of Γ_H measurements

The current state of the art of the Γ_H measurements is shown in Fig. 5.1. The constraint due to the average of all $B_s^0 \to J/\psi\phi$, $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to \psi(2S)\phi$ measurements, already mentioned in Sec. 4.1, and which depends on both on Γ_L and Γ_H, is shown as the red contour. The green, pink and blue bands are the constraints given by the effective lifetime measurements with CP-odd final states ($B_s^0 \to J/\psi f_0(980)$ [122] and $B_s^0 \to J/\psi\pi^+\pi^-$ [6], [103]), CP-even final states ($B_s^0 \to J/\psi\eta$ [123] and $B_s^0 \to D_s^+ D_s^-$[104]), and flavour-specific final states (mainly $B_s^0 \to D_s\ell X$ decays, where ℓ is a lepton), respectively. The average, taking all constraints into account, is shown as the black filled contour. A small tension in the Γ_H value between the $B_s^0 \to \bar{c}cK^+K^-$ and flavour specific measurements and the *CP*-odd ones is present. The current world average value of the heavy decay width is [2]

$$\Gamma_H = 1.616 \pm 0.010 \text{ ps}, \tag{5.1}$$

while the value obtained from *CP*-odd decays only is

$$\Gamma_H^{odd} = 1.666 \pm 0.024 \text{ ps}. \tag{5.2}$$

Therefore, improvements on the Γ_H precision will enable more stringent tests of the consistency between direct measurements of the heavy decay width and those inferred using B_s^0 decays into an admixture of *CP* eigenstates or flavour specific final states.

5.2 Selection of $B_s^0 \to J/\psi\pi^+\pi^-$ signal candidates

The selection of the signal candidates is developed following the same strategy explained in Sec. 4.2. It is divided in four main steps: a trigger, trigger and offline preselection, the rejection of misidentified decays, the suppression of the combinatorial background with a multivariate-based selection and the final subtraction of the remaining background sources.

5.2.1 Trigger and offline selection

The selection of the signal candidates is initially performed online. No L0 requirement is made, in order to retain more statistics. At the HLT1 stage, the charged particles tracks are reconstructed and information about their momentum and the PV position is obtained. Then, the events are selected by at least one of the following HLT1 trigger lines, as TOS: `Hlt1TrackMuonDecision`, `Hlt1DiMuonHighMassDecision` and `Hlt1TrackMVADecision`. The latter is similar to the `Hlt1TwoTrackMVADecision` line used in the $B_s^0 \to J/\psi K^+K^-$ analysis. Thanks to a multivariate algorithm, it manages to select the events in which a charged particle has been produced with a large transverse momentum and a displaced vertex with respect to any PV. The third step of the trigger selection exploits the complete reconstruction of the event performed by the HLT2 trigger stage. It consists into the selection of those events where two muons are present, coming from a J/ψ that is not consistent with being produced in the PV. For this purpose, only muon pairs that fulfil the HLT2 trigger line `HLT2DiMuonDetachedJPsi` requirements are kept. As explained in Sec. 4.2.1, the trigger selection introduces a non uniform efficiency as a function of the decay time of the b-meson. This effect has to be carefully studied, in order to rely on the performed measurement.
The $B_s^0 \to J/\psi\pi^+\pi^-$ candidate is created by combining, for each event selected by these trigger lines, two opposite charged muons and two opposite charged pions. In order to reject combinations that are not coming from a real decay, additional requirements are made. This is done offline by the LHCb experiment, requiring the events to pass the selection performed by the stripping line `StrippingB2JpsiHHBs2Jpsif0`, organised in the following hierarchical way:

- all the particles of the final state must have a good-quality track and a small probability to be a ghost particle. Then, some PID requirements are made, in order to reject possible background candidates caused by misidentified particles. Both muons and pions are required to have a relative high transverse momentum, in order to reduce the presence of soft particles coming from the PV. Since most of the particles produced in the PV are

pions, an additional cut on the χ^2_{IP} of selected pions is made, ensuring their displacement from the PV.

- The two muon tracks are used to build up a J/ψ candidate, on which other requirements are applied. Namely, the resulting invariant mass has to be around the nominal J/ψ mass and its decay vertex has to be of high quality. The two muon tracks must also have a relatively small distance of closest approach χ^2 (DOCA_χ^2), meaning that the two tracks should be close in space, keeping into account the uncertainty of the track position.

- The same requirements on vertex quality and DOCA_χ^2 are also made for the pions combination, plus an additional requirement in the $\pi^+\pi^-$ transverse momentum. No cut on the $\pi^+\pi^-$ invariant mass is performed, in order to select the full mass spectrum.

- The four tracks are then used to build up the B_s^0 candidate. The resulting decay vertex has to be of a good quality and the B_s^0 meson has to come from the PV and not from other secondary decay. This is achieved by requiring $\chi^2_{IP} < 25$ and a DIRA > 0.999, where the DIRA is the cosine of the angle between the flight-distance direction of the particle and the direction of its momentum. Finally, a cut on the minimum reconstructed DTF decay time is performed, to further reject wrong combinations of particles coming from the PV.

Table 5.1: `StrippingB2JpsiHHBs2Jpsif0` selection requirements, which are the same for all the 2015-2018 datasets.

#	Selection variables	Requirements
1	All tracks χ^2/ndf	< 5
2	All tracks GHOSTPROB	< 0.5
3	muon DLL($\mu - \pi$)	> 0
4	p_{T} of muon	> 500 MeV/c
5	J/ψ vertex χ^2_{vex}	< 16
6	J/ψ DOCA χ^2	< 20
7	J/ψ mass window	± 80 MeV/c
8	p_{T} of pion	> 250 MeV/c
9	PID of pion	DLL($\pi - K$) > -10 or `ProbNNpi`> 0.05
10	minimum π χ^2_{IP}	> 4
11	Sum p_{T} of $\pi^+\pi^-$	> 900 MeV/c
12	$\pi^+\pi^-$ vertex χ^2	< 16
13	$\pi^+\pi^-$ DOCA χ^2	< 20
14	B_s^0 Vertex χ^2_{vex}/ndf	< 10
15	B_s^0 χ^2_{IP}	< 25
16	B_s^0 Pointing, DIRA $\cos\theta_p$	> 0.999
17	B_s^0 decay time	> 0.2 ps
18	B_s^0 mass window	$[5150, 5670]$ MeV/c^2

Table 5.2: Analysis offline selection requirements

#	Selection variables	Requirements
1	π tracks GHOSTPROB	< 0.2
2	π $\chi^2_{\rm IP}$ OWNPV	> 0
3	J/ψ mass window	in $[3048.9, 3139.8]$ MeV/c^2
4	B_s^0 mass DTF with J/ψ mass constraint	> 0 MeV/c^2
5	B_s^0 $\chi^2_{\rm IP}$	> 0
6	B_s^0 Decay Vertex $\chi^2_{\rm vex}$/ndf	> 0
7	B_s^0 DTF Decay Vertex $\chi^2_{\rm vex}$/ndf	> 0
8	B_s^0 Flight Distance	> 0 mm
9	B_s^0 DTF decay time, t	> 0.3 ps
10	B_s^0 DTF decay time, t	< 14 ps
11	B_s^0 decay time error, δ_t	< 0.15 ps
12	Clone track rejection	$\theta_{\rm btw\, tracks\, in\, one\, candidate} > 0.5$ mrad

The details of the stripping selection are summarised in Tab. 5.1. After the stripping, a futher selection is performed, as summarised in Tab. 5.2. Sanity cuts on the B_s^0 flight distance, invariant mass, decay vertex, DTF χ^2, IP χ^2 and pions IP χ^2 are required, in order to reject poorly-reconstructed events. The number of combinatorial background events is further reduced by making tighter requirements on the ghost probability of pions and on the invariant mass of the J/ψ. Finally, fake candidates obtained from the combination of two or more clone tracks are rejected. Their presence in the selected sample is visible by looking at the φ helicity angle. When one positive (negative) final state track is a clone of the other positive (negative) track, the angle φ becomes equal to $\pm\pi$, creating some non-physical spikes. These events are vetoed requiring that the angle between the trajectories of the muon and the pion of the same charge is less than 5 mrad. The φ distribution in 2018, before and after the veto, is shown in Fig. 5.2 and the selection efficiency on signal decays, obtained from truth-matched* simulated samples, is 99.99%.

5.2.2 Misidentified background

Five main sources of misidentified backgrounds populate the $J/\psi\pi^+\pi^-$ invariant mass region below the B_s^0 mass peak; namely the $B^+ \to J/\psi K^+$, the $B_d^0 \to J/\psi K^{*0}$, the $B_d^0 \to J/\psi\pi^+\pi^-$, the $\Lambda_b \to J/\psi K p$ and the $B_s^0 \to J/\psi\eta'$ decays.
The $B^\pm \to J/\psi K^\pm$ decays can be wrongly identified as signal events if the kaon is misidentified as a pion and combined with a random pion produced in the pp collision. In order to identify the presence of the B^+ candidates in the data samples, the invariant mass is recomputed considering all the events in the samples as B^+ events. Two different masses can be recalculated, depending on the fact that we are considering the B^+ or the B^- decay mode. In the first case the negative

*A simulated event is considered a signal candidate if it has `BKGCAT` equal to 0 or 10 or 50. The category `BKGCAT` = 10 is considered to account for the complex resonant structure of the $\pi^+\pi^-$ system.

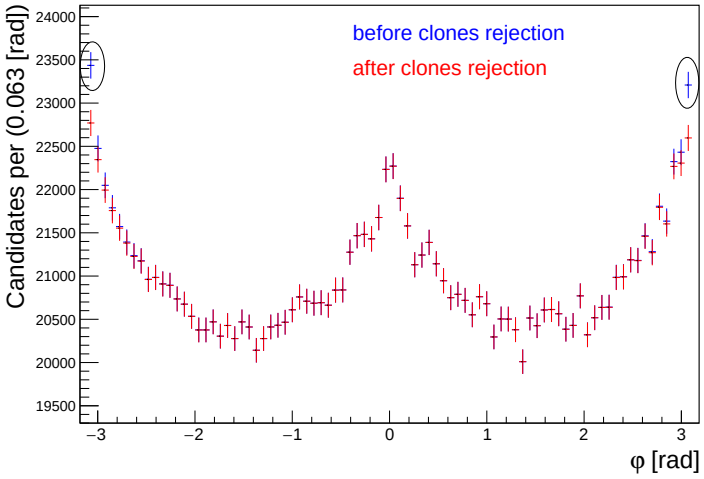

Figure 5.2: Distribution of the φ helicity angle in the 2018 dataset before and after the clone rejection. The clone contribution, before the application of the veto, is highlighted with a circle.

pion is not considered in the recalculation and the kaon mass is assigned to the positive one. In the second case, the positive pion is removed and the mass is recalculated assigning to the negative pion the kaon mass. Fig. 5.3 shows in blue the combined recomputed mass of $B^{\pm}$ candidate. An evident peak around the nominal mass of the B^+ meson, 5279.34 ± 0.12 MeV/c^2 [13] is visible, while the broad distribution below the peak is due to real B_s^0 decays and to combinatorial background events. This peaking background is largely suppressed by vetoing the candidate if one of the pion has `ProbNNk` > 0.1 and if the corresponding recalculated $J/\psi K^{\pm}$ mass is consistent with the known B^+ value within $\pm 3\sigma$, where the sigma is extracted by fitting the peak in Fig. 5.3 with a Gaussian function. The projection of the fit is shown in Fig. 5.3 by the red line. The fit, and consequently the veto, is performed separately year by year. The recalculated mass after the veto is showed in red in Fig. 5.3. The $B^{\pm}$ background is almost totally rejected, with an efficiency on the signal events of about 99.97% for all the years. This efficiency has been studied in truth-matched simulated samples, properly corrected to match PID and additional variables of the data. The correction of simulated samples is discussed in more detail in Sec. 5.2.3.

$B_d^0 \to J/\psi K^*(892)^0$ events, with the $K^*(892)^0 \to K^+\pi^-$, and its *CP* conjugated decay can be selected as signal decays if the kaon is misidentified as a pion. In order to recalculate the mass, the kaon mass is assigned to the pion with larger `ProbNNk`. The mass obtained is shown in blue in Fig. 5.4, and a clear peaking structure is present around the B_d^0 nominal mass value, 5279.65 ± 0.10 MeV/c^2 [13]. This background is vetoed requiring that all the events with a recalculated mass consistent with the B_d^0 mass value within $\pm 3\sigma$ are removed, and the distribution after the veto is shown in red in Fig. 5.4. Also in this case the σ is extracted by fitting the recomputed mass distribution with a Gaussian function. The veto efficiency on the

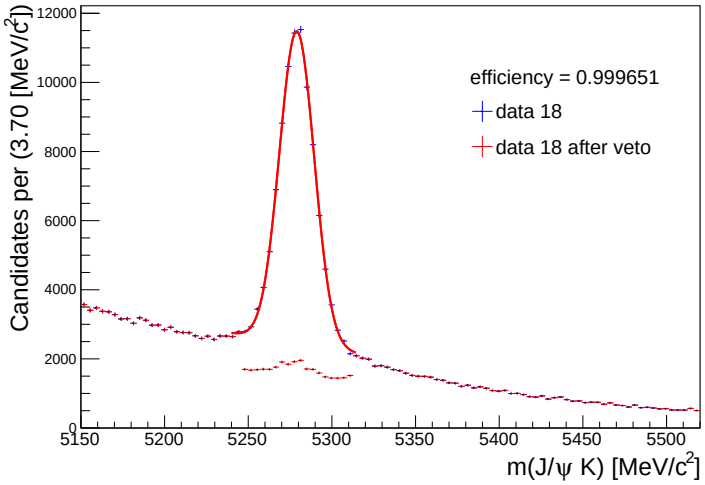

Figure 5.3: Distribution of the recomputed $m(J/\psi K)$ before, in blue, and after, in red, the application of the veto in 2018.

signal events is studied in corrected truth-matched simulated samples and it is about 99.87% for all the years.

The $\Lambda_b \to J/\psi K p$ decay is wrongly selected as a signal candidate if the kaon and the proton are misidentified as pions. The recomputed mass is shown in blue in Fig. 5.5. The red distribution shows the recalculated mass after the application of a veto and a not negligible number of events survive to the rejection. Indeed, this background is not efficiently removable by PID requirements without loosing too many signal candidates. Considering that its lifetime is very different from that of the B_s^0, a small contribution of this background could introduce a bias in the determination of Γ_H. For that reason, it has to be completely removed using a different strategy, discussed in the Sec. 5.2.5.1.

The $B_s^0 \to J/\psi\eta'$, with the $\eta' \to \rho(\to \pi^+\pi^-)\gamma$ is, instead, an irreducible background. It has the same visible final state of the signal, since the γ is not reconstructed, and can not be removed using the approach discussed so far. Thus, the same strategy adopted for the Λ_b background will be used for its rejection.

The $B_d^0 \to J/\psi\pi^+\pi^-$ background has the same final state of the B_s^0 decay and is shifted with respect to the nominal B_s^0 mass value by $\Delta m(B_s^0 - B_d^0) \sim 87$ MeV/c^2 . In order to remove this contribution it is enough to select only the signal events with an invariant mass in the region of ± 20 Mev/c^2 around the B_s^0 mass peak. This last selection will be applied at the end of the background subtraction procedure.

5.2.3 Correction of simulated samples

This analysis relies on simulated events in several places:

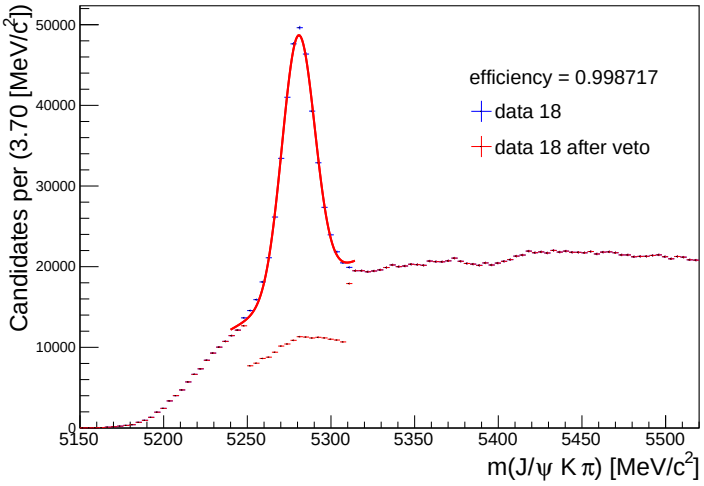

Figure 5.4: Distribution of the recomputed $m(J/\psi K\pi)$ before, in blue, and after, in red, the application of the veto in 2018.

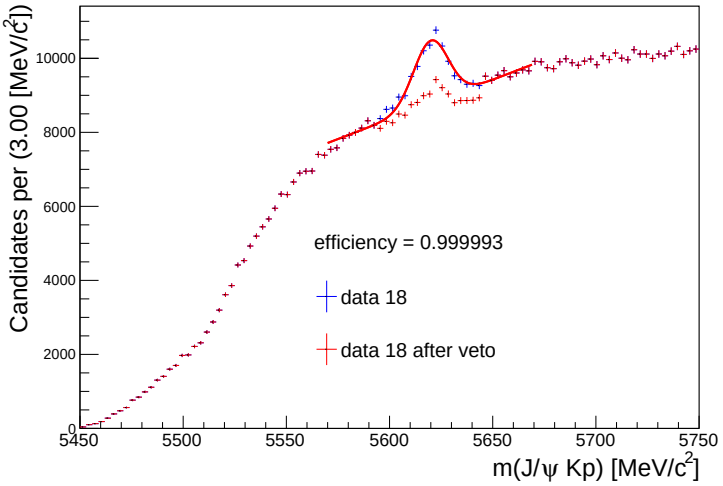

Figure 5.5: Distribution of the recomputed $m(J/\psi Kp)$ before, in blue, and after, in red, the application of the veto in 2018.

- evaluating the misidentified background rejection efficiencies;
- training a BDT algorithm to increase the combinatorial background rejection with high signal efficiency;
- studying the mass distribution shape;
- evaluating angular acceptances and determining a small correction to the decay-time acceptance, extracted with a data-driven method.

Therefore, it is essential that the simulation matches the data as closely as possible. Nevertheless, the B_s^0 production kinematics and the number of long tracks and PVs are usually not well described by the simulation. In addition, some variables that will be used to train the BDT show a poor agreement with data. Thus, a correction of the simulated samples is necessary. The same strategy discussed in Sec. 4.3.2 is used. Initially, all the PID variables used in the analysis are corrected using the `PIDCalib` package. Then, the simulated samples are weighted to match the respective data samples in the following variables:

- natural logarithm of the smallest IP χ^2 among the two pions (ln_minIPCHI2pi);
- B_s^0 transverse momentum (Bs_PT);
- the sum of the pions transverse momentum (sumPT);
- natural logarithm of the IP χ^2 of the B_s^0 (ln_BsIPCHI2);
- natural logarithm of the DTF fit χ^2 of the B_s^0 (ln_BsDTFVertexCHI2);
- the number of long tracks in an event (nLongTracks);
- the number of PVs (nPV).

The GBR method is used, with background subtracted data samples as target. These pure signal samples are obtained making usage of the *sPlot* technique, with the DTF B_s^0 invariant mass as a discriminating variable. The signal is modelled using a Gaussian function and the background is modelled using an exponential. The procedure is performed separately for all the years. The mass distributions, with the resulting fit projections superimposed, are shown in Fig. 5.6. The comparison between data and simulations before and after the correction is shown in Figs. 5.9, 5.7, 5.8 and 5.10, 5.12, 5.11. After the correction, the data-simulation agreement is significantly improved, allowing to rely on simulated samples in the different steps of the analysis.

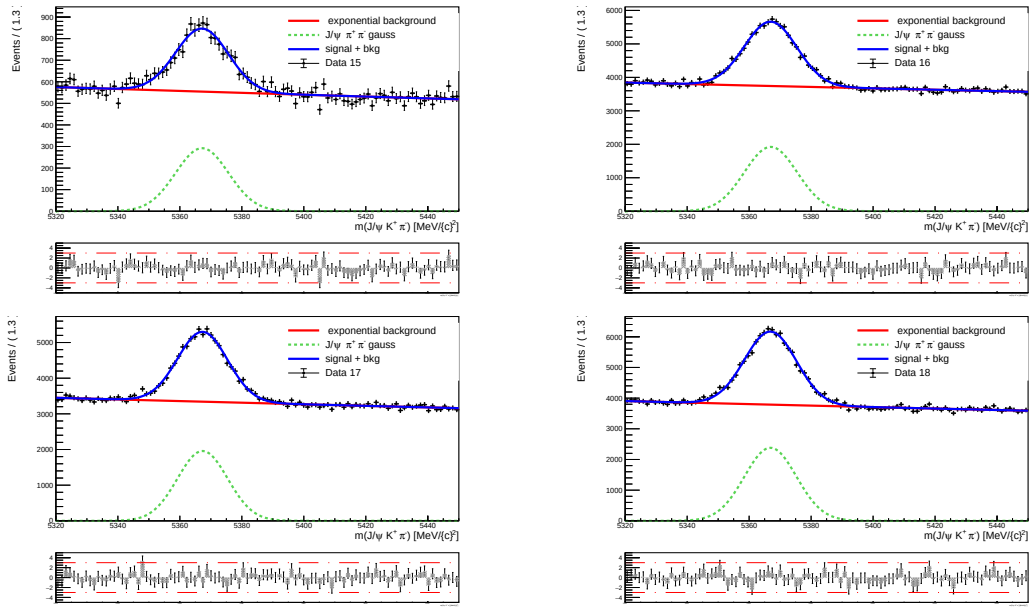

Figure 5.6: Distribution of the $J/\psi\pi^+\pi^-$ invariant mass with the fit projections superimposed for different years of data taking in Run 2.

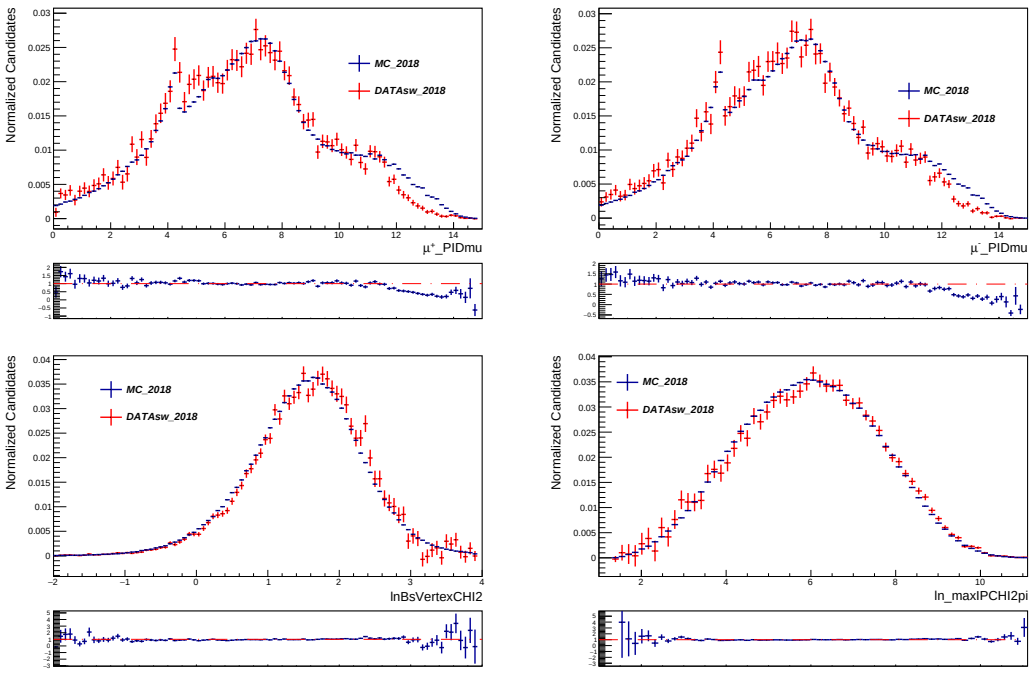

Figure 5.7: Distribution of selected variables used to train the BDT, before the data-simulation correction, for the 2018 dataset.

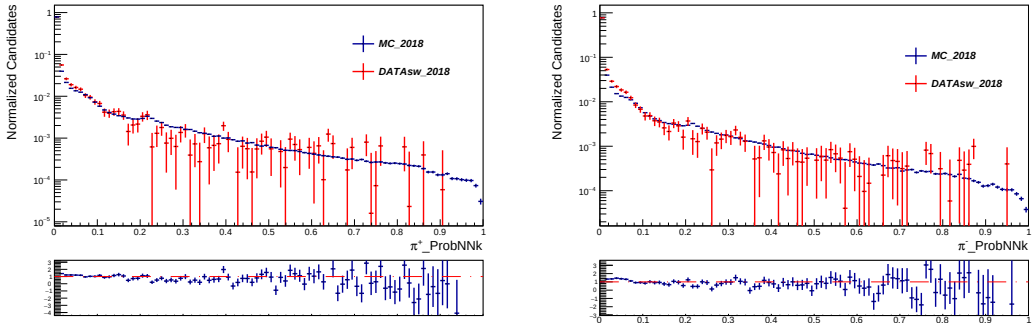

Figure 5.8: Distribution of PID variables used to veto misidentified backgrounds, before the data-simulation correction, for the 2018 dataset.

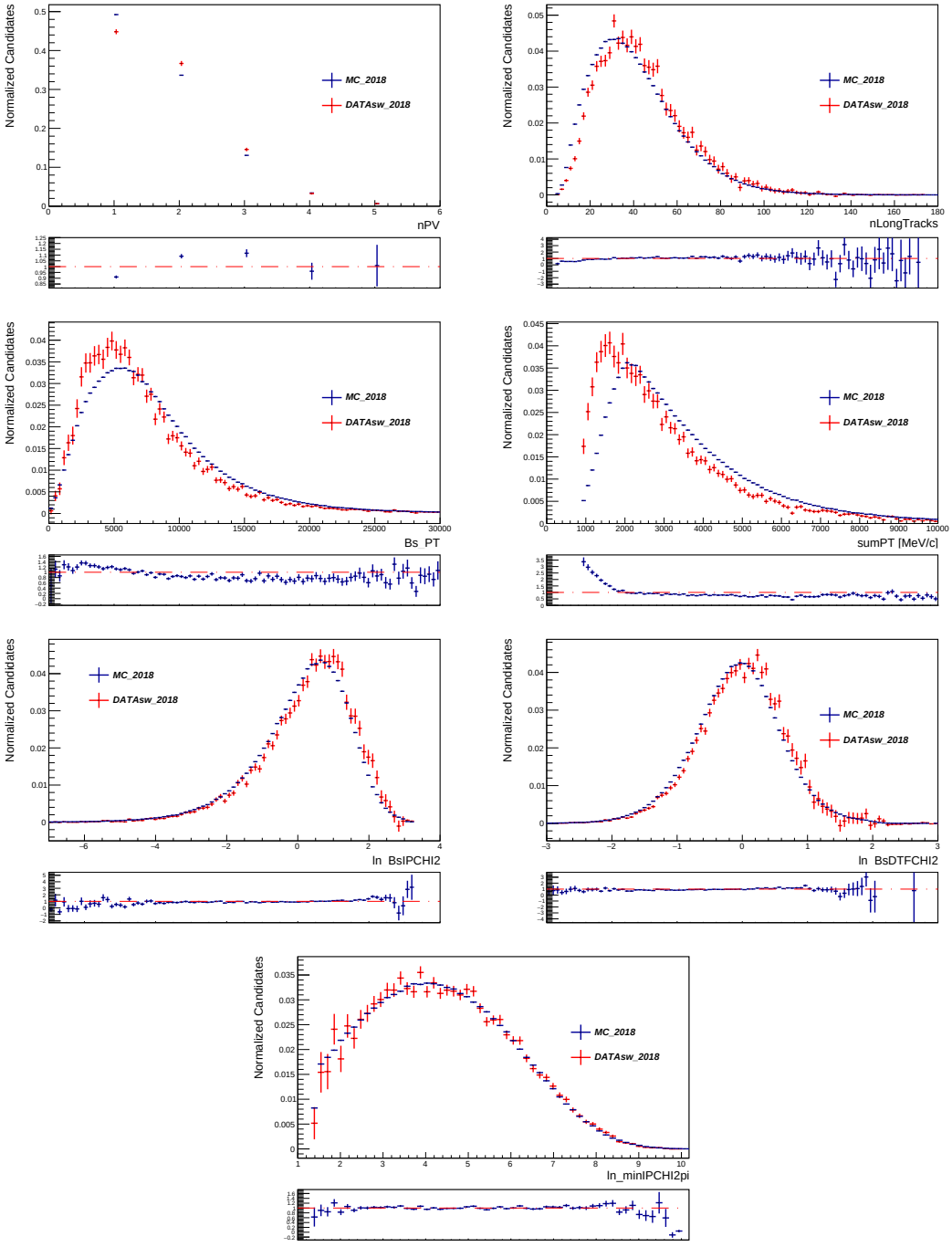

Figure 5.9: Distribution of the variables that show the largest disagreement between data and simulation, before any correction, for the 2018 dataset.

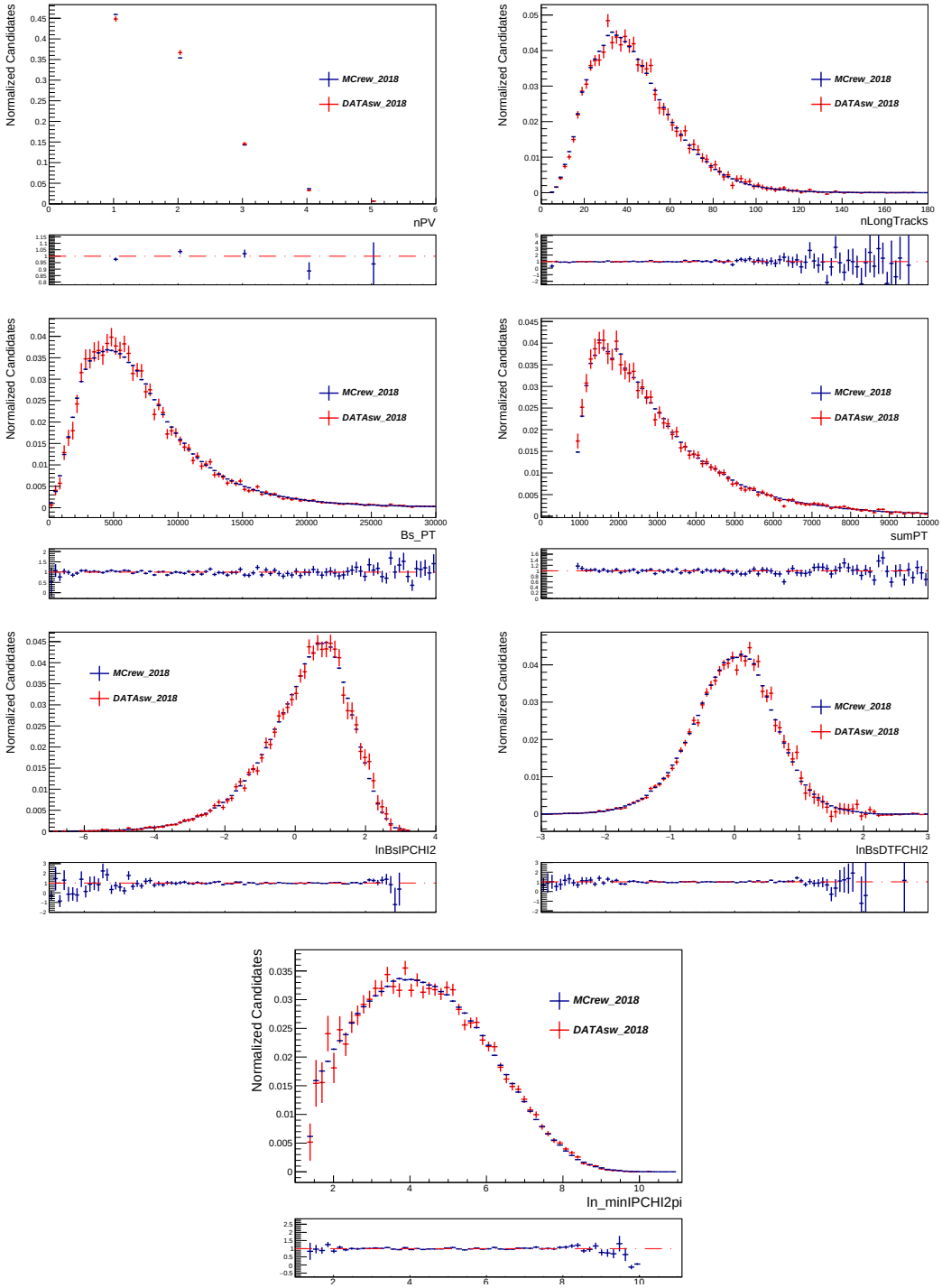

Figure 5.10: Distribution of the variables that show the largest disagreement between data and simulation, after the correction, for the 2018 dataset.

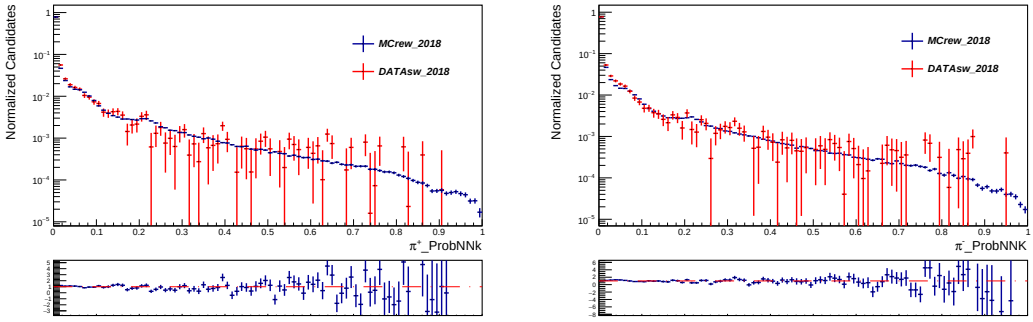

Figure 5.11: Distribution of PID variables used to veto misidentified backgrounds, after the data-simulation correction, for the 2018 dataset.

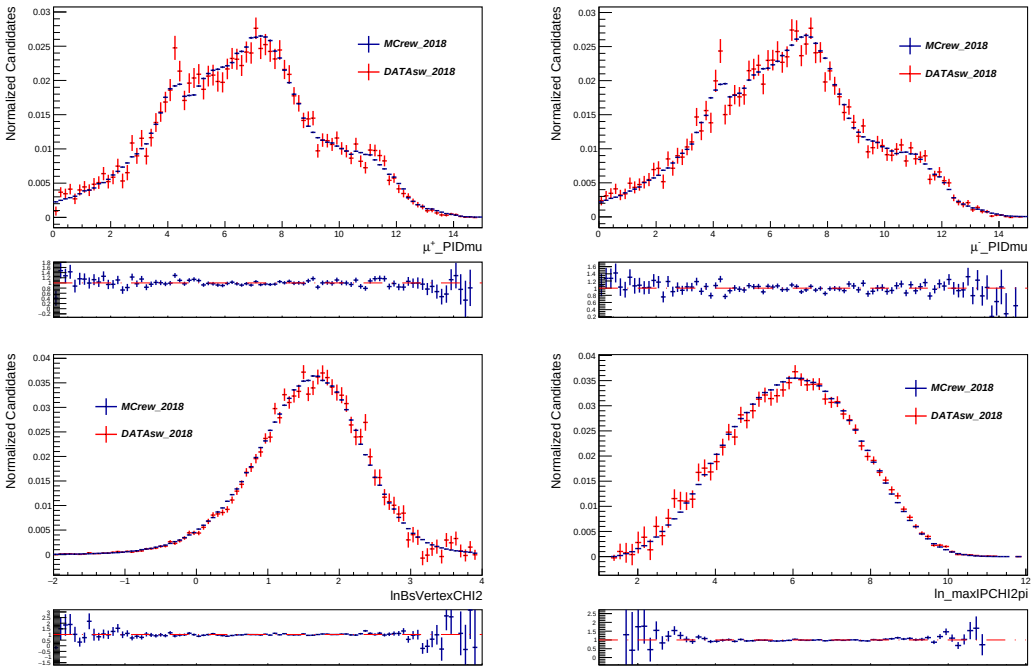

Figure 5.12: Distribution of additional variables used to train the BDT, after the data-simulation correction, for the 2018 dataset.

5.2.4 Multivariate-based selection

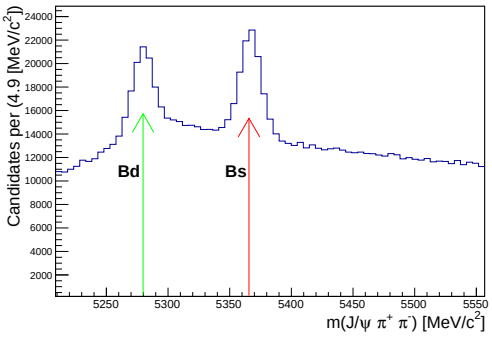

Figure 5.13: $J/\psi\pi^+\pi^-$ invariant mass distribution after the trigger, the offline selection and the rejection of misidentified backgrounds in the 2018 dataset.

The mass distribution of B_s^0 candidates after the trigger, the offline selection and the rejection of misidentified backgrounds is shown in Fig. 5.13 for the 2018 year. Large B_s^0 and B_d^0 contributions over a rather large background are visible. Thus, a multivariate selection is applied to further improve the signal to background ratio. A BDT method is used, called *uBoost* [124], that ensures uniform selection efficiency in a variable of interest. In particular, the BDT is trained in order to develop a uniform selection on the $\cos\theta_\pi$. To show this, the distributions of this variable with and without the BDT selection are shown in Fig. 5.14 using the signal MC. No acceptance effect is introduced by the BDT application. The training variables used in the

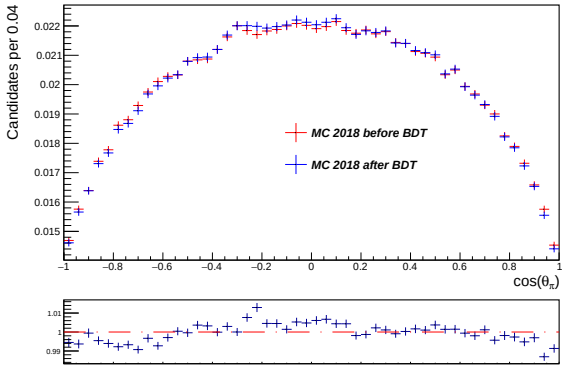

Figure 5.14: $\cos\theta_\pi$ distribution before and after the application of the BDT selection, for the 2018 simulated sample.

BDT are:

- smallest DLL($\mu - \pi$) among the two muons (minPIDmu);

- natural logarithm of the smallest IP χ^2 among the two pions;
- natural logarithm of the largest IP χ^2 among the two pions (ln_maxIPCHI2pi);
- natural logarithm of the vertex fit χ^2 of the B_s^0 (ln_BsVertexCHI2);
- natural logarithm of the IP χ^2 of the B_s^0;
- transverse momentum p_T of the B_s^0;
- sum of the transverse momentum of the pions;
- natural logarithm of the DTF fit χ^2 of the B_s^0.

The uBoost technique involves a training procedure. Corrected truth-matched simulated signal and sideband data with invariant mass inside the interval $[200-250]$ MeV/c^2 above the nominal B_s^0 mass are divided into two independent samples. One is used to train the classifier, while the other is used to test the performance. The BDT training is performed using the 2016 samples also for 2015, due to its low statistics. Different trainings are instead use for the 2017 and the 2018 samples. Signal and background distributions for each of the variables used in the uBoost are shown in Figs. 5.15 and 5.16.

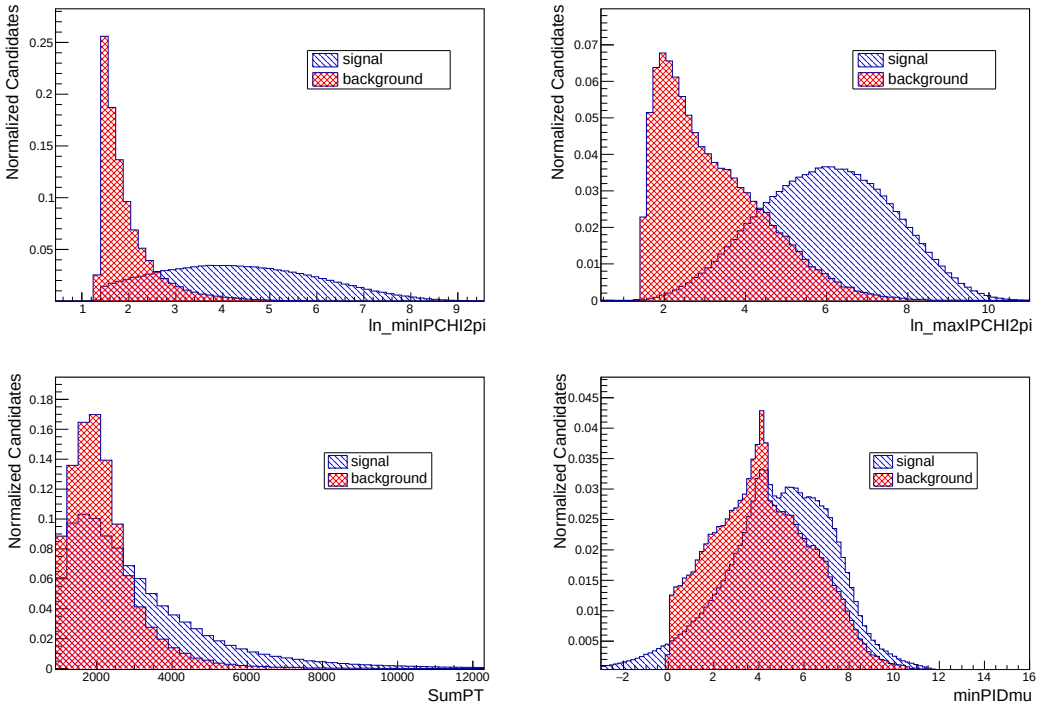

Figure 5.15: Signal and background distributions of final state particle variables used to train the BDT, for the 2018 dataset.

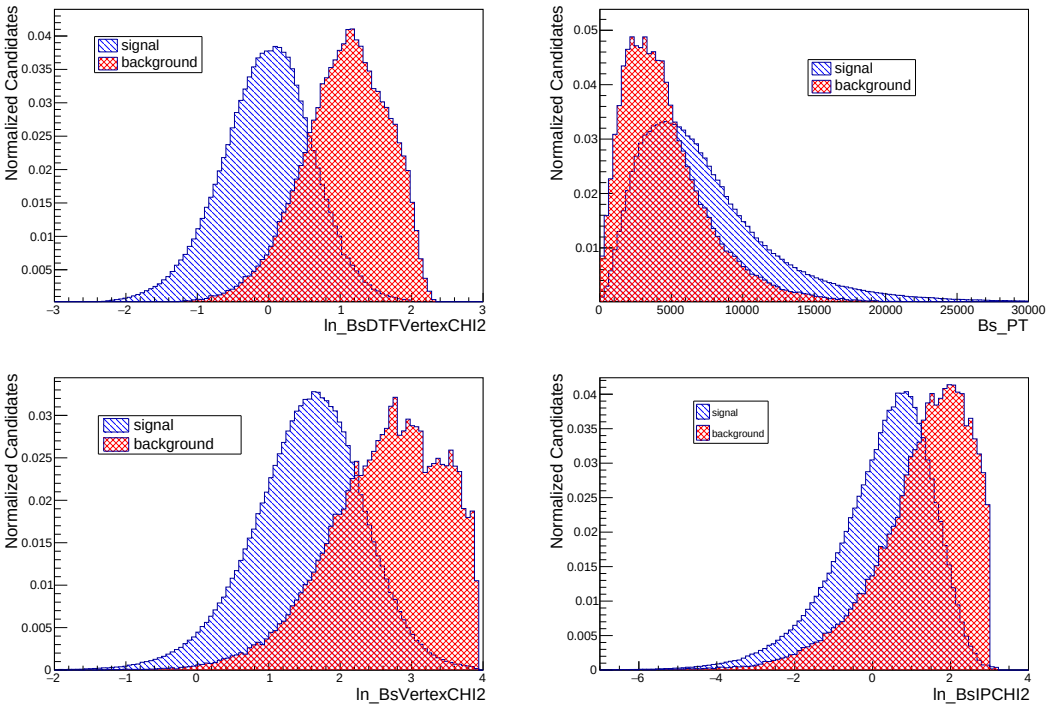

Figure 5.16: Signal and background distributions of B_s^0 variables used to train the BDT, for 2018 dataset.

As visible, there is discrimination power between signal and background in all of these variables. The BDT distributions for signal and background of training and testing subsamples, for the year with higher statistics (2018), are shown in Fig. 5.17.

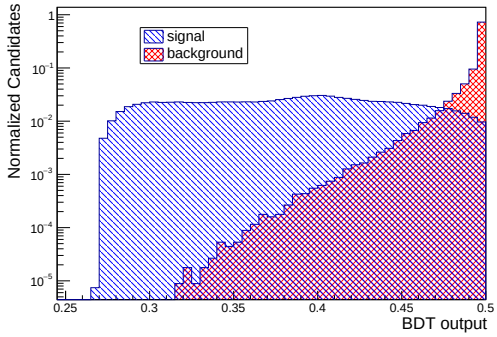

Figure 5.17: BDT output variable of signal and background component, for the 2018 dataset.

Good separation between signal and background component is visible. The cut on the BDT

output is chosen to maximise the following figure of merit (FOM)

$$\text{FOM} = \frac{(\sum_i \omega_i)^2}{\sum_i \omega_i^2}, \tag{5.3}$$

where w_i is the weight assigned to each event by the *sPlot* technique, determined from B_s^0 the invariant mass fit that is performed at each point in the scan over the BDT response. Fig. 5.18 shows the figure of merit as a function of the BDT cut for 2018. A similar distribution is observed for all the years. All the events with a BDT output value larger than 0.46 are selected as signal candidates.

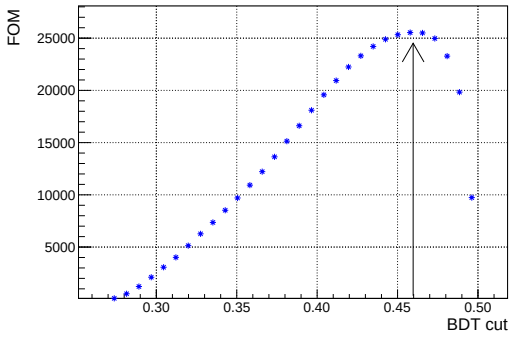

Figure 5.18: Figure of merit as a function of the BDT cut in 2018. The black arrow shows the optimisation point in which the FOM is maximised, equal to 0.46.

5.2.5 Subtraction of remaining backgrounds

After the application of the multivariate selection, the invariant mass below the B_s^0 mass region is still populated by three different sources of background: the $\Lambda_b \to J/\psi K p$, the $B_s^0 \to J/\psi\eta'$ and the combinatorial background. In order to remove them, a strategy similar to that developed for the Λ_b background in Chap. 4 is used.

5.2.5.1 Misidentified backgrounds

The Λ_b and the $J/\psi\eta'$ invariant mass shapes are studied using the related simulated samples. In order to obtain the invariant mass shape when the background is misidentified as signal, the simulated Λ_b events are reconstructed by misidentifying the kaon and the proton as pions. Then, the entire selection procedure applied to the signal data sample is used. Similarly, the $B_s^0 \to J/\psi(\eta' \to \rho\gamma)$ is reconstructed ignoring the γ in the final state and it is then selected as the signal sample. Their resulting invariant mass distributions in the B_s^0 mass region are shown in Fig. 5.19. The two background shapes are described using the `RooKeyspdf` tool, which is able to implement a one-dimensional kernel estimated PDF to model the distribution of an arbitrary input dataset.

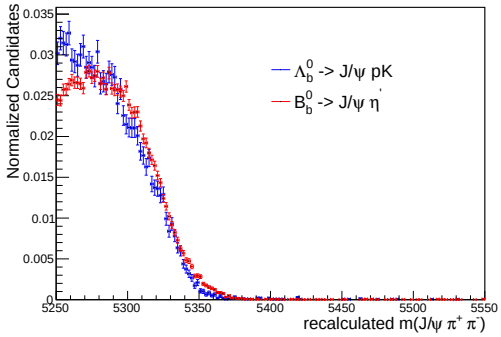

Figure 5.19: Mass distributions of the $\Lambda_b^0 \to J/\psi Kp$ and the $B_s^0 \to J/\psi(\eta' \to \rho\gamma)$ irreducible backgrounds, reconstructed as $m(J/\psi\pi^+\pi^-)$.

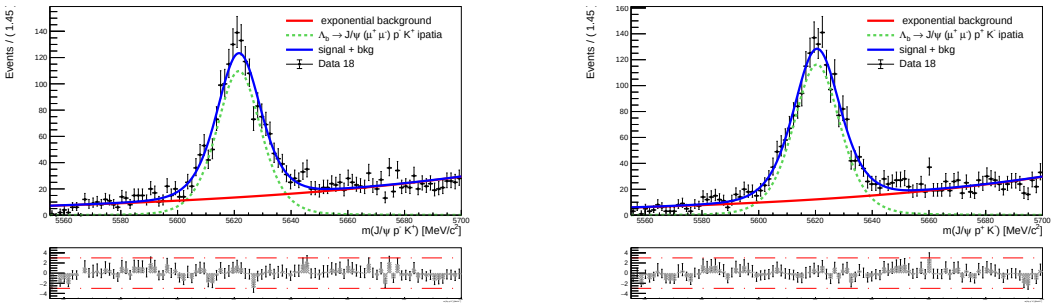

Figure 5.20: Distributions of the $\Lambda_b^0 \to J/\psi Kp$ (left) and $\overline{\Lambda}_b^0 \to J/\psi Kp$ (right) events in B_s^0 2018 signal sample, with fit projections superimposed.

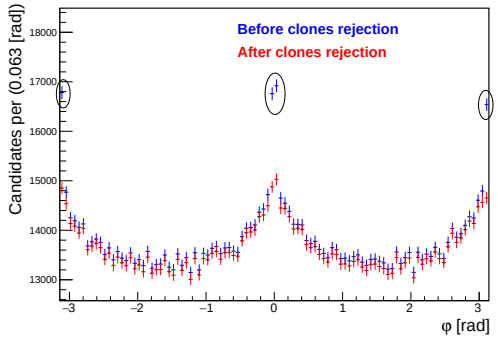

Figure 5.21: Distribution of the φ helicity angle of 2018 WS dataset before and after the clones rejection. The clone contributions, before the application of the veto, are highlighted with a circle.

The two components have a very similar shape. In order to avoid over-estimating one contribution at the expense of the other in the final background subtraction procedure, see Sec. 5.2.5.3, the amount of Λ_b events is fixed to the value obtained as explained in the following:

- the mass distribution of the Λ_b ($\overline{\Lambda}_b$) events present in the real data is obtained misidentifying the positive (negative) pion as a proton and the negative (positive) pion as a kaon. Only the invariant mass region $[5250, 5352]$ MeV/c^2 is considered, in order to suppress most of the real B_s^0 signal candidates. The efficiency of this requirement on the Λ_b^0 events has been obtained from studies on simulated Λ_b^0 samples and it is found to be equal to 1.

- the real events of B_s^0 are then further removed requiring that the `ProbNNk` of the kaon is greater the 0.5. The efficiency of this requirement in the Λ_b^0 events, studied on simulated samples, is found to be equal to 97.5%. This number will be used to correct the fitted yield of Λ_b^0 candidates;

- the mass distribution is then fitted using an Ipatia function for the Λ_b component and an exponential function for the background, where the Ipatia tails are fixed from studies on Λ_b simulated samples.

The resulting distributions of Λ_b and $\overline{\Lambda}_b$ events for 2018 are shown in Fig. 5.20, with the projections of the fit superimposed. The resulting yields for all the years are shown in Tab. 5.3. These yield are then divided for the total effciency of the procedure, namely 97.5%, and the corrected number of Λ_b^0 events are shown in Tab. 5.3.

Table 5.3: The estimated number of Λ_b^0 events in the $B_s^0 \to J/\psi\pi^+\pi^-$ sample for the different years. The uncertainty is the statistical uncertainty.

Year	The number of Λ_b^0	Corrected number of Λ_b^0
2015	458 ± 28	470 ± 29
2016	3028 ± 71	3106 ± 73
2017	2870 ± 71	2943 ± 73
2018	3375 ± 76	3462 ± 78

5.2.5.2 The wrong-sign sample

Since the left sideband of the $J/\psi\pi^+\pi^-$ invariant mass is populated by irreducible backgrounds, the mass shape of the combinatorial events can not be easily determined by fitting the signal invariant mass distribution without introducing correlations with the other background sources. This would lead to a very difficult evaluation of the possible systematic uncertainty due to the modelisation of the invariant mass contributions. For that reason, an independent sample of pure combinatorial background events is used. This is obtained by combining of a J/ψ with two

same-charged pions and it is called $wrong - sign$ (WS) sample. It is selected using the same selection applied to the signal. Due to the different final state, in the WS sample there are four ways to create a clone track. If the pions are positively (negatively) charged, both could be a clone of the positive (negative) muon or vice-versa. Since the definition of the helicity angles remains the same, the clone tracks in the WS sample cause also the presence of an non-physical spike at $\varphi = 0, \pm\pi$, as shown in Fig. 5.21 by the blue distribution. The same rejection strategy used for the signal is applied and the resulting φ distribution is shown in red in Fig. 5.21.

The WS sample is then weighted to match as much as possible the combinatorial background in data. For that, the final distributions used in the fir in the WS sample are compared to the signal one, in the upper mass sideband ($[5420, 5550]$ MeV/c^2). The comparison is shown in Figs. 5.22, 5.23 and 5.24, for the 2018 dataset. A good consistency for all the variables is found,

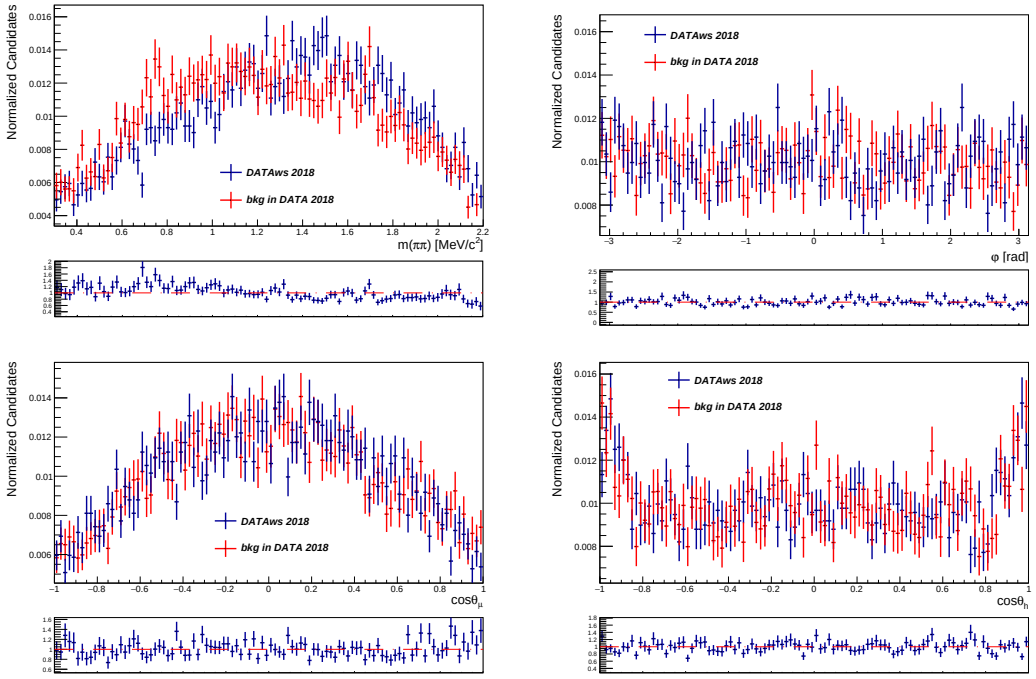

Figure 5.22: Signal and WS distributions of helicity formalism variables before the correction of the WS sample, for the 2018 dataset.

except for the two pions invariant mass, $m(\pi\pi)$, and decay-time, t, distributions. The difference in $m(\pi\pi)$ is due to the fact that the combinatorial background in the signal sample is populated also by events in which the pions come from the decay $\rho \to \pi^+\pi^-$, that cannot contribute to the WS combination. To correct for this, the ratio of the two-dimensional ($m(\pi\pi)$, t) distributions between the signal and the WS events are used to extract a per event weight. These weights are then used to correct the WS events. The comparison between the distributions of the signal and the WS samples is repeated after the reweighting procedure and it is shown in Figs. 5.25, 5.27

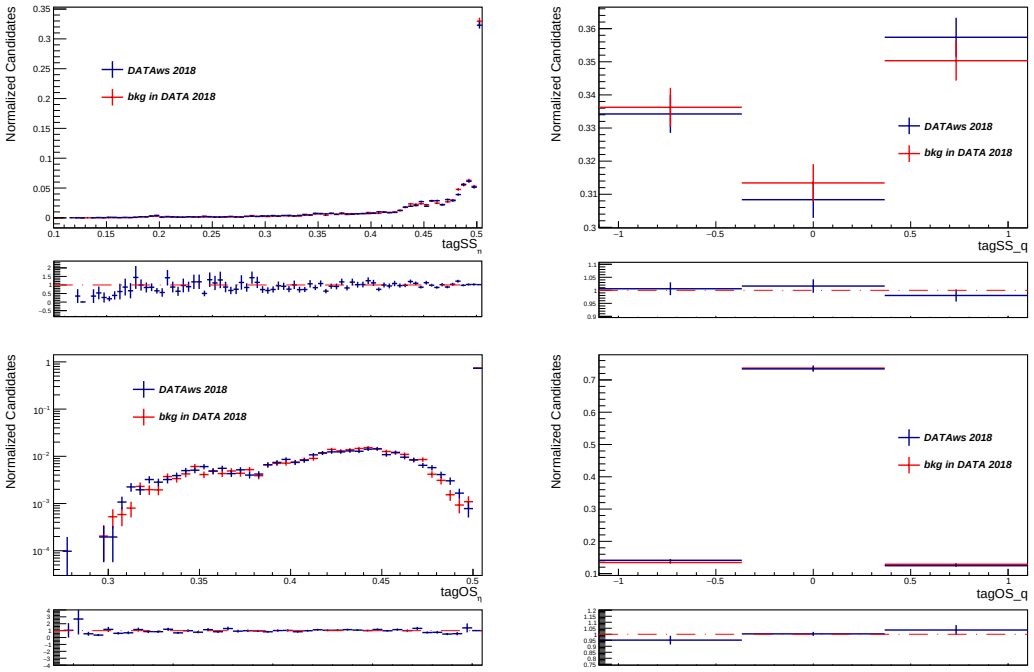

Figure 5.23: Signal and WS distributions of tagging variables before the correction of the WS sample, for 2018 dataset.

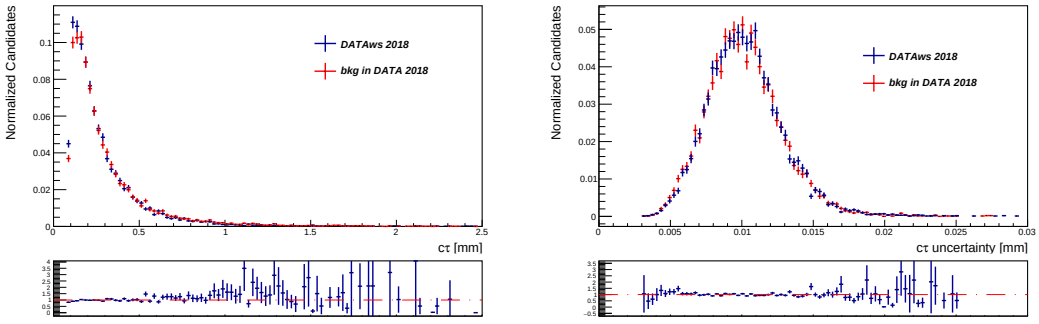

Figure 5.24: Signal and WS distributions of decay-time variables before the correction of the WS sample, for 2018 dataset.

and 5.26, for the 2018 dataset. A good agreement in all the variables is visible, allowing the usage of the WS sample as a proxy for the combinatorial background.

5.2.5.3 Background subtraction

A simultaneous fit of the invariant mass of the signal sample and the corrected WS sample, in the region $[5250, 5550]$ MeV/c^2 is performed. The signal component in the $J/\psi\pi^+\pi^-$ data

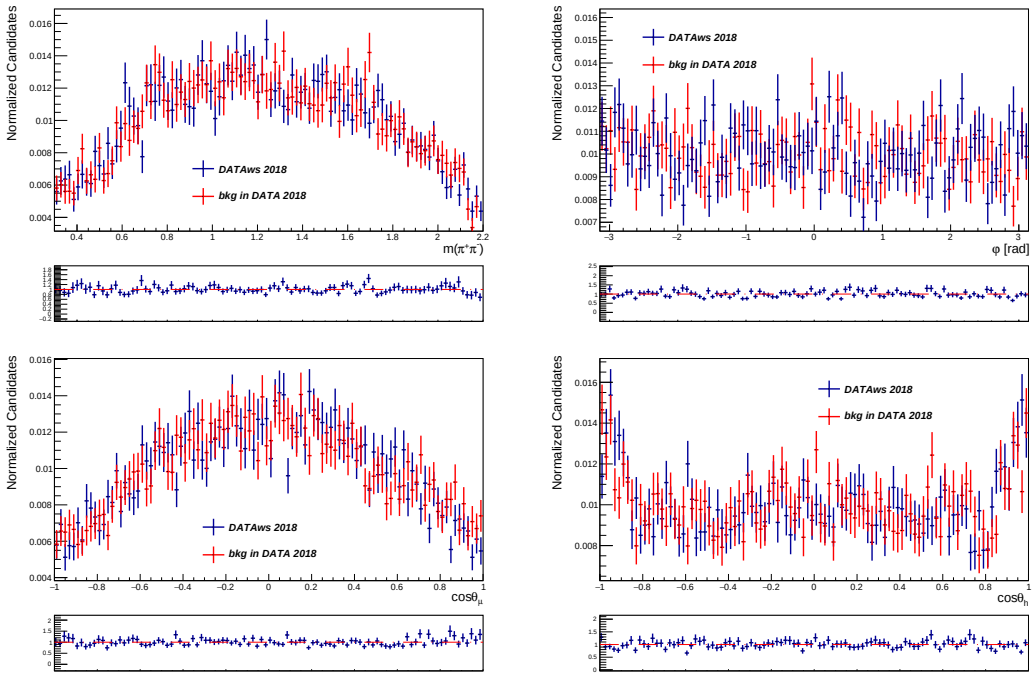

Figure 5.25: Signal and WS distributions of helicity formalism variables after the correction of the WS sample, for the 2018 dataset.

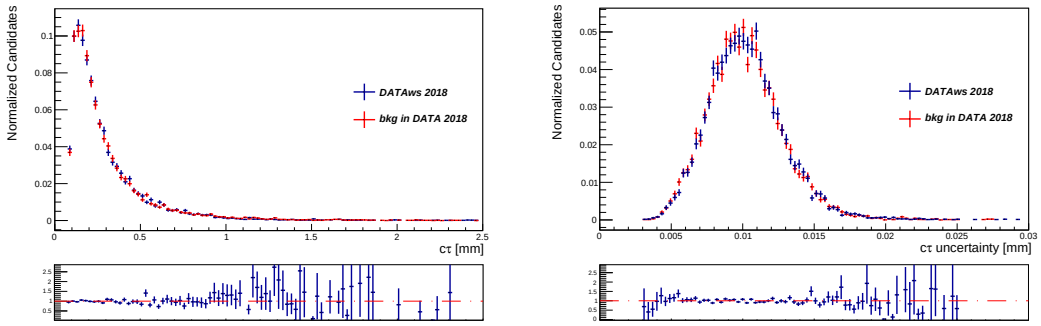

Figure 5.26: Signal and WS distributions of decay-time related variables after the correction of the WS sample, for the 2018 dataset.

distribution is modelled with an Ipatia function, while an exponential function is used for the combinatorial background. The two misidentified background contributions are fitted using the shapes obtained with the `RooKeyspdf` tool, with the Λ_b yield fixed to the corrected value obtained in Sec. 5.2.5.1. The tails of the Ipatia are fixed to the values obtained fitting the truth-matched and corrected simulation sample. Considering the mass region used for the fit, also the $B_d^0 \to J/\psi\pi^+\pi^-$ component has to be taken into account. It is fitted with an

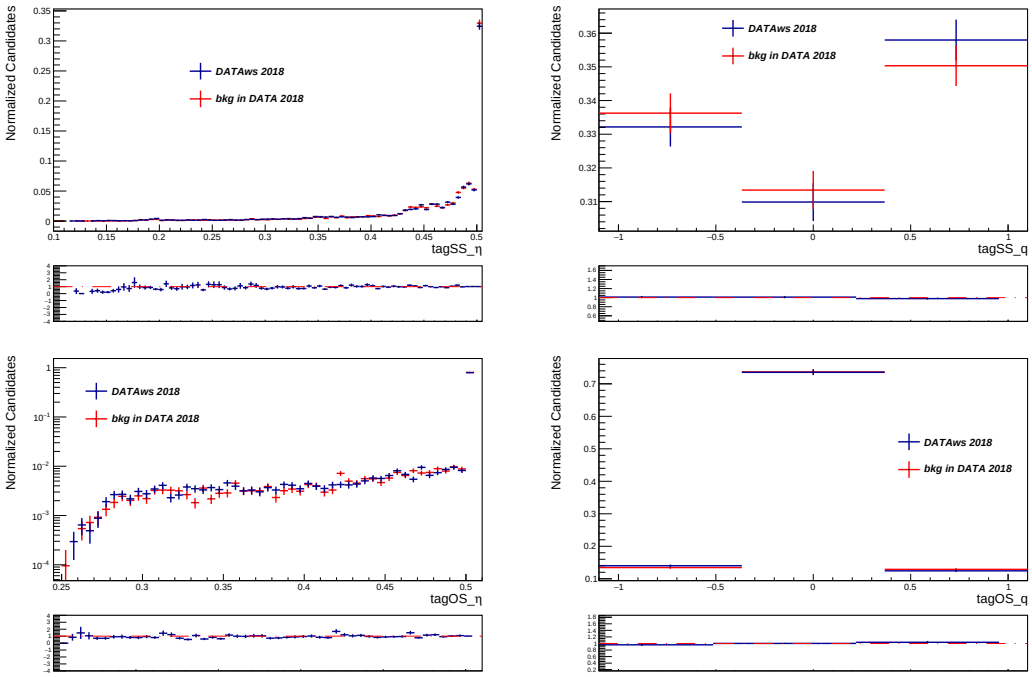

Figure 5.27: Signal and WS distributions of tagging variables after the correction of the WS sample, for the 2018 dataset.

Ipatia whose parameters are the same as the signal ones, except for the peak position, that is considered to be the B_s^0 peak position minus the nominal difference between the two meson masses $\Delta m(B_s^0 - B_d^0) \sim 87$ MeV/c^2 [13]. This can be done considering that the two decays have a very similar topology, resulting in a very similar mass resolution effect and, consequently, in a very similar mass shape. The fit of the combinatorial background in the $J/\psi\pi^+\pi^-$ sample is constrained by the WS sample, that is simultaneously fitted with an exponential function with the same parameters, expect for the normalisation coefficient. The results of the fit for all the years are shown in Fig. 5.28, at left for the signal sample and at right for the corrected WS sample. In Tab. 5.4, the resulting yields of the different components in the signal samples are shown, as a function of the year. A large amount of background is still present.

Table 5.4: The fitted yields of the different components that populate the signal dataset in the mass region $[5250, 5550]$ MeV/c^2, for the different years.

Year	$B_s^0 \to J/\psi\pi^+\pi^-$	$B_d^0 \to J/\psi\pi^+\pi^-$	$B_s^0 \to J/\psi\eta'$	$B_s^0 \to J/\psi\pi^\pm\pi^\pm$
2015	4764 ± 83	3094 ± 93	581 ± 115	4607 ± 117
2016	31069 ± 209	19652 ± 226	3994 ± 283	28948 ± 292
2017	30579 ± 202	19804 ± 218	4212 ± 260	22302 ± 261
2018	37949 ± 229	24448 ± 243	4939 ± 301	33508 ± 316

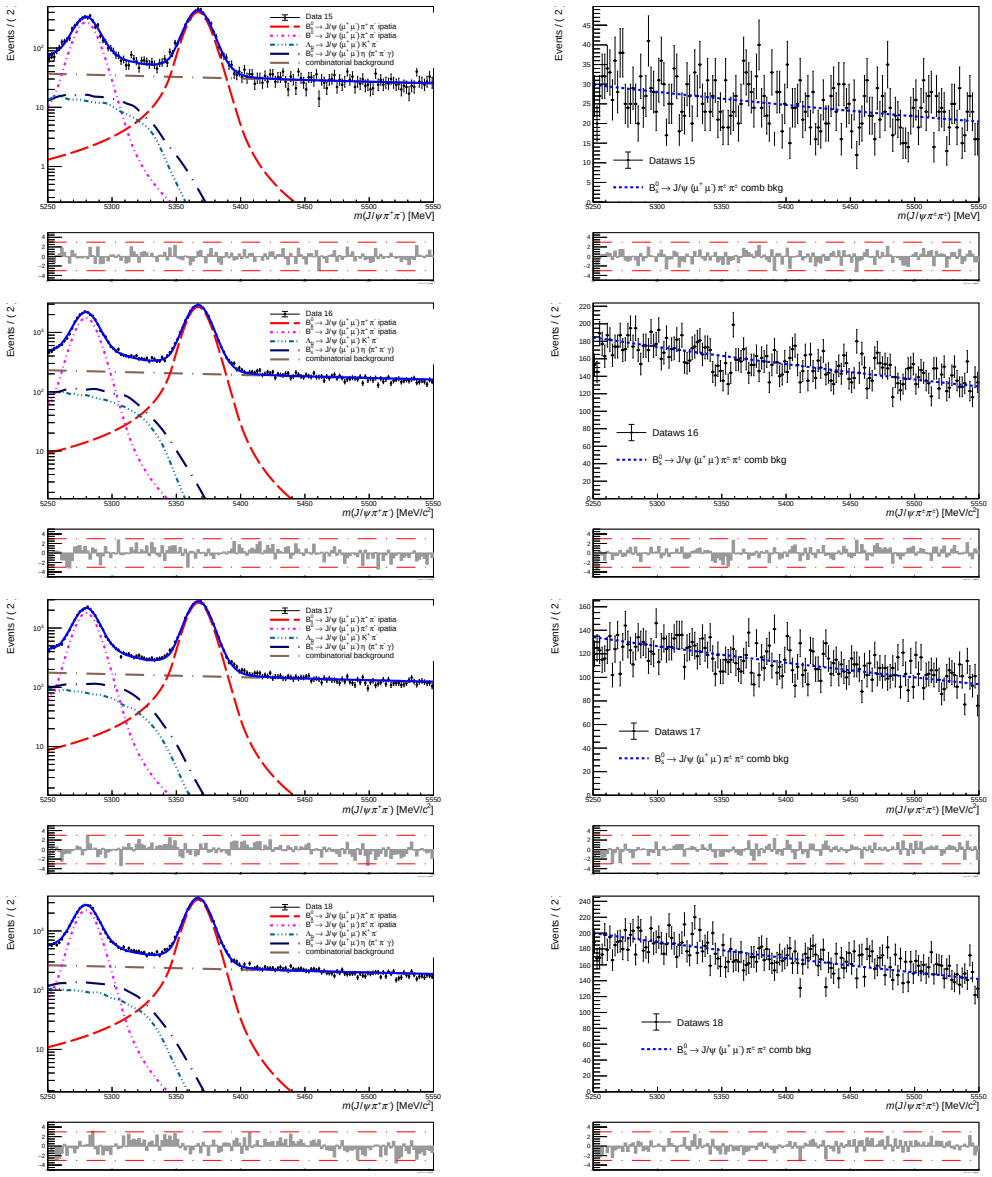

Figure 5.28: Simultaneous fit of the invariant mass of the $B_s^0 \to J/\psi\pi^+i^-$ signal sample (left) and corrected WS sample (right) for each year, with the fit projections superimposed.

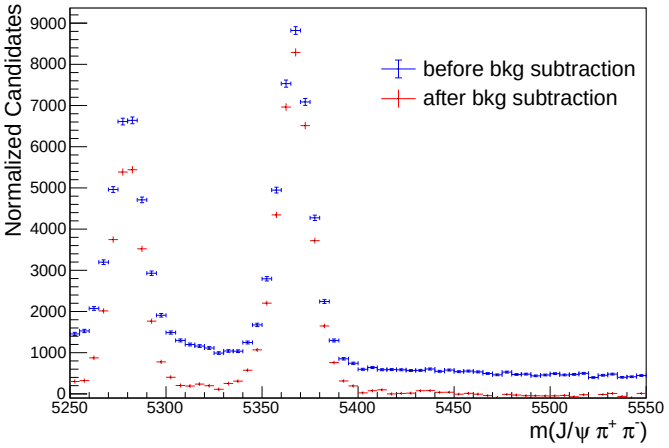

Figure 5.29: Invariant mass distribution of the $B_s^0 \to J/\psi\pi^+\pi^-$ signal sample before and after the background subtraction, for the 2018 dataset.

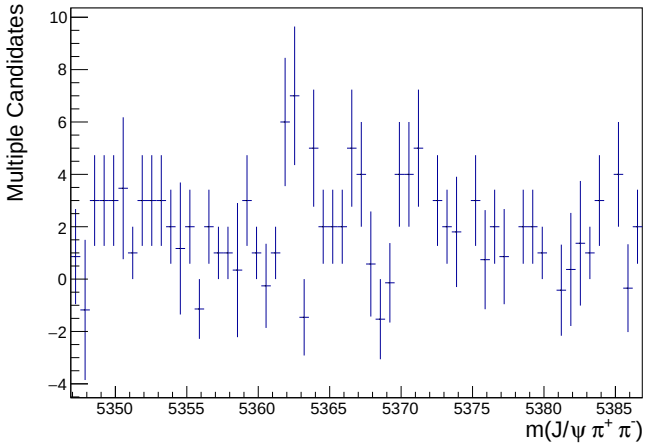

Figure 5.30: Invariant mass distribution of the $B_s^0 \to J/\psi\pi^+\pi^-$ signal candidates of events containing more than one candidate, for the 2018 dataset.

Therefore, in order to take properly into account all the possible correlations between the mass and the observables used in the final fit and to have a good control of the background related systematic uncertainty, the subtraction of the background is performed in a different way with respect to the ϕ_s measurement made with the $B_s^0 \to J/\psi K^+K^-$ decay, presented in the previous chapter. Instead of using the *sPlot* technique, the WS sample and the misidentified MC samples of Λ_b and $B_s^0 \to J/\psi\eta'$ are weighted to have a number of events equal to minus the values obtained from the simultaneous fit. Then, these samples are injected in the signal

data sample, in order to statistically subtract the background components. Fig. 5.29 shows the $J/\psi\pi^+\pi^-$ invariant mass distribution for the 2018 year, before (blue) and after (red) the background subtraction. The B_d^0 background is removed by performing the final measurement in the mass region of ± 20 MeV/c^2 around the B_s^0 nominal mass value.
At the end of the background subtraction procedure, even though the chance to produce two $B_s^0 \to J/\psi\pi^+\pi^-$ signal candidates within one pp collision is extremely small, the 0.3% of the events contains more than one candidate and their mass distribution in 2018 is shown in Fig. 5.30. It is possible to see that these multiple candidates are distributed like combinatorial background events, so they are mostly due to fake B_s^0 events, which have been created combined the J/ψ coming from the b-meson with at least one random pion. Since their contribution is really small, it is considered negligible for the scope of this measurement.

5.3 Selection of the $B_d^0 \to J/\psi K^+\pi^-$ control mode

The $B_d^0 \to J/\psi K^*(892)^0$ sample, with the $K^*(892)^0 \to K^+\pi^-$, will be used in the analysis as a control mode for the determination of the decay-time acceptance, following the same approach already discussed in Chap. 4. In order to be used in the acceptance determination as a proxy of the signal decay, a very similar selection strategy is used.

5.3.1 Trigger and offline selection

The same trigger selection is applied. Then, the stripping line `StrippingB2JpsiHHBs2JpsiKstarLine` is used in the offline selection. It applies almost the same requirements of the signal stripping line and the details are summarised in Tab. 5.5. The differences with the signal stripping line are highlighted in bold and are due to the different particles involved in the decay with respect to the signal one.

Table 5.5: `StrippingB2JpsiHHBs2JpsiKstarLine` selection requirements, which are the same for all the 2015-2018 datsets.

#	Selection variables	Requirements
1	All tracks χ^2/ndf	< 5
2	All tracks GHOSTPROB	< 0.5
3	muon DLL($\mu - \pi$)	> 0
4	p_{T} of muon	> 500 MeV/c
5	J/ψ vertex χ^2_{vex}	< 16
6	J/ψ DOCA χ^2	< 20
7	J/ψ mass window	± 80 MeV/c
8	p_{T} of pion	> 250 MeV/c
9	PID of pion	DLL($\pi - K$) > -10 or `ProbNNpi`> 0.05
10	PID of kaon	**PIDK > 0** and **(PIDK − PIDp) > −10** or `ProbNNk`**> 0.05**
11	p_{T} of kaon	> 250 MeV/c
12	Sum p_{T} of $K^+\pi^-$	> 900 MeV/c
13	$K^+\pi^-$ vertex χ^2	< 16
14	$K^+\pi^-$ DOCA χ^2	< 20
15	B_d^0 Vertex χ^2_{vex}/ndf	< 10
16	B_d^0 χ^2_{IP}	< 25
17	B_d^0 Pointing, DIRA $\cos\theta_p$	> 0.999
18	B_d^0 decay time	> 0.2 ps
19	B_d^0 mass window	**[5000, 5550] MeV/c²**

The same requirements summarised in Tab. 5.2 are then applied. In addition the kaon and pion must have the minimum $\chi^2_{\mathrm{IP}} > 4$, in order to harmonise the selection with the signal one, in which this cut is present at the stripping level. The pion is required to have `PIDK`< 0, in order to reduce the misidentified background contributions, mostly due to the misidentification of kaons as pions.

Also in this case, simulated samples of the $B_d^0 \to J/\psi K^+\pi^-$ decays are produced and selected using exactly the same selection of data. Since these samples are strongly used in the development of the analysis, they are corrected exploiting the same strategy explained for the signal channel. The agreement between the truth-matched simulated samples and the sWeighted data in the variables used during the selection procedure is shown in Figs. 5.33, 5.31, 5.32, 5.34.

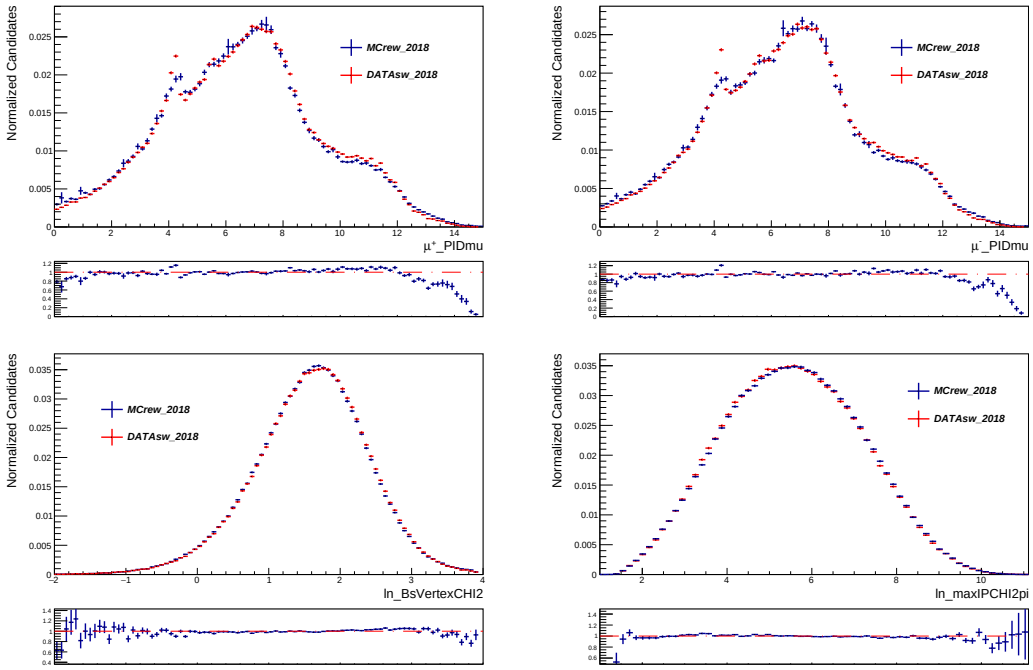

Figure 5.31: Distributions of additional variables used to train the BDT, after the data-simulation correction for the 2018 dataset.

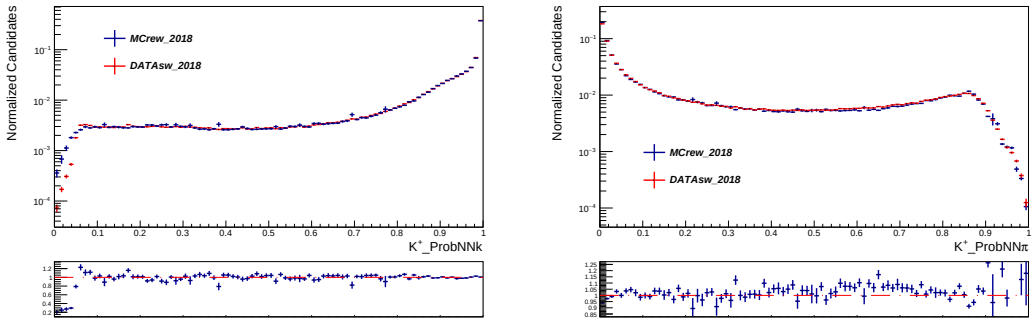

Figure 5.32: Distributions of kaon PID variables used to veto misidentified backgrounds, after the data-simulation correction for the 2018 dataset.

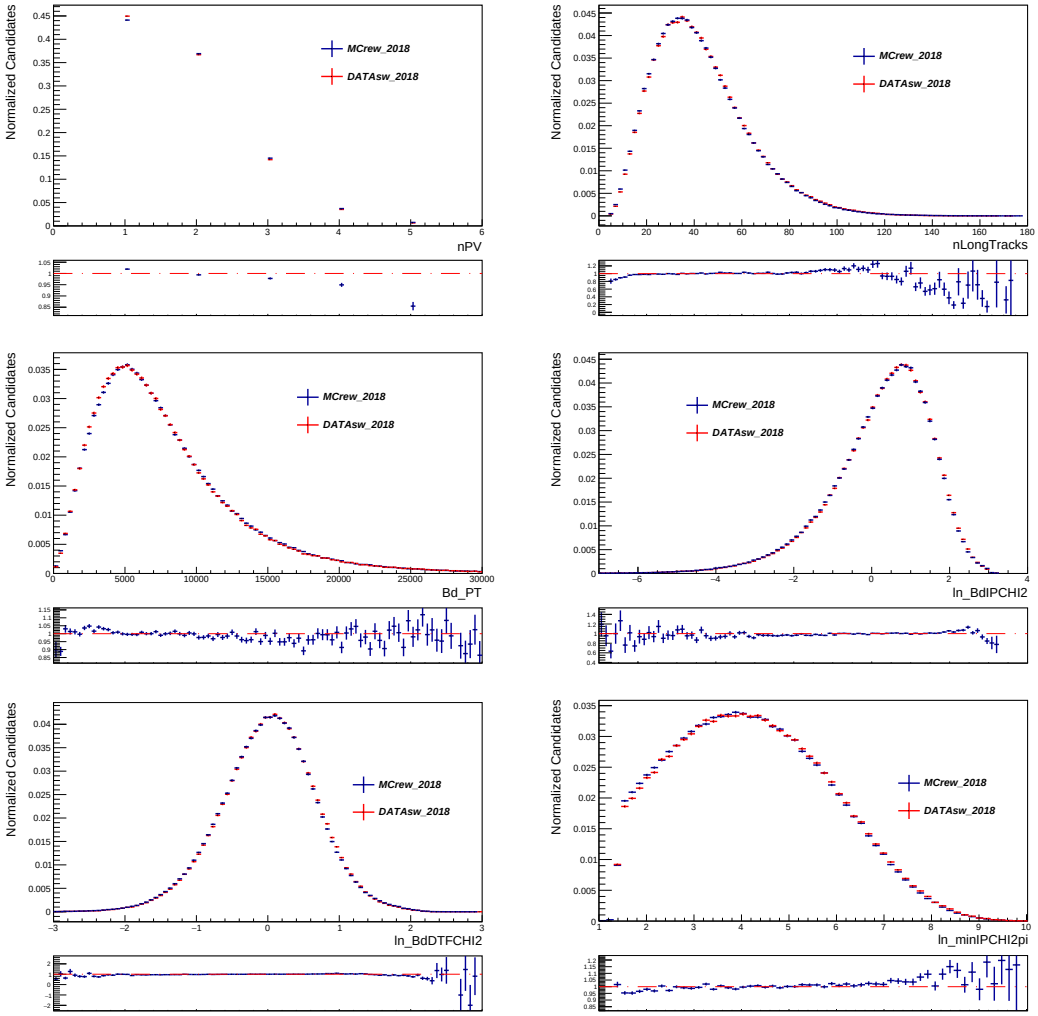

Figure 5.33: Distribution of weighting variables, after the data-simulation correction for the 2018 dataset.

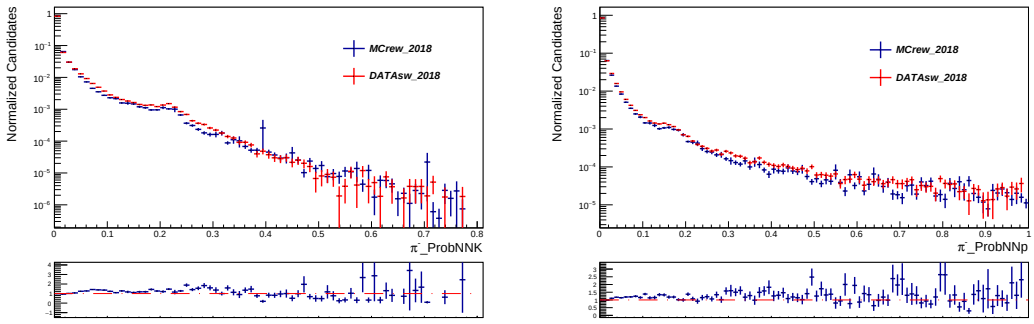

Figure 5.34: Distribution of pion PID variables used to veto misidentified backgrounds, after the data-simulation correction for the 2018 dataset.

5.3.2 Misidentified backgrounds

In this section, the presence of potential peaking backgrounds arising from the misidentification of final-state particles coming from other decays is investigated.

5.3.2.1 The $\Lambda^0_b \to J/\psi K^+ p^-$ background

Background events due to of $\Lambda^0_b \to J/\psi K^+ p^-$ decays can be reconstructed as a B^0_d candidates if the proton is misidentified as a π^-. Figure 5.35 shows the recomputed $J/\psi Kp$ mass distribution obtained from the B^0_d data sample, when interpreting the p as a π, and a peaking structure around the nominal Λ^0_b mass is visible. These events are vetoed removing all candidates with the pion `ProbNNp` > 0.2 and $|m(J/\psi Kp) - m(\Lambda^0_b)| < 3\sigma$, where the resolution σ is obtained from the fit of the vetoed distribution.

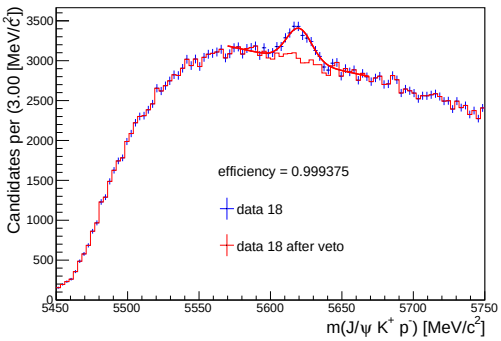

Figure 5.35: Distribution of the recomputed $m(J/\psi Kp)$ before, in blue, and after, in red, the application of the veto in 2018.

5.3.2.2 The $\bar{B}_d^0 \to J/\psi K^-\pi^+$ background

Doubly-misidentified background events due to $\bar{B}_d^0 \to J/\psi K^-\pi^+$ decays can be reconstructed as a B_d^0 candidate if the positive pion is misidentified as a K^+ and the negative kaon as a π^-. Figure 5.36 shows the recomputed $J/\psi K^-\pi^+$ mass distribution obtained from the B_d^0 data sample, when interpreting the kaon as a π and vice-versa, and no peaking structure around the nominal $\bar{B}_d^0$ mass is visible. Anyway, in order to suppress any possible $\bar{B}_d^0$ events, all candidates with $\texttt{ProbNNk}(\pi) > \texttt{ProbNNk}(K)$ and $\texttt{ProbNNk}(\pi) > 0.1$, $\texttt{ProbNNpi}(K) > 0.1$ and $|m(J/\psi K^-\pi^+) - m(\bar{B}_d^0)| < 3\sigma$ are rejected. The σ is obtained from the fit of the vetoed distribution, obtained without the requirement of PID$k < 0$ for the pion.

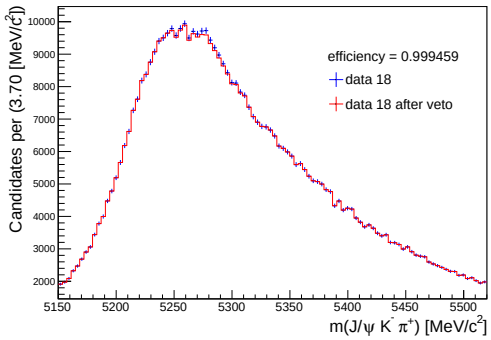

Figure 5.36: Distribution of the recomputed $m(J/\psi K^-\pi^+)$ before, in blue, and after, in red, the application of the veto in 2018.

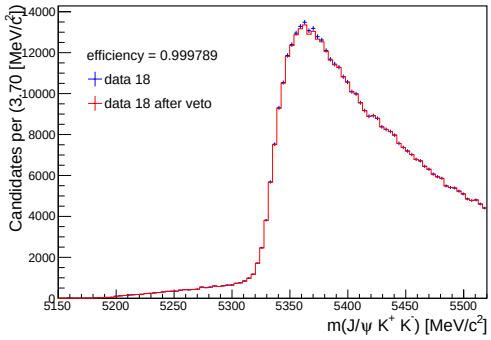

Figure 5.37: Distribution of the recomputed $m(J/\psi K^-K^+)$ before, in blue, and after, in red, the application of the veto in 2018.

5.3.2.3 The $B_s^0 \to J/\psi\phi$ background

Background events due to $B_s^0 \to J/\psi\phi$ decays, with the $\phi \to K^+K^-$, can be reconstructed as a B_d^0 candidate if the negative pion is misidentified as a K^-. Figure 5.37 shows the recomputed

$J/\psi K^- K^+$ mass distribution obtained from the B_d^0 data sample, when interpreting the kaon as a π, and no peaking structure around the nominal B_s^0 mass is visible. Anyway, possible events due to this background are vetoed removing all candidates with `ProbNNk`$(\pi) > 0.1$ and $|m(K^-K^+) - m(\phi)| < 3\sigma$, where σ is the nominal width of the ϕ resonance [13].

5.3.3 Mass fit and background subtraction

After the preselection and the subtraction of the peaking backgrounds, the same BDTs trained for the B_s^0 decay are applied to the B_d^0 samples, for the different years. Figure 5.38 shows, for each year, the invariant mass distribution of the selected candidates, which is fitted with an Ipatia function to model the B_d^0 signal contribution and an exponential function for the combinatorial background. The parameters of the tails of the signal function are fixed from studies on simulated samples. The total number of signal events for each year is found to be 64390 ± 263, 405274 ± 657, 412136 ± 663 and 500466 ± 730, respectively for 2015, 2016, 2017 and 2018 . The fit is then used to generate event-by-event weights with the *sPlot* technique to subtract the remaining combinatorial background. The obtained background-subtracted samples will be used to extrapolate the decay-time acceptance. At the end of the background subtraction,

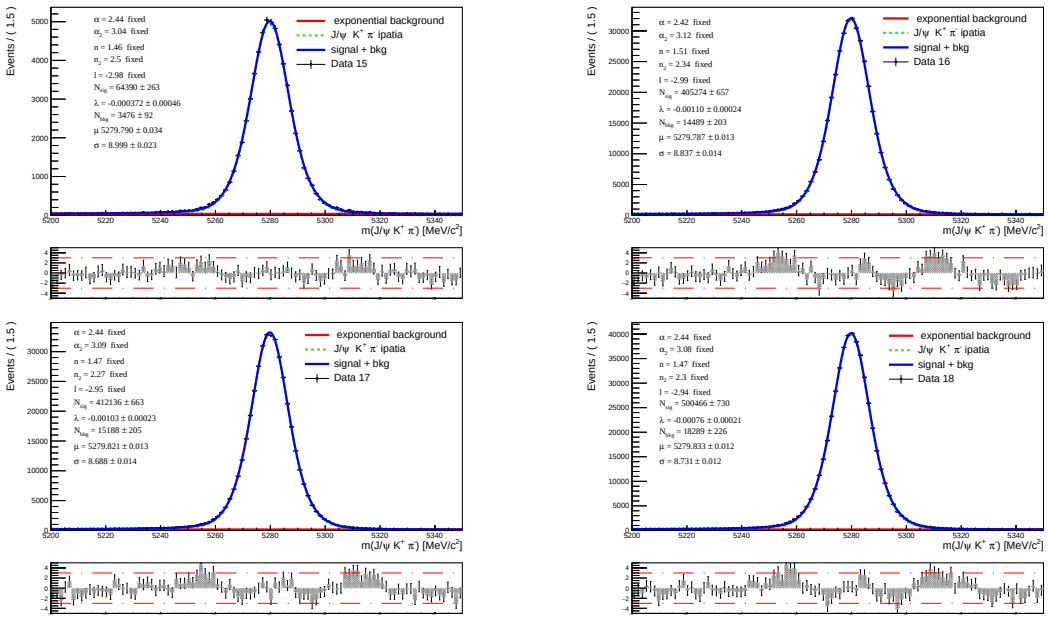

Figure 5.38: Fit of the invariant mass distribution of $B_d^0 \to J/\psi K^+\pi^-$ in 2015, 2016, 2017 and 2018.

the 0.9% of the events contain more than one candidate and their mass distribution in 2018 is shown in Fig. 5.39. These multiple candidates show peaking structure at the B_d^0 mass value, thus they are generally created combining the J/ψ coming from the b-meson decay with cloned tracks of the pion and/or of the kaon really coming from the B_d^0. As for the B_s^0 signal decay,

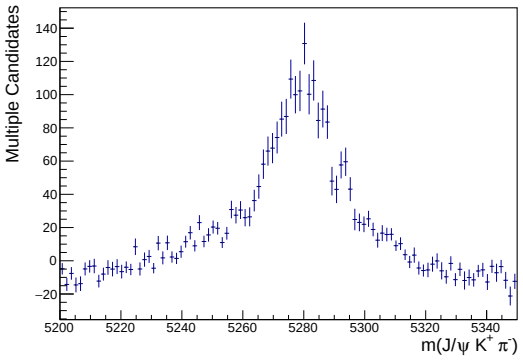

Figure 5.39: Invariant mass distribution of $B_d^0 \to J/\psi K^+\pi^-$ signal candidates of events containing more than one candidate, in 2018 dataset.

since the multiple candidates contribution is really small, it is considered negligible and its input will be evaluated as a systematic contribution.

5.4 Experimental effects

In this section the determination of the decay-time resolution is presented. The same strategy discussed in Sec. 4.5 is used. The other experimental effects, namely the determination of the angular and decay-time acceptances and the calibration of the flavour tagging algorithm are currently being developed by another LHCb group and will thus not be included in this thesis.

5.4.1 Decay-time resolution

A per-event decay-time resolution has been determined, calibrating the per-event DTF decay-time uncertainty σ_t. A data-driven method is used, involving a sample of fake $B_s^0 \to J/\psi\pi^+\pi^-$ prompt events, where the J/ψ and the pions are coming from the PV. This sample is expected to have decay time equal to zero, so the spread of its decay-time distribution can be used to evaluate the decay-time resolution effect. The calibration of σ_t is done studying the resolution effect in ten bins of the DTF decay-time uncertainty. The binning scheme has been chosen in order to have an almost equal number of events in each bin. The optimisation of the binning scheme is done independently for each year of data taking and the result is reported in Tab. 5.6.

Table 5.6: Binning scheme used for the DTF decay-time uncertainty, σ_t, in the calibration of the decay time resolution of for the different years.

# bin	2015 [ps]	2016 [ps]	2017 [ps]	2018 [ps]
1	(0 - 0.0258)	(0 - 0.0257)	(0 - 0.0263)	(0 - 0.0265)
2	(0.0258 - 0.0285)	(0.0257 - 0.0284)	(0.0263 - 0.0289)	(0.0265 - 0.0291)
3	(0.0285 - 0.0305)	(0.0284 - 0.0304)	(0.0289 - 0.0309)	(0.0291 - 0.0311)
4	(0.0305 - 0.0323)	(0.0304 - 0.0322)	(0.0309 - 0.0326)	(0.0311 - 0.0328)
5	(0.0323 - 0.0341)	(0.0322 - 0.0340)	(0.0326 - 0.0343)	(0.0328 - 0.0345)
6	(0.0341 - 0.0359)	(0.0340 - 0.0358)	(0.0343 - 0.0360)	(0.0345 - 0.0362)
7	(0.0359 - 0.0379)	(0.0358 - 0.0379)	(0.0360 - 0.0379)	(0.0362 - 0.0381)
8	(0.0379 - 0.0404)	(0.0379 - 0.0405)	(0.0379 - 0.0402)	(0.0381 - 0.0404)
9	(0.0404 - 0.0442)	(0.0405 - 0.0442)	(0.0402 - 0.0436)	(0.0404 - 0.0439)
10	(0.0442 - 0.15)	(0.0404 - 0.15)	(0.0436 - 0.15)	(0.0439 - 0.15)

In order to keep the resolution effect in the prompt sample as similar as possible to that present in the signal sample, a very similar selection procedure is applied. The requirements that can create a distortion of the decay-time distribution are not used: namely #10, #15, #16 and #17 of Tab. 5.1, #9 in Tab. 5.2 and the BDT selection. In addition, the mass of the prompt candidate is required to be in a range of 20 MeV/c^2 around the B_s^0 nominal mass.

In order to evaluate the combinatorial background component, due to muons not really coming from a J/ψ, the invariant mass of the J/ψ is studied and its distribution is shown in Fig. 5.40 for the 2018 year. A large amount of combinatorial background events is present. In order to

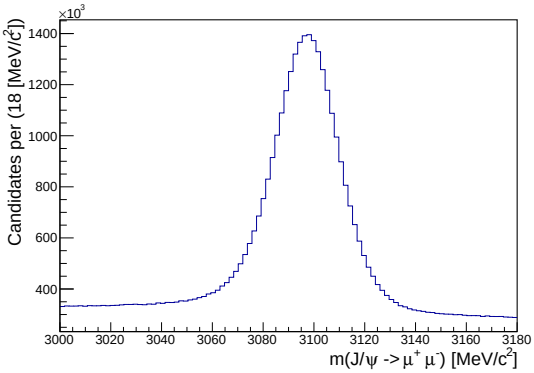

Figure 5.40: Invariant mass distribution of prompt $J\psi \to \mu^+\mu^-$ in 2018.

remove it, the *sPlot* technique is used, by fitting the J/ψ invariant mass to generate event-by-event sWeights to statistically subtract the combinatorial background component. Since the measurement of the decay-time resolution is done in ten bins of σ_t, the fits are also performed in bins. Two single-sided Crystal Ball functions that share the same mean value but have different widths are used to describe the signal, while the combinatorial background is described with an exponential function. The tails of the Crystal Ball functions are fixed from studies performed in truth matched prompt MC samples[†]. The fits to the dimuon invariant mass for 2018 in the 1$^{\text{st}}$, 4$^{\text{th}}$, 7$^{\text{th}}$ and 10$^{\text{th}}$ bin are shown as an example in Fig. 5.41.

In order to properly apply the calibration function obtained in the prompt decay to the signal decay, the differences between signal and prompt samples need to be corrected. They are generated from the differences in the selection procedure and from the different phase space. The sWeighted prompt events are thus weighted to match the background subtracted signal distributions, corresponding to the number of long tracks and the DTF χ^2 of the B_s^0. The validity of this weighting procedure is discussed in Sec. 5.4.1.5. The comparison between signal and prompt variables, before and after the reweighting, are shown in Fig. 5.42 for the 2018 year.

5.4.1.1 Measurement of the decay time resolution

The decay time resolution is extracted by evaluating the dilution D due to the events with decay-time value $t < 0$, following the same strategy discussed in Sec. 4.5.1. The dilution, in each bin of σ_t, is then numerically computed as

$$D = \frac{1}{N}\sum_{i=1}^{N}\cos(\Delta m_s(t - t_0)) \quad \text{if } t < 0, \tag{5.4}$$

[†](`BKGCAT`(J/ψ) ≤ 10 or `BKGCAT` $== 50$ or `BKGCAT`($B_s^0 \leq 10$ or `BKGCAT`(B_s^0) $== 50$). The last requirements on the B_s^0 `BKGCAT` account for the unavoidable presence of a long-lived component in the data prompt sample.

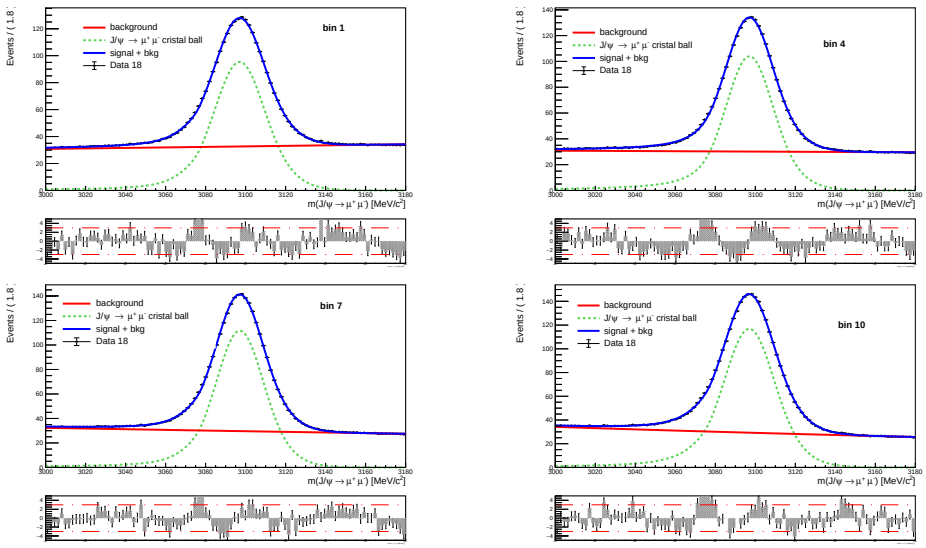

Figure 5.41: Dimuon invariant mass for 2018 in the 1st, 4st, 7th and 10th bin, with fit projections superimposed.

where $\Delta m_s = 17.74$ ps^{-1} and t_0 is the mean of the resolution model obtained from the fit explained in Sec. 5.4.1.3. In this way, it accounts for the presence of a decay-time bias due to the VELO misalignment. In order to take into account the small contribution of long-lived candidates present at negative decay-time values, the long-lived component shape is extracted from the fit described in Sec. 5.4.1.3. Then the dilution due only to the real prompt events, D_{corr}, is calculated as

$$D_{corr} = \frac{1}{1 - I_{ll}} D - \frac{I_{ll}}{1 - I_{ll}} D_{ll} \tag{5.5}$$

where D is the original computed dilution of the sample, I_{ll} is the fraction of the long-lived component in the negative decay-time range and D_{ll} is its contribution to D. For the estimation of D_{ll}, a sample of long-lived candidates is randomly generated using the shape obtained by the fit. As an example, the obtained long-lived shape for the fifth bin in 2018 is shown in Fig. 5.43. The generation is repeated 1000 times for each σ_t bin, leaving the number of long-lived events floating, following a Poissonian distribution with mean equal to the number of long-lived events extracted by the fit of the decay-time distribution. In this way, the error associated to the fitted PDF, due to the finite size of the prompt sample, is propagated to the D_{ll} estimation and, therefore, to D_{corr}. The long-lived component represents only about $2-5\%$ of the negative side of the decay-time distribution in the different years and bins.
In each bin of σ_t, the dilution can be translated into the effective single Gaussian resolution width σ_{eff} as

$$\sigma_{eff} = \sqrt{-2/\Delta m_s^2 \, \ln D}. \tag{5.6}$$

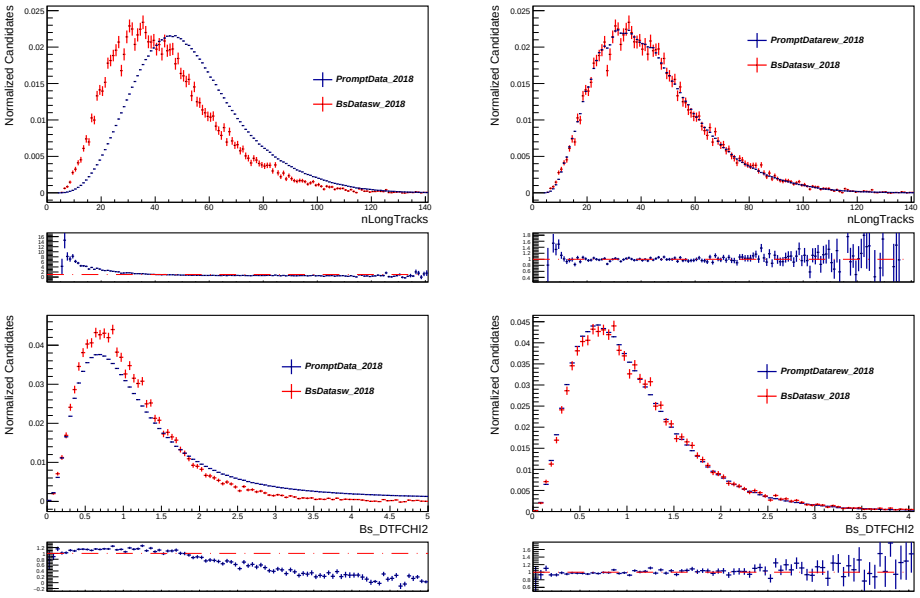

Figure 5.42: The comparison between signal and prompt variables, before (left) and after (right) the reweighting procedure in the 2018 year.

Then, the calibration of σ_t can be performed.

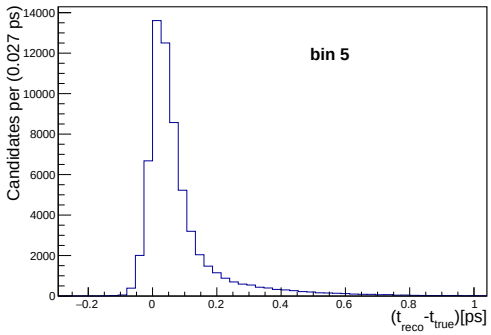

Figure 5.43: The long-lived shape for the fifth bin in 2018, obtained from the fit described in Sec. 5.4.1.3.

5.4.1.2 Calibration of the DTF decay-time uncertainty

The trend of effective signal Gaussian resolution width, as a function of σ_t, is fitted in order to extrapolate the calibration function. A linear function with coefficients p_0 and p_1 is used

$$PDF(\sigma_t) = p_0 + p_1\sigma_t. \tag{5.7}$$

The obtained calibration function for all the years, together with the resulted parameter values, is presented in Fig. 5.44.

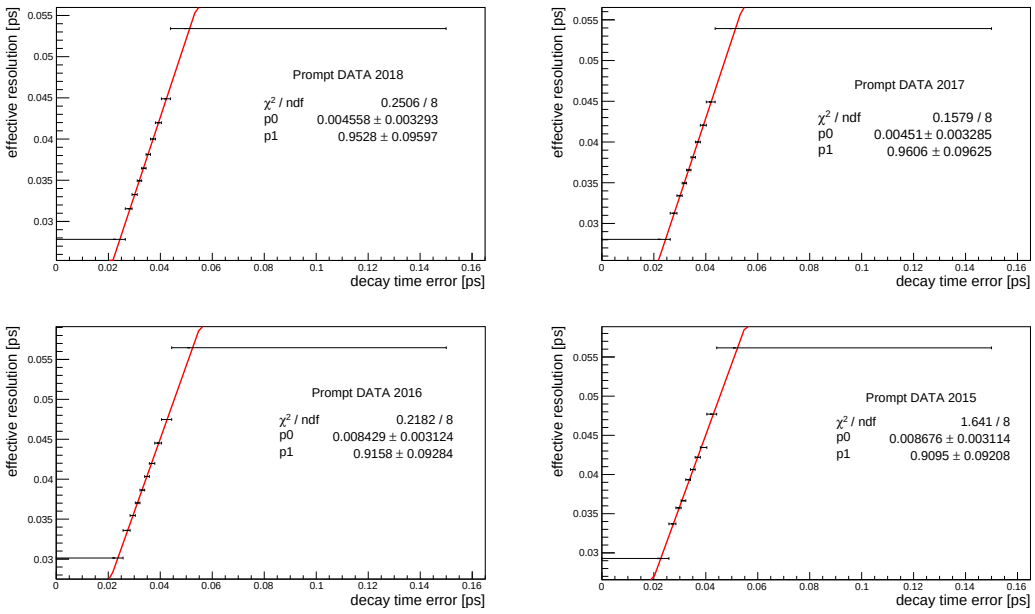

Figure 5.44: DTF decay-time uncertainty calibration functions for the different years.

The average calibrated σ_t of the prompt and signal data samples, for the different years, are summarised in Tab. 5.7. As already seen in the measurement of ϕ_s in $B_s^0 \to J/\psi K^+K^-$, there is an improvement in the resolution effect in 2017 and 2018 years.

Table 5.7: Average values of the calibrated σ_t in prompt and signal decay modes for the different data taking years.

decay mode	2015	2016	2017	2018
Prompt $J/\psi\pi^+\pi^-$	40.8 fs	40.9 fs	38.5 fs	38.3 fs
$B_s^0 \to J/\psi\pi^+\pi^-$	40.0 fs	40.4 fs	36.1 fs	36.7 fs

5.4.1.3 Resolution model

In order to extract the dilution due to the long-lived component for $t < 0$, the difference between the reconstructed and the true decay time $\delta t = t_{reco} - t_{true}$ distribution is fitted. For the prompt sample $t_{true} = 0$ by definition, therefore, t_{reco} is fitted directly. Different components contribute to the decay time distribution: one due to the real prompt events, another representing the events coming from long-lived decays and finally the events that are reconstructed using a wrongly-associated PV. The model used is similar to that presented in Chap. 4,

$$\mathcal{P}(t) = (1 - f_{wpv})\mathcal{R}(t) \otimes [f_{prompt}\delta(t) + (1 - f_{prompt})\mathcal{L}(t)] + f_{wpv}W(t), \tag{5.8}$$

where $\mathcal{L}(t)$ is the long-lived component, represented by the two exponential functions

$$\mathcal{L}(t) = (f_{sl}e^{-t/\tau_1} + (1 - f_{sl})e^{-t/\tau_2}), \tag{5.9}$$

and $\mathcal{W}(t)$ is the wrong PV component. The lifetime and relative fractions of the long-lived exponential functions are allowed to vary during the fit. The resolution $\mathcal{R}(t)$ is modelled as the sum of three Gaussians with shared means and different widths and two relative fractions, which are allowed to vary in the fit, specifically

$$\mathcal{R}(t) = \sum_{i=1}^{3} f_i \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-\frac{1}{2}(\frac{t-\mu}{\sigma_i})^2}. \tag{5.10}$$

The width of the first and second Gaussian are reparametrised to reduce correlations between the fit parameters, to further improve the convergence of the fit, for instance

$$\sigma_1 = s_1 - \sqrt{\frac{f}{1-f}}\, s_2, \tag{5.11}$$

$$\sigma_2 = s_1 + \sqrt{\frac{1-f}{f}}\, s_2, \tag{5.12}$$

where s_1, s_2 are coefficients of the parametrisation and f is the fraction of the first Gaussian. The model of the wrongly-associated PV component is determined from data, using a sample of candidates with a wrong PV association, obtained with the `PVMixer` tool. Due to lack of statistics, for the fit in 2015 the wrong PV shape obtained in the 2016 sample is used. The study of this component is discussed in the following section. The wrong PV shape is fixed in the fit, while its fraction is allowed to float in a range of about $(0.3 - 1.7)\%$, as the σ_t bin increases. Fig. 5.45 shows the fits of the decay time distribution in different bins of δ_t for 2018. In simulated signal samples, the long-lived and wrong PV components do not contribute to the $\delta t = t_{reco} - t_{true}$ distribution. Nevertheless, events where the B_s^0 candidate is originated from a B_c^+ meson have been taken into account in the fit, because of the non-negligible lifetime of the B_c^+ meson. This contribution is modelled as an exponential function convoluted with the resolution effect, where the average time of the exponential is fixed to the nominal lifetime of the B_c^+ [13].

The wrong PV component The sample obtained with the `PVMixer` tool is selected using the same requirements asked for the standard prompt sample. The combinatorial background is then removed exploiting the *sPlot* technique, with the J/ψ invariant mass as a discriminating variable. It is then divided in bins of σ_t, accordingly to the binning scheme shown in Sec. 5.4.1. Independent fits are performed in each bin. The subsamples are then weighted to match the B_s^0 signal distributions of the number of long tracks and the B_s^0 DTF χ^2/ndof. The wrong PV component shape, for each bin, is extracted by fitting the decay time distribution with a double-sided Crystal Ball function, in order to take into account the asymmetrical tails of the distribution. The shapes obtained from the fit of the first, fourth, seventh and tenth bins in the 2018 sample are shown as an example in Fig. 5.46.

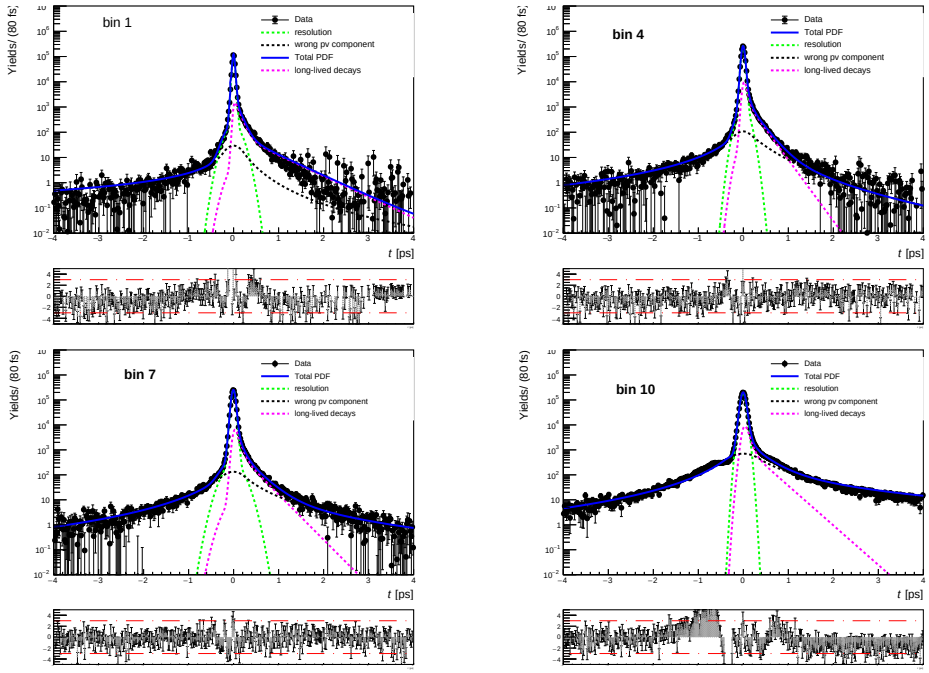

Figure 5.45: Fits of the decay time distribution in different bins of δ_t for 2018.

5.4.1.4 Measurement of the decay time resolution in simulations.

In order to validate the method used to determine the decay time resolution of B_s^0 signal decays, as explained in more details in Section 5.4.1.5, a measurement of the decay time resolution in the MC signal and prompt samples has been performed.

Prompt MC sample The MC prompt sample is selected using the same strategy developed for the data prompt one. Before performing the fits to the decay time distribution, the MC prompt sample is weighted to match the prompt data sample with the GB Reweighting method, using the same variables as for the B_s^0 signal MC. In Figs. 5.51 and 5.52 the comparison between prompt data and MC samples for 2018 is shown, before and after the reweighting.
In addition, the MC prompt sample is weighted to correct also for the differences with the signal sample and the comparison between the two samples before and after the reweighting is shown in Fig. 5.53.
The numerical computation of the dilution, in each bin of σ_t has been performed, subsequently corrected for the long-lived component, as explained in the previous chapter. The same binning scheme as in the previous section has been used. The calibration fit of the effective resolution as a function of the estimated DTF decay-time error is shown in Fig. 5.47, together with the resulted calibration parameters. The obtained average values of the effective resolution width in prompt simulated sample are shown in Tab. 5.8.

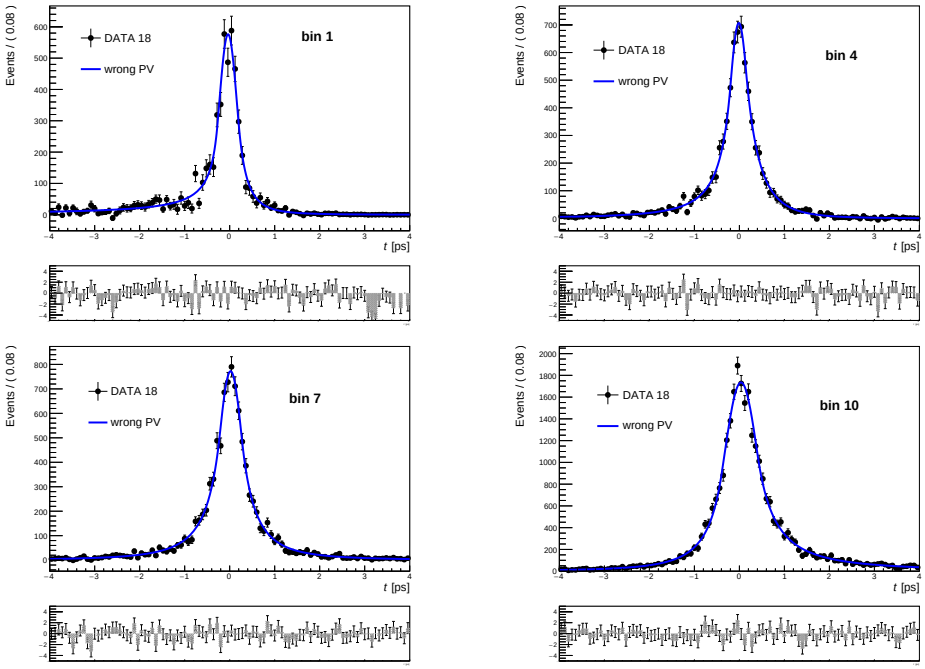

Figure 5.46: Fits of the decay time distribution of the WP sample in different bins of δ_t for 2018.

Table 5.8: Average values of the calibrated σ_t in prompt simulated samples for the different data taking years.

decay mode	2015	2016	2017	2018
MC Prompt $J/\psi\pi^+\pi^-$	37.38 fs	36.95 fs	36.765 fs	36.93 fs

Signal MC sample For comparison, in MC the dilution can be extracted also directly using the generated information. In this case, the dilution is extracted by using the $\delta t = t_{reco} - t_{true}$ distribution, so that the contribution due to the long-lived is not considered. The full selection chain used for the data signal sample is applied to the simulated signal events, before evaluating the calibration function. The same binning scheme as in the previous sections is used. The amount of B_s^0 candidates originated from the B_c^+ meson in the negative decay-time range is negligible, so the dilution is not corrected for any long-lived component. The calibration of the effective resolution as a function of the DTF decay-time uncertainty is shown in Fig. 5.48, together with the resulted calibration parameters. The obtained average effective resolutions of the selected simulated signal samples are shown in Tab. 5.9.

Table 5.9: Average values of the calibrated σ_t in signal simulated samples for the different data taking years.

decay mode	2015	2016	2017	2018
MC $B_s^0 \to J/\psi\pi^+\pi^-$	37.00 fs	36.92 fs	36.65 fs	36.77 fs

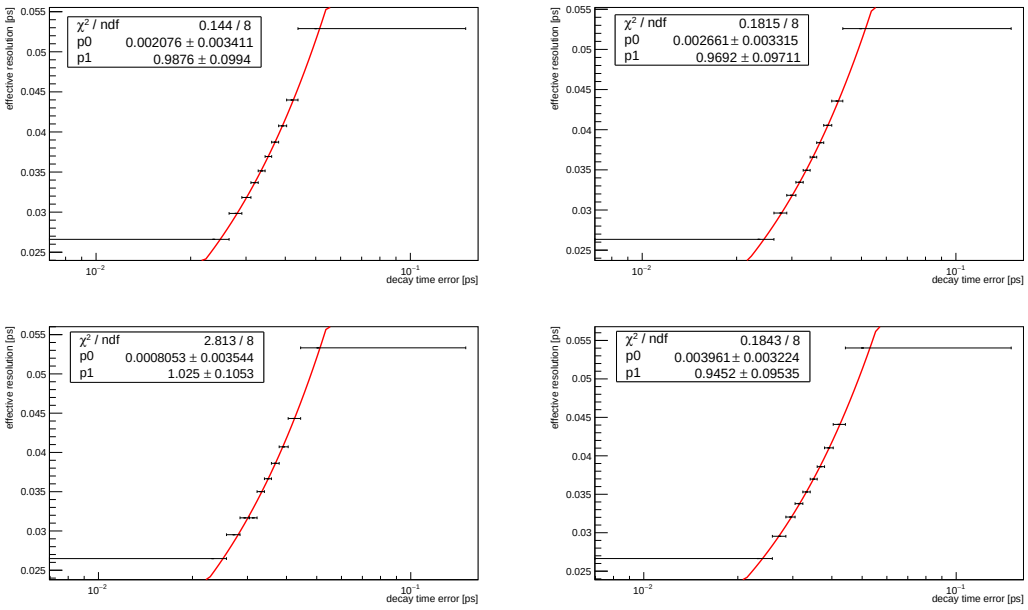

Figure 5.47: DTF decay-time uncertainty calibration functions of MC prompt samples.

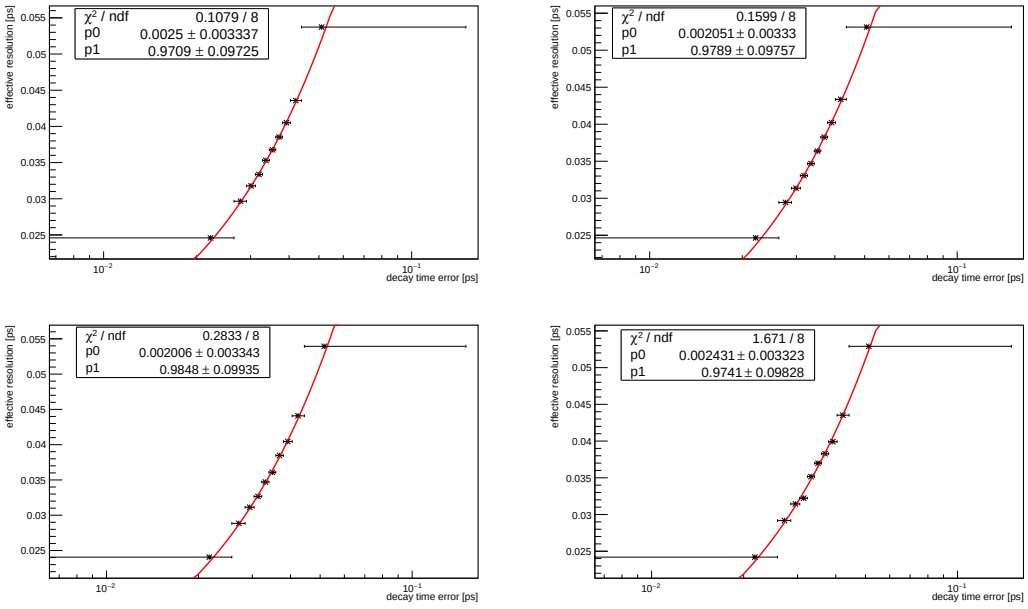

Figure 5.48: DTF decay-time uncertainty calibration functions of MC signal samples.

5.4.1.5 Validition of the calibration method

The goal of the calibration procedure is to properly estimate the calibration curve to evaluate the decay-time resolution of the signal decay on data. In order to validate the method, different checks are performed, using truth matched MC samples. Initially, the distributions of the decay-time pull $(t_{reco} - t_{true})/\delta_t$ of the prompt samples and signal MC are compared, considering $t_{true} = 0$ for the prompt samples. The pull distribution gives the average scaling between the decay-time resolution effect and the decay-time estimated error. Therefore, the more the prompt pull distributions match the corresponding data ones, the more the calibrated resolution of the data samples is a good approximation of the real resolution value. The check is initially made without reweighting the prompt samples to match the corresponding data samples in the `nLongTracks` and B_s^0`_DTF_`χ^2 variables, the so-called portability reweighting. This is done in order to evaluate the differences between the samples before the correction. Fig. 5.49 shows in the left plot the comparison of the pull distributions for 2018. The prompt data resolution shows a slight offset from zero, due to the VELO misalignment. In order to better compare the pulls, the prompt distribution is shifted to be centred at zero and the resulted comparison is shown in the right plot. The signal distribution is narrower than that of the prompt samples, especially for the data. This is expected due to different reconstructed quantities present in the two simulation samples, as no BDT is applied to the prompt one. Moreover, the positive tail in the prompt sample is due to the presence of long-lived components from b-hadron decays. From this quick comparison, it is clear that a difference in the pull distribution of the samples is present. In order to better check the precision of the calibration method, without applying the portability reweighting, the average effective resolution of the simulated signal decay, obtained using both the calibration curves from the simulated signal and the prompt samples, are compared, where the decay-time resolution is measured as explained in Section 5.4.1.4. The results for all the years are summarised in Tab. 5.10. The effective decay-time resolution of the signal decay could be estimated using the prompt sample with an accuracy of about 10%. If the portability reweighting is instead applied to the calibration sample, a significant improvement of the accuracy of the method is obtained. In fact, the results on the effective decay time resolutions of the simulated signal decay, after the reweighting of prompt samples, are shown in Tab. 5.11, and an average accuracy of about 0.6% is achieved. The pull distributions after the portability reweighting are shown in Fig. 5.50 for 2018 and a very good agreement between the samples can be observed. From these tests, a very good consistency between the calibrated curves of simulated signal and prompt samples after the portability reweighting is visible, validating the technique of using the reweighted prompt data sample to calibrate the time resolution of the signal decay.

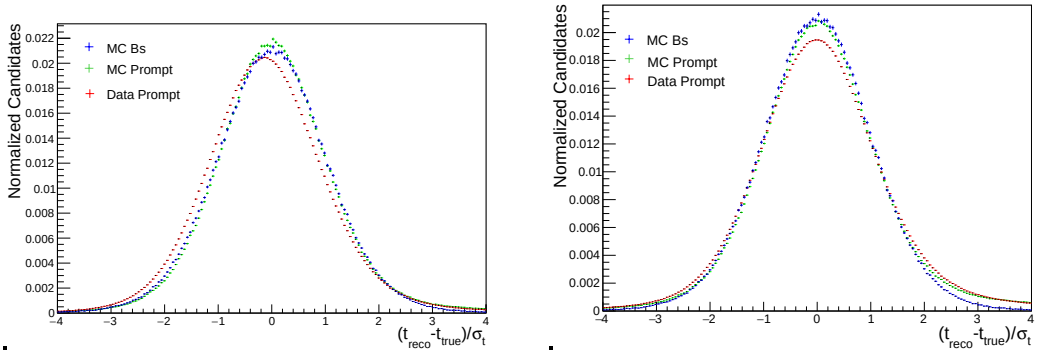

Figure 5.49: Pull distribution for 2018 dataset, $(t_{reco} - t_{true})/\delta_t$, of MC Prompt and signal modes and signal DATA samples, before the correction for the portability.

Table 5.10: Average values of the calibrated σ_t in signal simulated samples for the different data taking years, using both the prompt and signal simulated calibration curves. These results are obtained without weighting the prompt samples in the `nLongTracks` and B_s^0`_DTF_`χ^2 variables.

calibration decay mode	2015	2016	2017	2018
MC Prompt	40.94 fs	40.78 fs	40.43 fs	40.13 fs
MC $B_s^0 \to J/\psi\pi^+\pi^-$	37.00 fs	36.92 fs	36.65 fs	36.77 fs

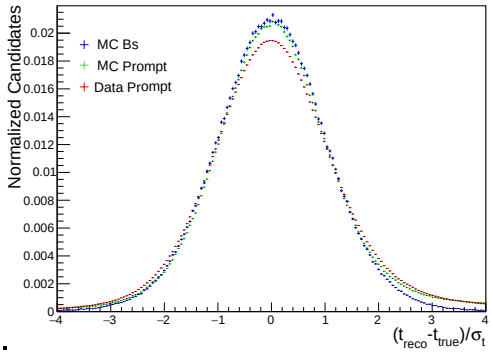

Figure 5.50: Pull distribution for the 2018 dataset, $(t_{reco} - t_{true})/\delta_t$, of MC Prompt and signal modes and signal DATA samples, after the correction for the portability.

Table 5.11: Average values of calibrated σ_t in signal simulated samples for the different data taking years, using both the prompt and signal simulated calibration curves. These results are obtained weighting the prompt samples in the `nLongTracks` and B_s^0`_DTF_`χ^2 variables.

calibration decay mode	2015	2016	2017	2018
MC Prompt	37.50 fs	37.16 fs	36.92 fs	36.94 fs
MC $B_s^0 \to J/\psi\pi^+\pi^-$	37.00 fs	36.92 fs	36.65 fs	36.77 fs

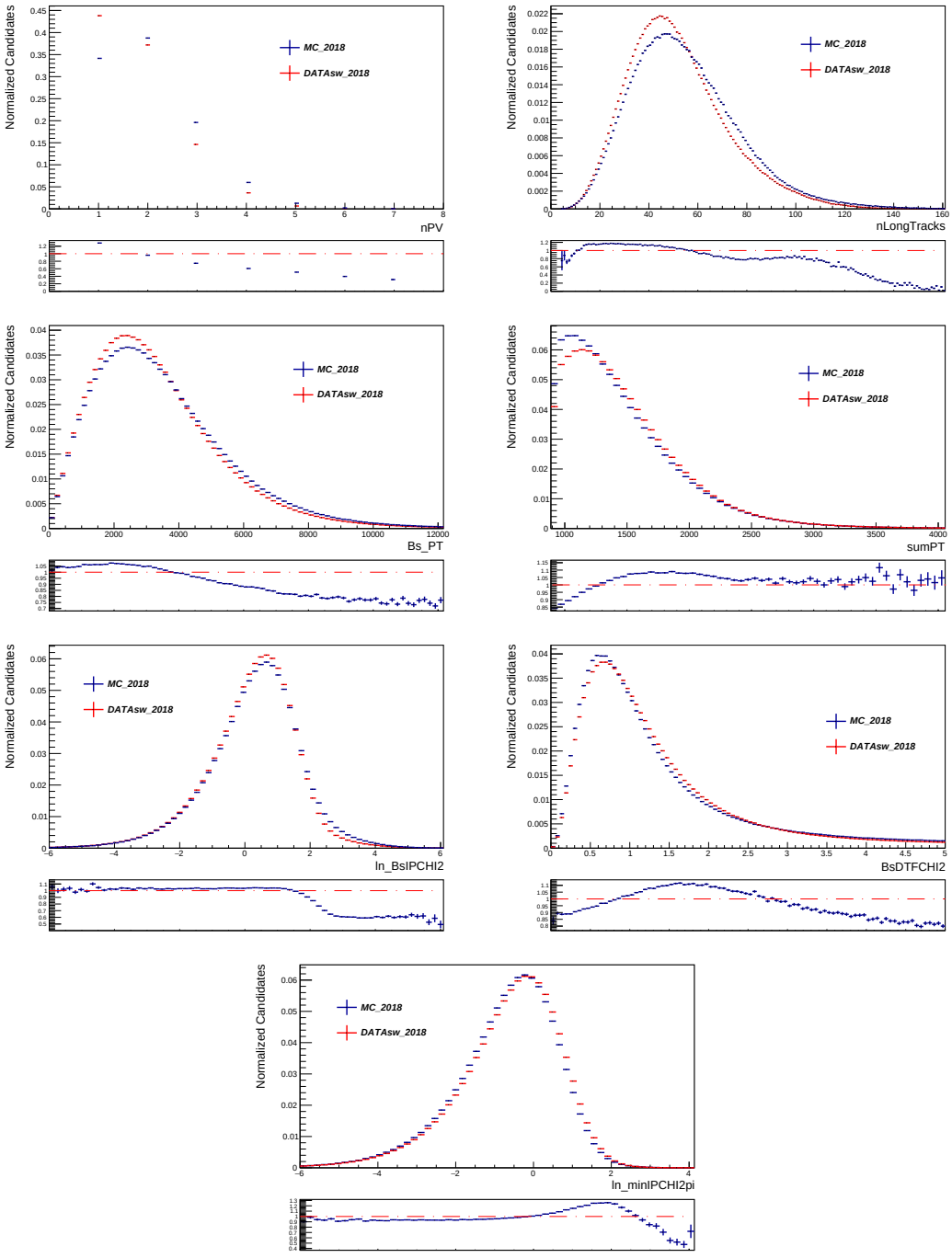

Figure 5.51: Distributions of weighting variables in Prompt MC and DATA samples, before the data-simulation correction for the 2018 dataset.

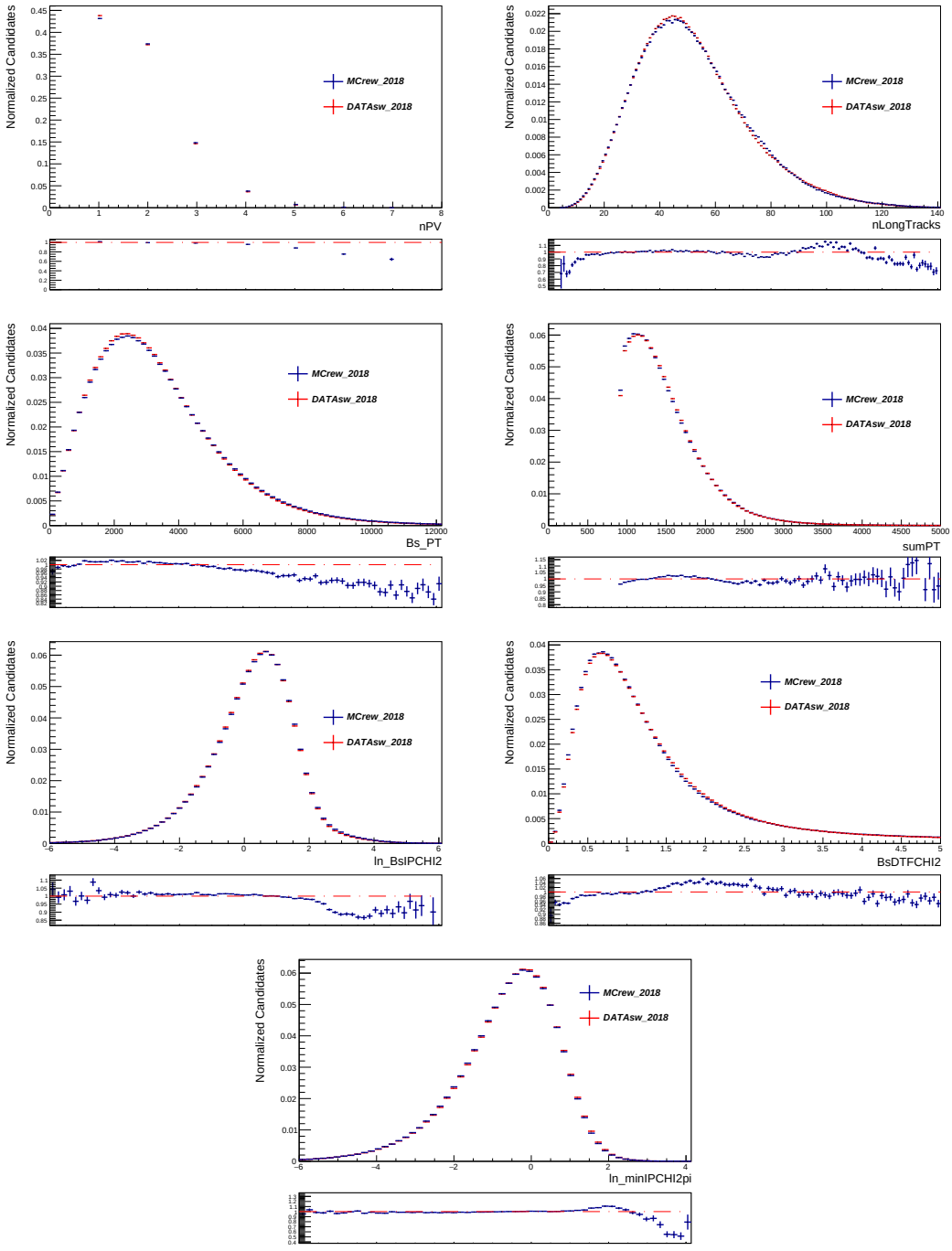

Figure 5.52: Distributions of weighting variables in Prompt MC and DATA samples, after the data-simulation correction for the 2018 dataset.

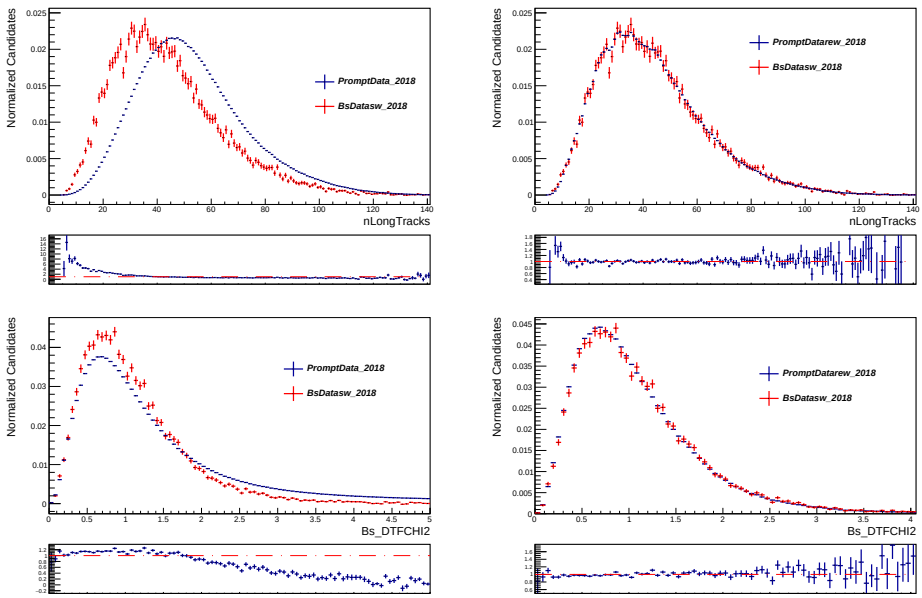

Figure 5.53: The comparison between simulated signal and prompt variables, before (left) and after (right) the reweighting procedure in the 2018 year.

CHAPTER 6

Measurement of the mixing parameter $\Delta\Gamma_d$

In this chapter, the analysis strategy developed for the ongoing measurement of the mixing parameter $\Delta\Gamma_d$ using the $B^0_d \to J/\psi K^*(892)^0$ and $B^0_d \to J/\psi K^0_s$ decays, with full Run 2 data, is presented, even though only the preliminary results obtained with the 2016 dataset will be shown in this thesis.

Currently, the measured value of $\Delta\Gamma_d$ is equal to $\Delta\Gamma_d = 0.0007 \pm 0.0066$ ps^{-1} [2] and it is compatible with zero and with the SM prediction, $\Delta\Gamma_d^{SM} = (2.6 \pm 0.4) \times 10^{-3}$ ps^{-1} [46]. However, the measured uncertainty is still one order of magnitude higher than the SM one. It is thus clear that, in order to test the SM theory prediction, futher measurements of $\Delta\Gamma_d$ are of fundamental importance.

The chapter is organised as in the following. In the first section, the selection of the $B^0_d \to J/\psi K^*(892)^0$ and $B^0_d \to J/\psi K^0_s$ candidates is explained in details. Then, in the second section, a deep look at the experimental effects that distort the decay-time distributions is given. Finally, after a discussion of the method developed to perform the final measurement and its validation, a preliminary estimate of the precision obtained with the 2016 dataset is presented, together with the expected precision reachable with the full Run 2 dataset.

I am the main contributor and proponent of the $\Delta\Gamma_d$ analysis, for which I worked on all the aspects of the analysis.

6.1 Selection of the $B_d^0 \to J/\psi K^*(892)^0$ and $B_d^0 \to J/\psi K_s^0$ signal candidates

The first step of the analysis is the reconstruction and the selection of the $B_d^0 \to J/\psi K^*(892)^0$ and $B_d^0 \to J/\psi K_s^0$ candidates, with the purpose of enhancing the signal content of the used samples with respect to the background contamination. Since the $\Delta\Gamma_d$ parameter is extracted from the decay-time distributions of the two different b-decays, as explained in Sec. 1.4, the selection strategy has been developed in order to perform a so-called time-unbiased candidate selection. This means that the selection of the candidates is performed with minimal distortion of the decay-time distributions. For this purpose, requirements on variables that are highly correlated with the B_d^0 decay time, such as the impact parameter of final state particles or of the b-meson with respect to the primary vertices, are employed as less as possible. In addition, the reconstruction of the K_s^0 decay products is made with particular attention, due to the long lifetime of this resonance. Indeed, the K_s^0 is one of the two mass eigenstates of the K^0 meson and it is a long-lived particle, with a lifetime $\tau_{K_s^0} = (0.8954 \pm 0.0004)\ 10^{-10}$ s [13]. Therefore, it flies on average few meters before it decays into a pair of oppositely charged pions. Because of this, almost two thirds of the K_s^0 decay outside of the VELO. Their decay products cannot be therefore reconstructed as Long tracks, but are reconstructed as Downstream tracks. From studies performed in simulated signal samples, see Chap. 6.2.1, it has been demonstrated that a distortion of the B_d^0 decay time is is introduced by splitting the sample into two categories, where the K_s^0 decay products are reconstructed either as Long tracks when the K_s^0 decays within the VELO or as Downstream tracks otherwise. Indeed, the probability of ending up in one of the two categories is highly correlated with the B_d^0 decay time, such that the Long track sample is enriched of short-living B_d^0 candidates and viceversa for the Downstream track sample. For this reason, all the pions coming from the K_s^0 are reconstructed as downstream tracks, performing offline the full reconstruction chain on the LHCb detector, such that no statistics is lost and, at the same time, the distortions to the decay time distribution are mitigated. The worst momentum resolution effect has been properly taken into account during the analysis, see Chap. 6.2.2.
In this section, the selection strategy is explained in details. It is divided into four main steps: a trigger and offline selection, the suppression of the combinatorial background with a multivariate-based selection, the rejection of possible misidentified decays and the final subtraction of the remaining background sources.

6.1.1 Trigger and offline selection

The first step of the selection is performed by the HLT trigger algorithms. No specific L0 line is required, since no significant decay-time distortion is introduced at this step. From studies performed on simulated samples, this allows to gain a $2-3\%$ of signal statistics with respect

to using only muon-related L0 lines. In the first stage of the high level trigger selection, only the so-called `HLT1DiMuonHighMass` trigger line is used, since it allows to select the signal with high efficiency almost without introducing distorting effects in the B_d^0 decay time distributions. In the second stage, the `HLT2DiMuonDetachedJPsi` line is employed. For both the trigger lines, only the events passing the trigger requirement as TOS are kept in the selected sample.
Because of the requirement applied by this line to the FD significance of the J/ψ with respect to the PV, it introduces mainly an acceptance effect at low decay time that has to be corrected. in order to obtain an overall trigger acceptance as similar as possible in the two decay modes, the trigger requirements are applied on the J/ψ candidates and not on the B_d^0 mesons. This means that only the muons coming from the J/ψ could be responsible for firing the trigger line. The details about the decay-time acceptance effects created in the trigger selection and their correction strategy will be discussed in Sec. 6.2.1, while more details about the trigger requirements can be found in Sec 4.2.1.
At the end of the trigger selection, only detached J/ψ particles decaying into two opposite charged muons, with a good quality decay vertex, are selected. Then, the stripping selection is performed. As it has been shown in the previous analyses, lines that directly select the candidates of interest for the specific decay mode are generally used. However, in the previous round of this measurement [1], it was found that the requirements of these stripping lines introduced several decay-time acceptance effects non-trivial to correct. For that reason, for both the decay modes, a so-called inclusive stripping line which selects only detached J/ψs is used, namely `FullDSTDiMuonJpsi2MuMuDetachedLine`, whose requirements are summarized in Tab. 6.1. Note that, the stripping selection applies a looser cut on the FDS with respect to the HLT2 line. Therefore, no additional acceptance effect is introduced at this stage.
On top of this selection, some additional requirements are applied in order to further reject the

Table 6.1: Stripping selection criteria used to identify $J/\psi \to \mu^+\mu^-$ candidates.

	Variable	Stripping
$J/\psi \to \mu^+\mu^-$	$\Delta\ln\mathcal{L}_{\mu\pi}$ (μ^+)	> 0
	p_T (μ^+)	$> 500\,\mathrm{MeV/c}$
	χ^2_{DOCA}	< 30
	$\chi^2_{vtx}/nDoF$	< 20
	$m(\mu^+\mu^-)$	$\in [2997, 3197]\,\mathrm{MeV/c^2}$
	FD significance (FDS)	$\lvert\mathrm{FDS}\rvert > 3$

background candidates or candidates that are not well reconstructed. The J/ψ $\chi^2_{vtx}/nDoF$ is required to be less than 16 and the J/ψ invariant mass has to be in the range $(3030-3150\ \mathrm{MeV/c^2})$. Then, the reconstructed and selected J/ψ candidates are combined with two oppositely charged hadrons in order to reconstruct the $B_d^0 \to J/\psi K_s^0(\to \pi^+\pi^-)$ and the $B_d^0 \to J/\psi K^*(892)^0(\to K^+\pi^-)$ decays. Since the reconstruction efficiency of a track becomes worst if the track crosses the edges of the LHCb detector, a so-called *fiducial selection* is applied to the muons, kaons

and pions, in order to have a homogeneous track reconstruction efficiency in the selected data samples. The particles of the final states are required to have a pseudorapidity in the range $2.0 < \eta < 5.0$ and a z-position of the PV within 100 mm of the nominal interaction point. Additionally, the muons transverse momentum is required to be greater than 500 MeV/c, while the pions and kaons one has to be grater than 300 MeV/c. This allows to suppress some of the particles coming from the PV, that generally have a lower momentum. In order to suppress misidentified candidates, the negative pion of both the decays is required to have $\Delta\ln\mathcal{L}_{K\pi} < 0$. No additional particle identification requirements are used on the positive hadron at this stage of the selection, to avoid to introduce different acceptance effects in the decay-time distribution of the two decay modes. In order to reconstruct the $B_d^0 \to J/\psi K_s^0$ decay, two opposite charged pions, with combined invariant mass around the K_s^0 nominal value [13] and namely in the range $422 - 572$ MeV/c^2 and total transverse momentum greater than 1200 MeV/c, are selected and combined to the J/ψ candidates. As already mentioned at the beginning of the chapter, the pions tracks are always reconstructed as Downstream tracks.
The same approach is used to reconstruct the $B_d^0 \to J/\psi K^{*0}$ decay, where a positive kaon and a negative pion, with combined invariant mass around the K^{*0} nominal value [13] and namely in the range $826 - 966$ MeV/c^2 and total transverse momentum greater than 1200 MeV/c, are selected and combined to the J/ψ candidates. Since the K^{*0} decays immediately, all the decay products are reconstructed as Long tracks.
At this point, a selection on the goodness of the $B_d^0 \to J/\psi K_s^0(\pi^+\pi^-)$ and $B_d^0 \to J/\psi K^*(892)^0(K^+\pi^-)$ candidates can be performed. For both the decay modes, a fit of the entire decay chain is performed using the DTF algorithm, in order to obtain the best determination of the decay time and mass value of each candidate. A selection on the $\chi^2/nDoF$ returned by the fit is then used to reject poorly reconstructed B_d^0 candidates or combinations of particles not really coming from a B_d^0 decay. Namely, only the B_d^0 candidates with DTF $\chi^2/nDoF < 5$ are selected. However, this requirement introduce a small acceptance effect of about 3 fs for the $B_d^0 \to J/\psi K^{*0}$ decay mode, causing a different final acceptance effect in the two modes. This acceptance is due to the small opening angle between the hadrons coming from the K^{*0} decay. If the opening angle is very small, the corresponding tracks are not enough separated inside the VELO, causing on average the share of the first hits. Then, the quality of the K^{*0} vertex becomes worse, causing therefore a consequently worsening of the B_d^0 DTF $\chi^2/nDoF$. The effect is more evident for high decay-time values, since it is more likely for the kaons and the pions to share the first hits in the VELO if the B_d^0 meson decays far from the beam axis. In order to correct for this effect, the same approach used in [1] has been employed. An alternative decay tree fit is used, in which the assigned track parameter uncertainties of the kaon and pion are increased in such a way that their contribution to the B_d^0 vertex position is negligible. After the correction, any bias effect is visible, as showed in Sec. 6.2.1.
The χ^2 of the IP, IP$_{\chi 2}$, of the B_d^0 candidate is required to be smaller than 25. This is done in

order to ensure that the b-meson originates from the PV. Since in a pp collision more than one PV is often produced and reconstructed, the b-meson candidate is associated to the PV with the smallest IP, often referred as the *best PV*. Anyway, sometimes the b-meson is associated to the wrong PV, especially when two PVs with similar distance from the b-meson candidate are reconstructed, causing a bias in the determined decay time. To minimize the possibility of misidentification of the PV, the b-meson candidates are removed if they have a $\mathrm{IP}_{\chi 2}$ with respect the second nearest PV, so-called next best PV, smaller than 50. In addition, all the events with just one PV are kept, requiring that the $\mathrm{IP}_{\chi 2}$ with respect the second nearest PV is equal to -1. This requirement is found to distort the decay-time distribution and the related effect will be discussed in more details in Sec. 6.2.1 for both the decay modes. Finally the B_d^0 DTF invariant mass is required to be in the range $(5200 - 5350)$ MeV/c^2, to further suppress the contribution of combinatorial background candidates.
The same selection, unless for the requirement on the invariant mass of the final state hadrons, is so far applied to both the decay modes, in order to take advantage from their similar topology, obtaining very similar distortions of the decay-time distributions.

6.1.2 Correction of simulated signal samples

In this analysis the simulated signal samples are extensively used:

- to develop a multivariate-based selection used to reduce the amount of combinatorial background events, keeping a high signal efficiency;
- to evaluate the efficiency on the signal of the misidentified backgrounds rejection requirements;
- to study the signal invariant mass shape;
- to measure the decay-time resolution effect, see Sec. 6.2.2;
- to determine the decay-time acceptances, see Sec. 6.2.1.

Therefore, the simulated samples have to match the data as closely as possible. As explained in Sec. 4.3.2, two different sets of corrections applied at the end of the offline selection. Initially, the PID variables, that are generally poorly described in the simulations, are corrected using the `PIDCalib` package [113]. Then, a weighting procedure is applied to the simulated samples in order to match the data in kinematics and multiplicity variables and in the variables used to train the multivariate algorithm:

- B_d^0 transverse momentum (Bd_PT);
- the sum of final state hadron transverse momentum (sumPT);
- natural logarithm of the DTF fit $\chi^2/nDoF$ of the B_d^0 (ln_BdDTFCHI2);

- the number of long tracks in an event (nLongTracks);
- the number of PVs (nPV).

The GBR method is used, with background subtracted data samples as target. These pure signal samples are obtained using the *sPlot* technique, with the DTF B_d^0 invariant mass as discriminating variable. The signal is modeled using an Ipatia function with tail parameters fixed from studies in simulated samples. The background is modeled using an exponential. The invariant mass distributions, with the resulting fit projections superimposed, are shown in Fig. 6.13 for both the decay modes. The comparison between data and simulation before and after the correction is shown in Figs. 6.2, 6.4 and 6.3, 6.5, respectively for the two channels. An evident improvement on the data and simulation agreement is visible after the correction, allowing to rely on simulated samples in the different steps of the analysis. No PID correction has been applied, since the PID simulated variables used in the analysis show already a good agreement with the data ones.
For both the decay channels, only pure signal and well reconstructed simulated events, with background category equal to 0 or 50 (see Sec. 4.2.3 for the details), are used, referred as truth-matched samples.

6.1.3 Multivariate-based selection

The invariant mass distributions of the two decay modes, shown in Fig. 6.13, highlight the presence of a large combinatorial background contribution in the two decay mode data samples. In order to enhance the signal with respect to the background, keeping a high efficiency on the signal, a boosted decision tree algorithm is trained for each channel. As usual, a corrected and truth-matched simulated signal sample is used as a proxy for the signal and the data mass sideband, namely in the range $(5380 - 5550)$ MeV/c^2, as a proxy for the combinatorial background. The training variables used in the BDT are:

- natural logarithm of the vertex fit χ^2 of the B_d^0 (ln_BdVertexCHI2);
- transverse momentum p_T of the B_d^0;
- sum of the transverse momentum of the hadrons in the final state;
- natural logarithm of the DTF fit $\chi^2/nDoF$ of the B_d^0.

No IP related variables are used in the training, to avoid to introduce big distortions of the decay-time distributions.
The signal and background samples are divided in two independent samples. One is used to train the classifier, while the other is used to test its performance. The signal and background distributions for each of the variables used in the BDT are shown in Figs. 6.7 and 6.6, for the

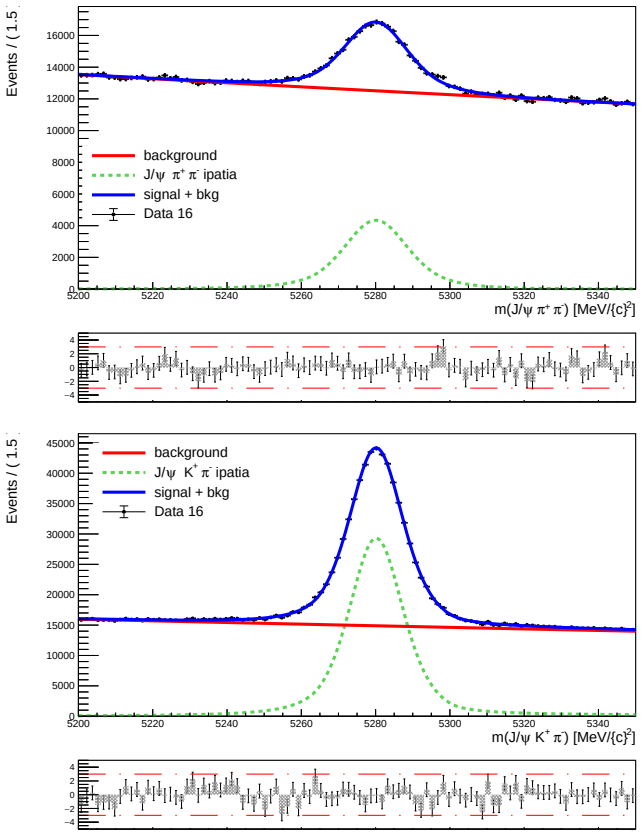

Figure 6.1: Distributions of the $J/\psi\pi^+\pi^-$ invariant mass (top) and $J/\psi K^+\pi^-$ invariant mass (bottom) with the fit projections superimposed.

two decay modes. A good discrimination power between signal and background is present in all of these variables. In order to evaluate the BDT performance, the BDT distributions for signal and background of training and testing subsamples, are shown in Fig. 6.8, for the two decay channels. The ROC curves are instead visible in Fig. 6.9, at right for the $B_d^0 \to J/\psi K_s^0$ mode and at left of the $B_d^0 \to J/\psi K^*(892)^0$. The BDTs show good performances in terms of background rejection and signal efficiency. In addition, no overtraining is present. The selection on the BDTs output is chosen to maximise the signal to noise ratio, calculated as the following figure of merit (FOM)

$$\text{FOM} = \frac{S}{\sqrt{S+N}}, \tag{6.1}$$

where S and N are, respectively, the signal and background yields, determined from the invariant mass fit that is performed at each point in the scan over the BDT response. The yields are evaluated in a region of $\pm 35\,\text{MeV/c}^2$ around the B_d^0 nominal mass [13], in order to maximise the FOM only in the region populated by the signal events. This final mass region will be used to perform the final measurement of $\Delta\Gamma_d$. The figure of merit as a function of the BDT cut for

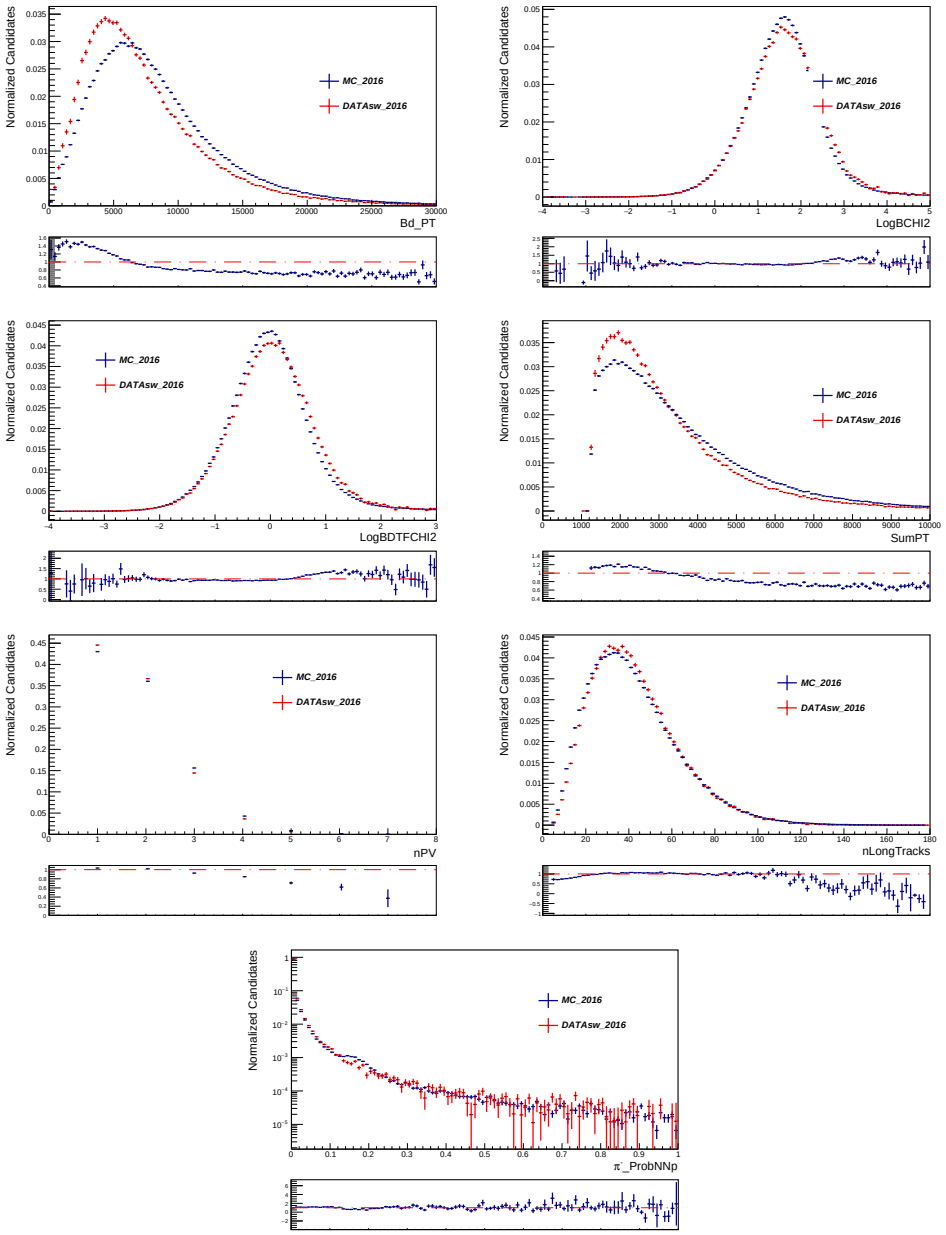

Figure 6.2: Distributions of PID and GB reweighting variables of $B_d^0 \to J/\psi K^{*0}$ channel of the background subtracted data (red) and simulation (blue) samples, before the GB correction.

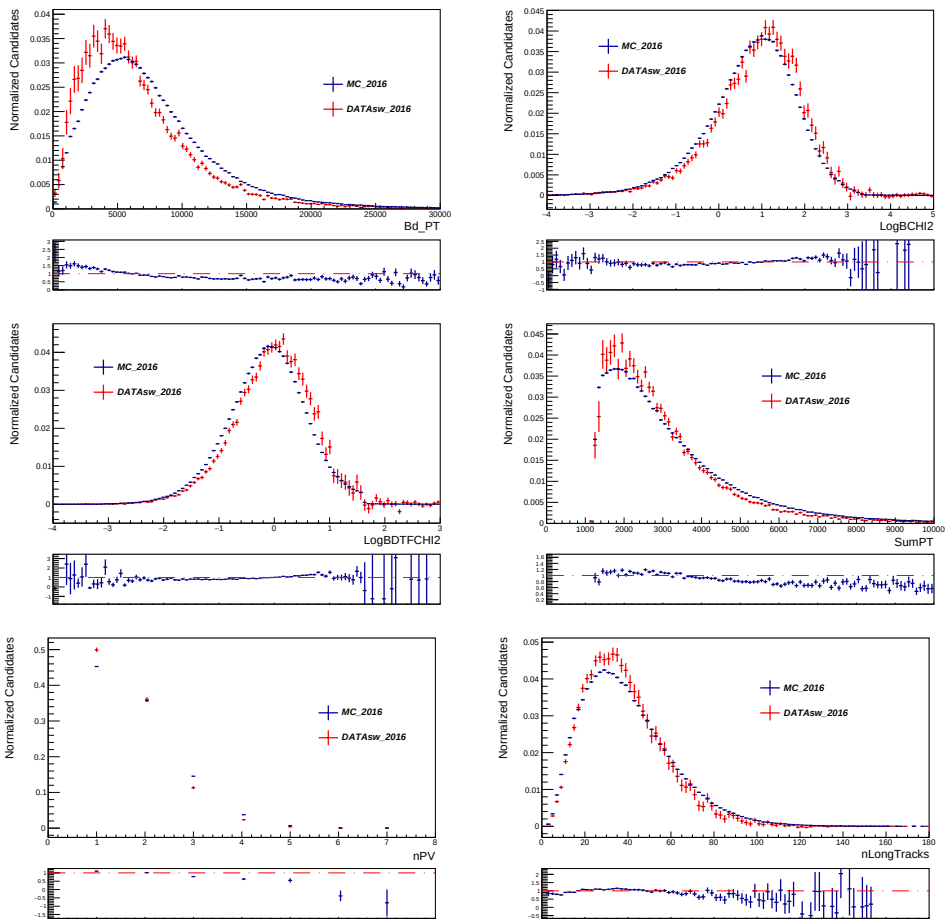

Figure 6.3: Distributions of GB reweighting variables of $B_d^0 \to J/\psi K_s^0$ channel of the background subtracted data (red) and simulation (blue) samples, before the GB correction.

the two decay modes is shown in Fig. 6.10. The optimized cut is found to be equal to -0.85 and to -0.45, respectively for the $B_d^0 \to J/\psi K^*(892)^0$ and $B_d^0 \to J/\psi K_s^0$ modes. All the events with a BDT output value larger than the optimized cut are kept in the selected samples.

6.1.4 Misidentified backgrounds

Apart from the combinatorial background candidates, the selected samples could be populated by additional contributions from physics backgrounds. These are candidates from other decay channels that pass the B_d^0 selection, because of the wrong identification of one or more of their decay products. The presence of many possible sources of misidentified decays has been investigated for both the decay modes. In particular, five main sources of misidentified background could populate the $J/\psi K^+\pi^-$ invariant mass region below the B_d^0 mass peak in the range $(5200 - 5350)$ MeV/c^2: the $B^+ \to J/\psi K^+$ and its conjugated decay, the $B_d^0 \to J/\psi \pi^+\pi^-$, the $B_s^0 \to J/\psi K^+K^-$, the $\overline{\Lambda}_b^0 \to J/\psi K^+ p^-$ with its conjugated decay, and the $\overline{B}_d^0 \to J/\psi \pi^+ K^-$.

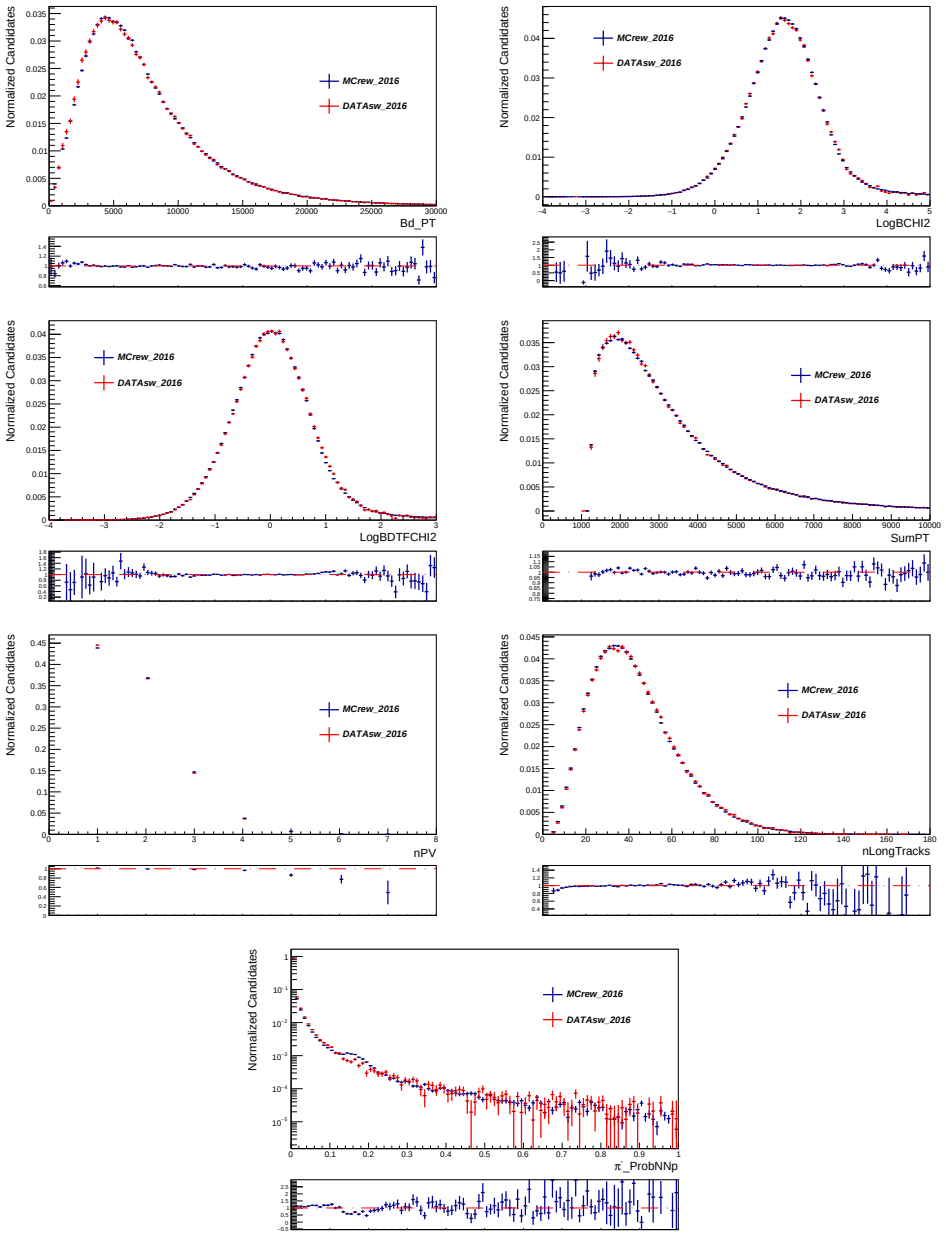

Figure 6.4: Distributions of PID and GB reweighting variables of $B_d^0 \to J/\psi K^{*0}$ channel of the background subtracted data (red) and simulation (blue) samples, after the GB correction.

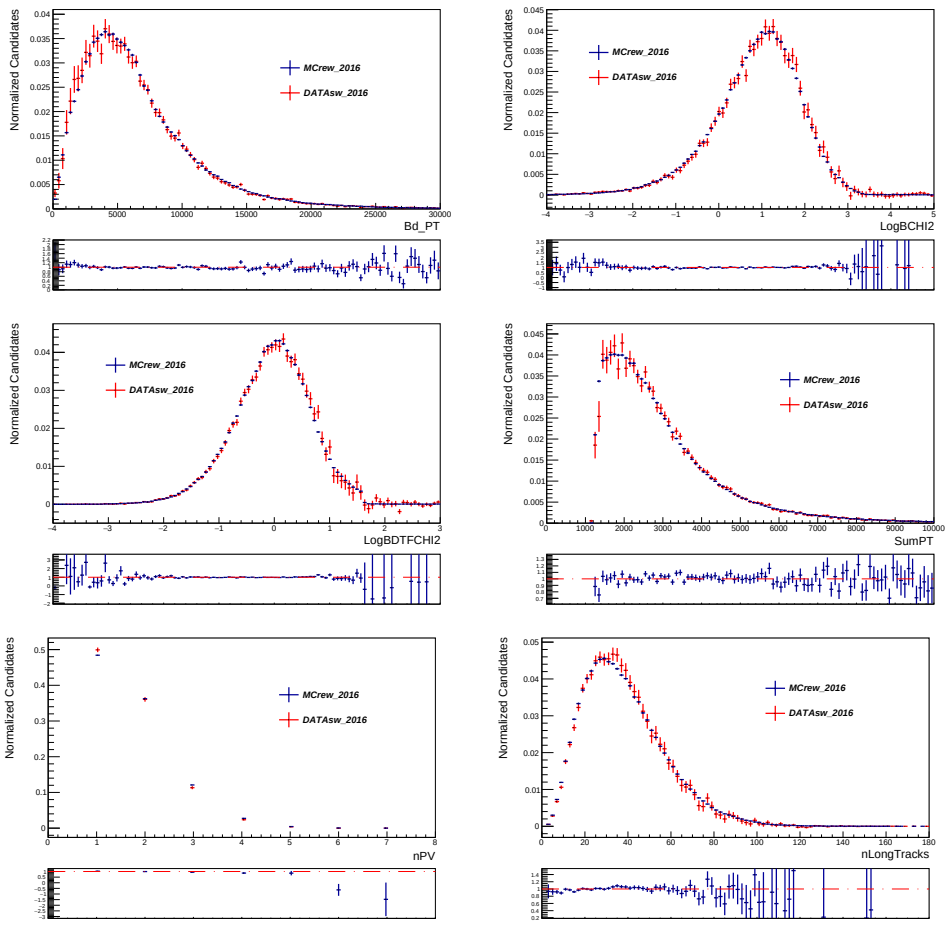

Figure 6.5: Distributions of GB reweighting variables of $B_d^0 \to J/\psi K_s^0$ channel of the background subtracted data (red) and simulation (blue) samples, after the GB correction.

Except for the $B_d^0 \to J/\psi \pi^+\pi^-$ channel, these sources can contribute also to the data sample of $B_d^0 \to J/\psi K_s^0$, plus the additional possible contribution of the $B_d^0 \to J/\psi K^+\pi^-$ decay. However, thanks to the requirement $\Delta\ln\mathcal{L}_{K\pi} < 0$ applied to the negative pions during the offline selection, most of the misidentified background events have been already rejected. No visible contributions have been found for the $B_d^0 \to J/\psi K_s^0$ decay channel, while a small contribution of $B_s^0 \to J/\psi K^+K^-$ and $\overline{\Lambda}_b^0 \to J/\psi K^+p^-$ are found in the $B_d^0 \to J/\psi K^{*0}$ sample.
The $B_s^0 \to J/\psi K^+K^-$ can be wrongly identified as a $B_d^0 \to J/\psi K^+\pi^-$ signal candidate if the negative kaon is identified as a pion. The mass distribution of this contribution has been recalculated from the B_d^0 data sample, assigning to the negative pion the kaon mass. The resulting plot is shown in Fig. 6.11 and a small peaking structure around the nominal B_s^0 mass [13] is visible. A fit of the peak has been performed and 1560 ± 63 events of $B_s^0 \to J/\psi K^+K^-$ have been found. The application of rejection requirements, based on particle identification variables, has been found not enough efficient to remove the background. Therefore, they have

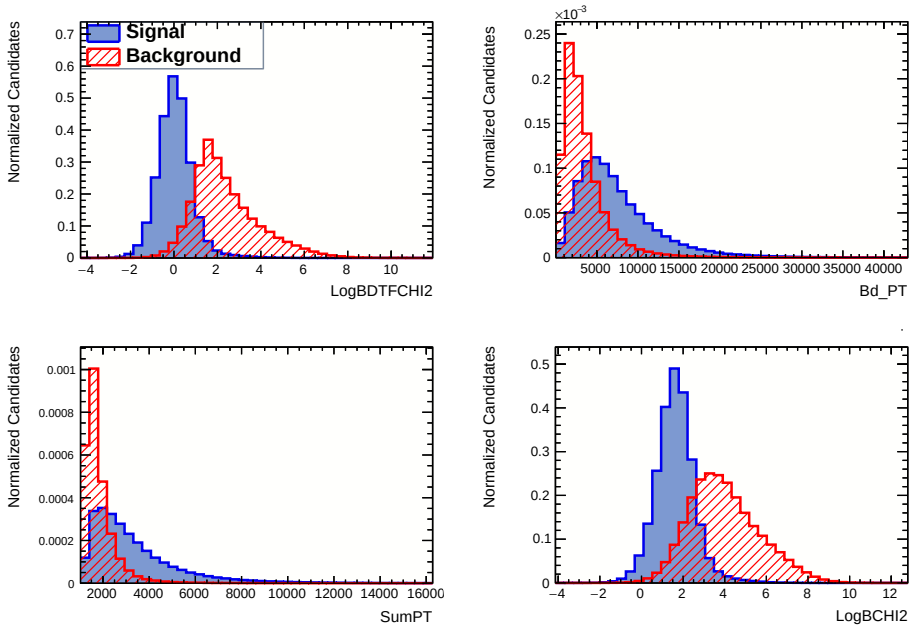

Figure 6.6: Signal and background distributions of the variables used to train the BDT algorithm for the $B^0_d \to J/\psi K^{*0}$ decay mode.

to be subtracted using a different strategy, and the studies on this topic are still ongoing.
Background events of $\overline{\Lambda}^0_b \to J/\psi K^+ p^-$ can be reconstructed as a B^0_d candidate if the proton is misidentified as a pion. Figure 6.12 shows in blue the recomputed $J/\psi Kp$ mass distribution obtained from the B^0_d data sample, when interpreting the p as a π, and a relative wide distribution of events is visible around the nominal Λ^0_b mass [13]. This peaking background is largely suppressed by vetoing the candidate if the pion has `ProbNNp` > 0.2 and if the corresponding recalculated $J/\psi K^+ p^-$ mass is consistent with the known Λ^0_b value within $\pm 3\sigma$, where the sigma found for the same background in the $B^0_s \to J/\psi \pi^+ \pi^-$ analysis is used. The recalculated mass distribution after the veto is shown in red in Fig. 6.12. The veto efficiency on the signal events is studied in corrected truth-matched simulated samples and it is about 99.97%.

6.1.5 Subtraction of the remaining background candidates

After the selection, a large amount of combinatorial background events is still present in the data samples, in particular for the $B^0_d \to J/\psi K^0_s$ channel, as visible in Fig. 6.13 for both the decay modes. Futher BDT training variables are under investigation, in order to optimise the rejection of the background in both the samples causing, at the same time, a similar decay-time acceptance due to the final BDT selection.
This background component has a very different decay-time distribution with respect to the

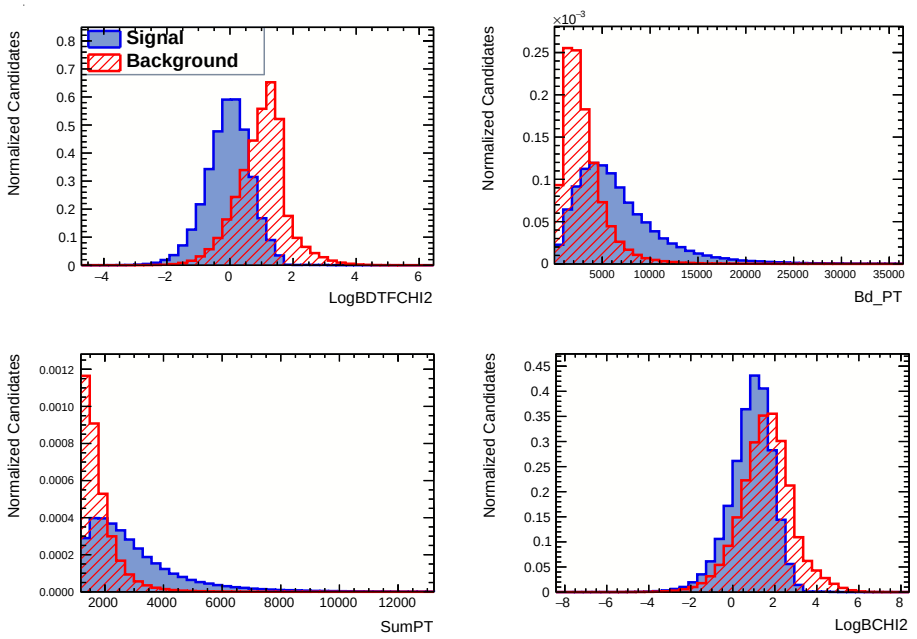

Figure 6.7: Signal and background distributions of the variables used to train the BDT algorithm for the $B_d^0 \to J/\psi K_s^0$ decay mode.

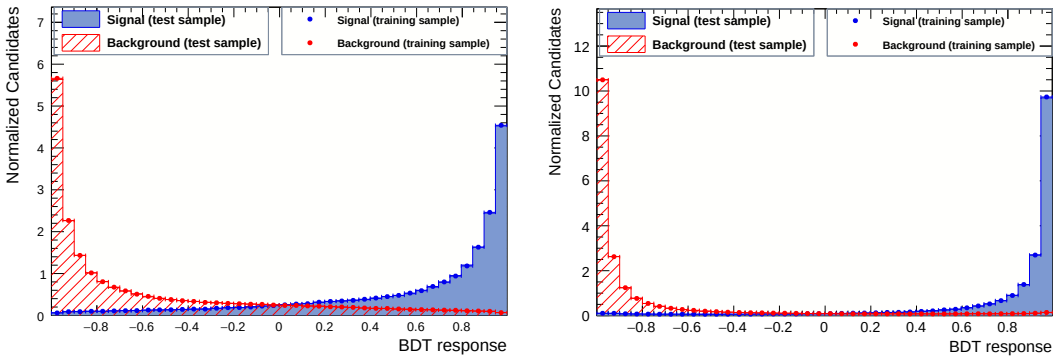

Figure 6.8: BDT distributions for signal and background of training and testing subsamples for the $B_d^0 \to J/\psi K_s^0$ channel (left) and $B_d^0 \to J/\psi K^{*0}$ channel (right).

signal, with a sensibly lower lifetime. To avoid the introduction of a bias in the measurement of $\Delta\Gamma_d$, the *sPlot* technique is used to statistically subtract the combinatorial background events. The B_d^0 invariant mass is exploited as target variable, where the signal and background shapes can be easily separated. The signal is described with an Ipatia function, while the background component is modeled as an exponential. The Ipatia tails are fixed from studies performed on the selected, truth-matched and corrected simulated signal samples, in order to avoid to wrongly treat signal events as background ones. The resulting fitted function are shown in Fig. 6.13 and

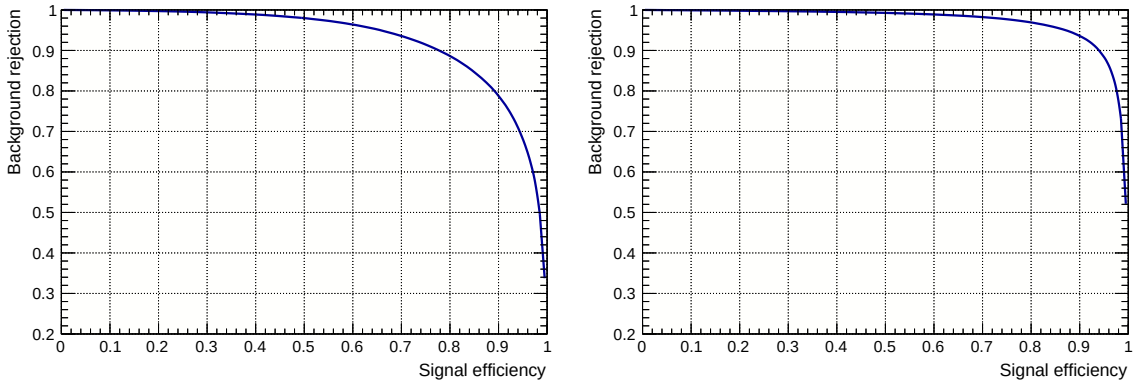

Figure 6.9: ROC curves for the $B_d^0 \to J/\psi K_s^0$ channel (left) and $B_d^0 \to J/\psi K^{*0}$ channel (right).

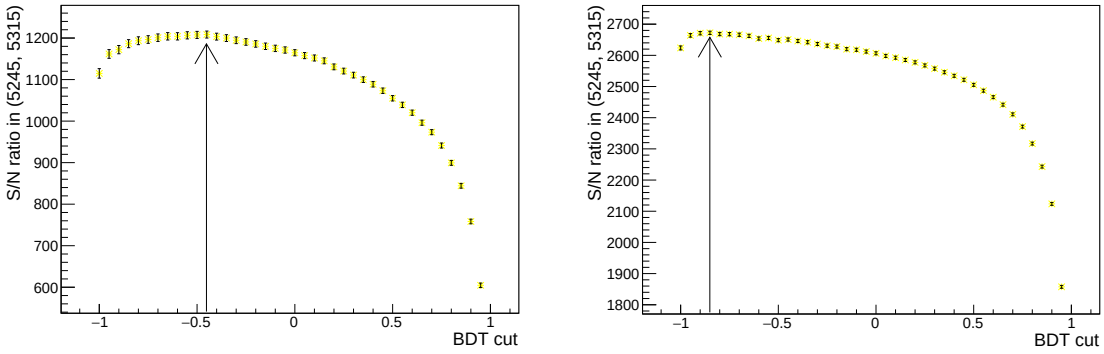

Figure 6.10: Distributions of the figure of merit for the $B_d^0 \to J/\psi K_s^0$ channel (left) and $B_d^0 \to J/\psi K^{*0}$ channel (right). The black arrow highlights the optimized cut.

a signal yield of about 365800 ± 2200 and 69580 ± 740 has been obtained, respectively in the $B_d^0 \to J/\psi K^{*0}$ and $B_d^0 \to J/\psi K_s^0$ decay modes.

6.2 Experimental effects

The most important and challenging part of this analysis is understanding and controlling the distortions of the decay-time distributions caused by the reconstruction and selection procedure. In Sec. 6.2.1, the different acceptance effects are analysed in detail for both the decay modes and classified into two main categories: the so-called lower and upper decay-time acceptance, depending if more events are removed at low or high decay times by a particular step in the reconstruction or selection procedure. In addition, a comparison of the acceptance effects between the two decay modes is made for all the analysis steps, in order to evaluate the similarity between the two samples.

Moreover, in Sec. 6.2.2, the measurement of the decay-time resolution width is also provided, in

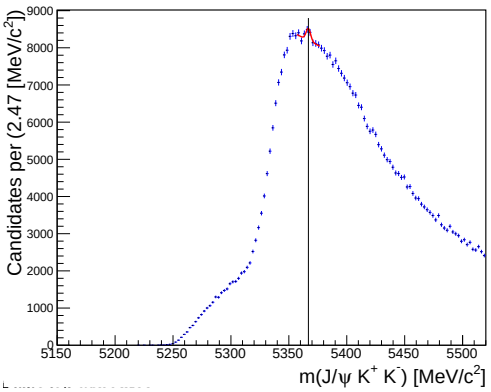

Figure 6.11: Distribution of the recomputed $m(J/\psi K^+K^-)$ in blue, with the Gaussian resulted from the fit shown in red. The black line highlights the central value of the B_s^0 invariant mass [13].

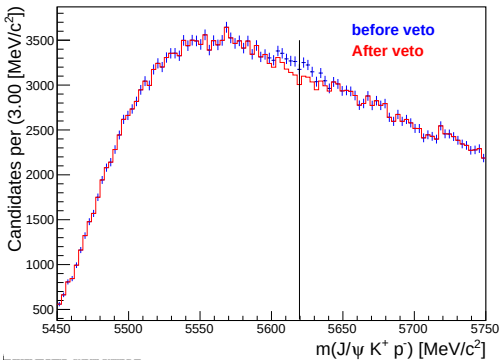

Figure 6.12: Distribution of the recomputed $m(J/\psi K^+p^-)$ before, in blue, and after, in red, the application of the veto. The black line highlights the central value of the $\bar{\Lambda}_b^0$ invariant mass [13].

order to take this effect under control during the final measurement.

6.2.1 Decay-time acceptance

In order to study the different sources of the decay-time acceptance it is useful to look at the simulated samples. The main advantage of simulated samples is the possibility to study, directly looking at the decay-time distribution, the effect due to the reconstruction procedure, information not accessible by using the data samples. In addition, in the simulated samples, the distortion effect due to the decay-time resolution can be isolated, since it allows to access the generated value of the decay time of each candidate, the so-called t_{true}. However, the reliability of simulated samples is strictly subordinated on how good is their description of the real signal data. For that reason, only truth-matched and corrected simulated samples are

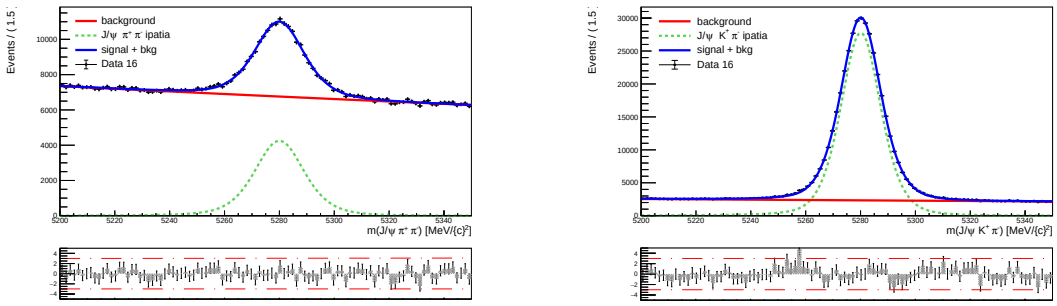

Figure 6.13: Distributions of the $J/\psi\pi^+\pi^-$ invariant mass (top) and $J/\psi K^+\pi^-$ invariant mass (bottom) with the fit projections superimposed, after the entire selection procedure.

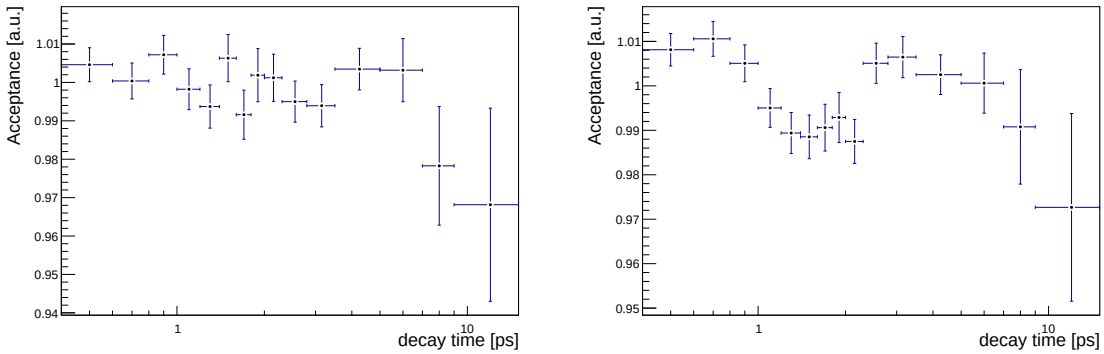

Figure 6.14: Acceptance effect on the decay time distribution after the reconstruction and selection procedure, for the $B_d^0 \to J/\psi K_s^0$ (left) and $B_d^0 \to J/\psi K^{*0}$ (right) decays, studied in simulated samples.

used to extract the total acceptance. The later is obtained by dividing the t_{true} decay-time distribution of the B_d^0 meson that survives the reconstruction and selection procedures by the generated distribution. The latest represents the decay-time distribution of the candidates before the application of any reconstruction or selection requirements. The LHCb framework allow to perform a match between the generated candidates and those that survive the reconstruction. If a generated candidate and a reconstructed candidate share a certain percentage of hits, a specific matching value is assigned to the so-called `IsReconstructed` variable in the generated simulated sample. The value 1 corresponds to a track reconstructed as long, while the value 2 to a track reconstructed as downstream. This allows to study individually the reconstruction effects due to each particle in the final state, as it will be shown in the following.
The individual simulated distributions are normalised to the same area. The obtained decay-time acceptance is shown in Fig. 6.14, on the left for the $B_d^0 \to J/\psi K_s^0$ channel and on the right for the $B_d^0 \to J/\psi K^*(892)^0$ one. The most evident effect, common to both the decay modes, is the decrease of the acceptance to reconstruct and select the B_d^0 candidates with a large decay time,

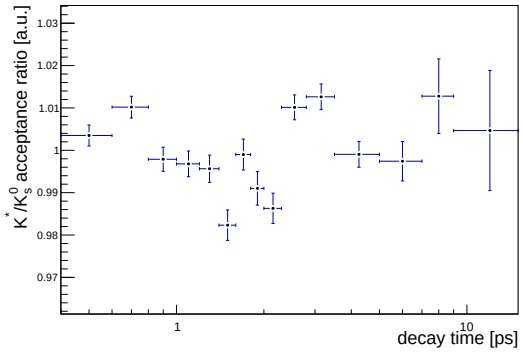

Figure 6.15: Ratio between the total acceptances on the decay time distribution of the $B^0_d \to J/\psi K^{*0}$ and $B^0_d \to J/\psi K^0_s$ decays, studied in simulated samples.

and it is referred to as upper decay-time acceptance. It is mostly due to the track reconstruction efficiency of muons and it has a similar behaviour for the two channels. The K^{*0} acceptance shows a drop around 2 ps, not present in the K^0_s channel. The drop effect is due to the fact that the B^0_d decay products can be wrongly used to form an additional PV, as already explained in Sec. 4.5.3. If this happens, the reconstructed B^0_d decay-time is smaller than the true value. Therefore, the candidate does not pass the displacement requirements made by the trigger and stripping selections. This happens more likely for B^0_d with low decay time, even thought if the decay-time value is very low the B^0_d decay products are more often used to reconstruct the true PV, explaining why this effect is not present at very low decay-time values. In the case of the decay involving the K^0_s resonance, since the K^0_s is a long-lived particle, its decay products are unlikely used to reconstruct an additional PV, avoiding the rejection of the events because of the displacement requirements. The ratio between the total acceptances of the two decay modes is shown in Fig. 6.15. Except for the drop around 2 ps, that is anyhow small, only very minimal deviations from the flatness are visible, ensuring a very similar acceptance effect in the two decay-time distributions.

In order to better understand the impact of the different selection steps in distorting the decay time distributions, the lifetime, for each step, is extracted by means of a fit to the true decay-time distribution in simulated samples, using a single exponential function. The result of the scan, for both the decay modes, is shown in Fig. 6.16, where all the final state particles are reconstructed as long tracks, except for the pions of the $B^0_d \to J/\psi K^0_s$ decay mode that are instead reconstructed as downstream. The first point represents the fitted value of the lifetime of the generated decay-time distribution, when all the final tracks are within the LHCb detector acceptance, and it is the same for both the decay modes. The last point refers to the lifetime value obtained at the end of the entire selection procedure and it is evaluated on the corrected reconstructed simulated samples. The comparison between the first and the last steps gives a quantitative estimation of the bias introduced by the reconstruction and selection procedures,

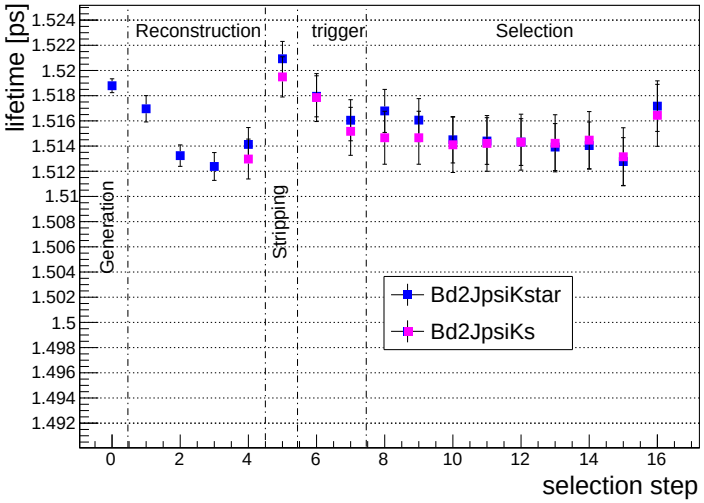

Figure 6.16: Measured lifetime in simulated samples of $B_d^0 \to J/\psi K^{*0}$ and $B_d^0 \to J/\psi K_s^0$ decay channels as a function of the different reconstruction, selection and trigger requirements used in the analysis. For a detailed explanation of the analysis steps see B.

that is of about 2 fs and 3 fs respectively for the $B_d^0 \to J/\psi K^{*0}$ and $B_d^0 \to J/\psi K_s^0$ channels. This result is really promising, showing not only the almost perfect similarity between the two decay modes, but more importantly the fact that all the distortion effects are rather small, especially when compared to Run 1, where overall acceptance effects of about 20 fs were found [1].

All the steps, except for the first, contain the fiducial selection and the selection applied in the previous step. In the steps 1 and 2, the effect due to the reconstruction of the final state hadrons and muons, respectively, is shown. It is evident that the muons cause the most important effect. Step 3 reports the lifetime obtained when all the final state particles are reconstructed. It is almost the same lifetime obtained requiring the reconstruction of only the muons, highlighting the almost negligible reconstruction effect due to the hadrons, when the muons have been already reconstructed. The steps from 0 to 3 are only shown for the $B_d^0 \to J/\psi K^{*0}$ decays, because the $B_d^0 \to J/\psi K_s^0$ simulated samples at generator level with the pions reconstructed as downstream tracks are still in production. The reconstruction effect in the reconstructed simulated sample is shown at step 4 and, considering the $B_d^0 \to J/\psi K^{*0}$ mode, it is in agreement with what has been found with the generated sample at step 3, ensuring the validity of the `IsReconstructed` variable. From this step, all the following steps are made by using the reconstructed simulated sample. Step 5 refers to the stripping selection, that creates a significant bias on the decay time due to the displacement requirements applied on the J/ψ. In addition to the reconstruction and stripping acceptance, the most biasing effects are due to the trigger selection. The HLT1 line used does not contain, in principal, decay-time biasing requirements. So the effect seen

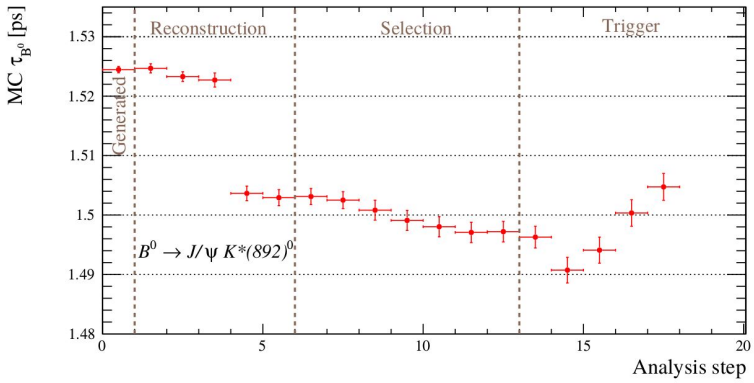

Figure 6.17: Measured lifetime in Run 1 simulated sample of $B_d^0 \to J/\psi K^{*0}$ channel as a function of the different reconstruction, selection and trigger requirements used in the analysis. For a detailed explanation of the analysis steps see [1].

at point 6 is probably due to the difference between the reconstruction algorithm used during the online reconstruction performed by LHCb in the HLT1 selection and the offline algorithm used when the trigger line selection is applied by the users. In fact, in order to be faster and usable during the data taking phase, the online algorithm is simpler than the offline one and, therefore, it performs a less sophisticated and precise selection, causing the loss of some events. The HLT2 effect studied in point 7 is, instead due to the stricter requirements on the J/ψ displacement with respect what has been applied during the stripping. Steps 8, 9 and 10 refer to the selection performed respectively on the final state hadrons, on the muons and on the resonance. In particular, steps 8 and 10 create a small bias in the $B_d^0 \to J/\psi K^{*0}$ decay mode, respectively because of the particle identification requirement asked for the negative pion and for the cut on the transverse momentum of the K^{*0}. From point 11 to 14 the lifetime obtained respectively after the offline selection on the J/ψ, the requirement on the B_d^0 invariant mass, the selection on the shifted DTF χ^2 of the B_d^0 and the requirement $B_d^0\ IP_{\chi^2} < 25$ is showed and no biasing effect is visible. The last two steps refer respectively to the requirement on the $B_d^0\ IP_{\chi^2} next\ best > 50$ or $= -1$ and to the BDT selection. In the former, the bias is due to the requirement $B_d^0\ IP_{\chi^2} next\ best = -1$. It is equivalent to ask that only events with one PV are kept in the sample and the motivation of the bias are still under study. The BDT selection bias in, instead, due to the fact that, in this preliminary stage, it has been training not using the B_d^0 shifted DTF χ^2 variable but the normal one. This is the reason why the effect is smaller for the $B_d^0 \to J/\psi K_s^0$ channel than for the $B_d^0 \to J/\psi K^{*0}$ one.

In Fig. 6.17, the lifetime scan as a function of the selection for the $B_d^0 \to J/\psi K^{*0}$ decay channel in the $\Delta\Gamma_d$ measurement performed with the 2011 dataset is shown. A total biasing effect of about 18 fs was found, mostly due to a rather large VELO reconstruction efficiency effect that is responsible of 20 fs of bias. For that reason, a most complicated analysis has been performed, with the development of a data-driven method for the correction of the reconstruction acceptance.

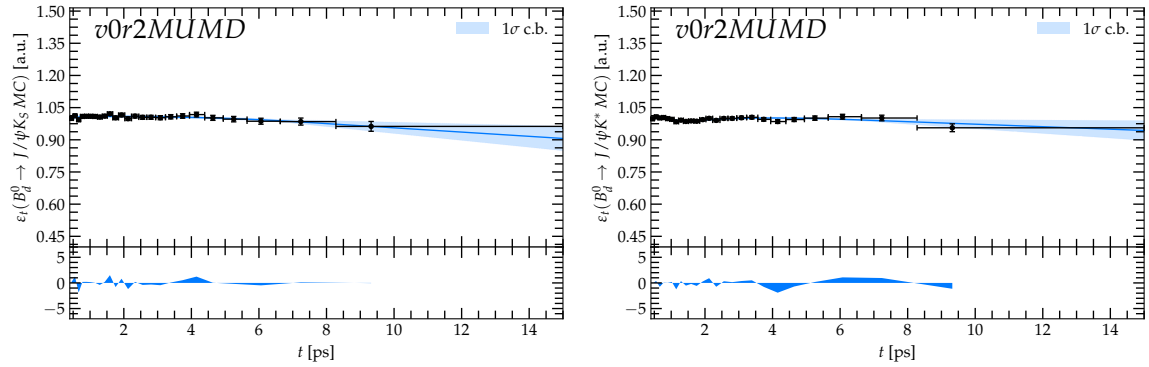

Figure 6.18: Decay-time acceptance of corrected simulated samples of $B_d^0 \to J/\psi K_s^0$ on the left and $B_d^0 \to J/\psi K^{*0}$ on the right, with the fit projections superimposed.

Since then, a big technical effort has been done by the LHCb collaboration, in order to minimize this effect in the Run2 data taking period, by the optimisation of the VELO reconstructed algorithm, reaching a total bias due the reconstruction of about 7 fs.
Considering the significantly smaller bias effect visible in Run 2, a simpler analysis has being carried out with respect the one performed in Run 1, keeping into account that while with the 2011 dataset also the absolute lifetime of B_d^0 mesons was measured, in this analysis we will measure only $\Delta\Gamma_d$ exploiting the cancellation of effects between the two decay modes. The acceptances are directly measured in truth-matched and corrected simulated samples. The decay-time acceptance of each decay mode is modeled as a cubic spline function, with knots positioned at [0.3, 0.91, 1.96, 9.00] ps, in order to account for the exponential nature of the distribution. Since the simulated samples are generated with $\Delta\Gamma_d = 0$, in the absence of acceptance and resolution effects, their decay-time distributions can be modeled with a single `true` exponential function with lifetime fixed to the known B_d^0 lifetime value. Then, each experimental decay-time distribution is described by the true exponential function convoluted with a Gaussian resolution function and multiplied with the acceptance. The Gaussian resolution model is centred at zero and its width is fixed from simulation studies, as shown in Sec. 6.2.2. The measured decay-time acceptances for the two decay modes are shown in Fig. 6.18 and the obtained coefficients will be used as fixed parameters in the final measurement.

6.2.2 Decay-time resolution

The decay-time resolution is a necessary component of the probability density function used to describe the selected decay-time distribution, in particular when performing precise *CP* or mixing parameters measurements. However, since no $B - \bar{B}$ oscillations have to be resolved in the measurement of $\Delta\Gamma_d$, a simpler approach is used with respect to the ϕ_s analyses described in the thesis.
The average resolution width is directly measured in the truth-matched and corrected simulated

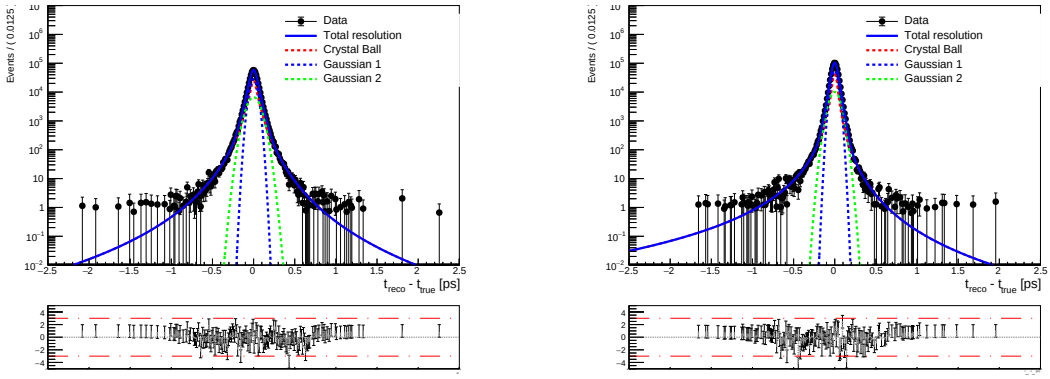

Figure 6.19: Distributions of the $t_{reco} - t_{true}$ observable, with fit projections superimposed, for the $B_d^0 \to J/\psi K_s^0$ channel (left) and $B_d^0 \to J/\psi K^{*0}$ channel (right).

signal samples, where the difference between the reconstructed and generated decay time $t_{reco} - t_{true}$ is studied. The resolution model $\mathcal{R}(t)$ composed of a double-sided Crystal Ball $\mathcal{CB}(x)$ function plus two Gaussians, with shared means and different widths, is used

$$\mathcal{R}(t) = f_{CB} \times \mathcal{CB}(t|\mu, \sigma_{CB}) + \sum_{i=1}^{2} f_i \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}(\frac{t-\mu}{\sigma_i})^2}. \tag{6.2}$$

As already done for the ϕ_s measurements, the widths of the Crystal Ball and first Gaussian are reparametrised to reduce correlations between the fit parameters, to further improve the convergence of the fit, namely

$$\sigma_{CB} = s_1 - \sqrt{\frac{f}{1-f}}\, s_2, \tag{6.3}$$

$$\sigma_1 = s_1 + \sqrt{\frac{1-f}{f}}\, s_2, \tag{6.4}$$

where s_1, s_2 are coefficients of the parametrisation and f is the fraction of the Crystal Ball. The Crystal Ball function is used in order to better fit the not completely symmetric tails of the $t_{reco} - t_{true}$ distribution. The results of the fits are shown in Fig. 6.19 for both the decay modes and the standard deviation σ_t of the fitted $\mathcal{R}(t)$ model is used as measured value of the average resolution effect, found to be of about 58 fs and 45 fs in the $B_d^0 \to J/\psi K_s^0$ and $B_d^0 \to J/\psi K^{*0}$ decay channel, respectively. The worst resolution effect in the $B_d^0 \to J/\psi K_s^0$ channel is due to the fact that the pions are reconstructed as downstream tracks, obtaining on average a worst vertex resolution.

6.3 Preliminary results

Two different fitters have been studied in order to perform the measurement. The validation of the fitters has been performed by using the corrected simulated samples, for which $\Delta\Gamma_d$ is

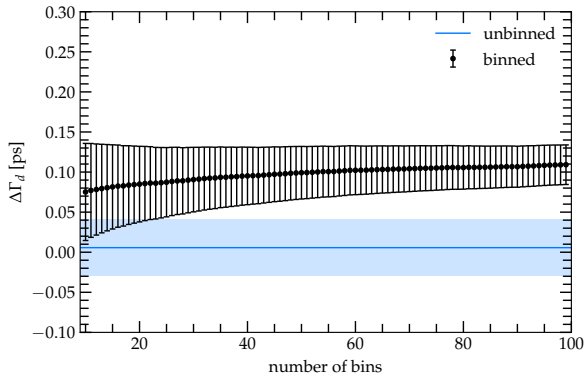

Figure 6.20: The measured $\Delta\Gamma_d$ value in simulated samples for the different fitters. The unbinned method do not depend from the binning, so it has been shown as a bar at the obtained value from the fit.

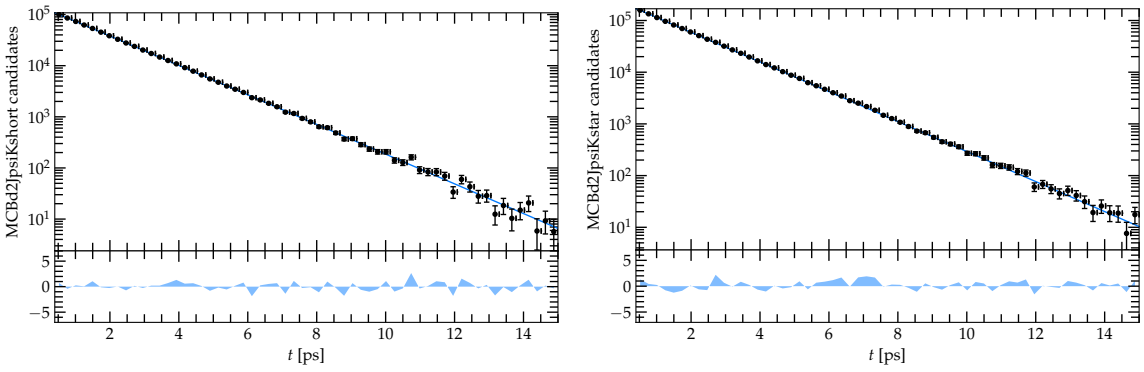

Figure 6.21: Decay-time distribution of corrected simulated samples of $B_d^0 \to J/\psi K_s^0$ on the left and $B_d^0 \to J/\psi K^{*0}$ on the right, with the fit projections resulting from the simultaneous fit superimposed.

generated equal to zero.

The decay-time distribution of the two channels $T(t)$ is described as a single exponential function, with the effective lifetime defined in Sec. 1.4, convoluted by a single Gaussian function, with width equal to the measured value of the average resolution effect

$$T(t) = e^{-t/\tau_{eff}} \otimes G(t|\sigma_t). \tag{6.5}$$

In all the fitters, the Γ_d and β parameters are fixed to the known values [2], namely $\beta = (22.2 \pm 0.7)^\circ$ and $\Gamma_d = 0.6583 \pm 0.0017$ ps^{-1}. Initially, a binned maximum likelihood fit of the ratio between the $B_d^0 \to J/\psi K_s^0$ and $B_d^0 \to J/\psi K^{*0}$ decay-time distributions has been performed. In this case, the total probability density function $P(t)$ used in the fit is defined as the ratio between the $T(t)$ function of the $B_d^0 \to J/\psi K_s^0$ and $B_d^0 \to J/\psi K^{*0}$ decay modes,

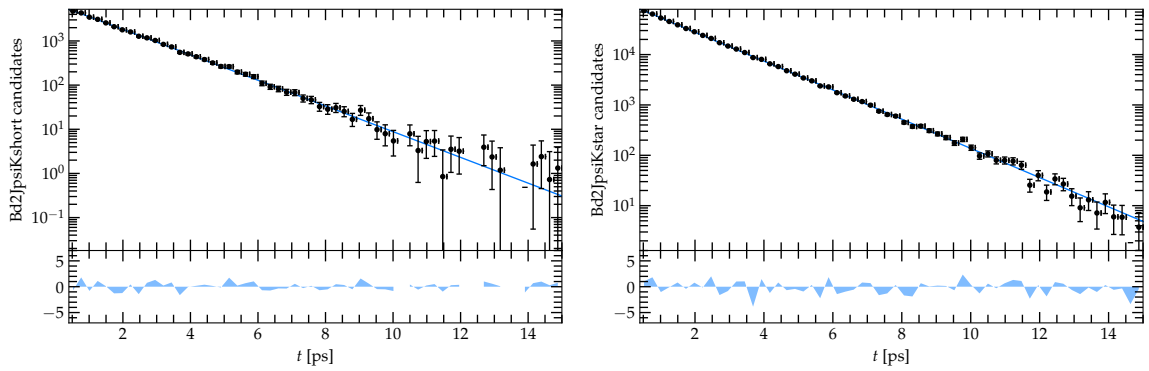

Figure 6.22: Decay-time distribution of background subtracted signal samples of $B^0_d \to J/\psi K^0_s$ on the right and $B^0_d \to J/\psi K^{*0}$ on the right, with the fit projections resulting from the simultaneous fit superimposed.

times the ratio between the measured acceptances of the two modes

$$P(t) = \frac{e^{-t/\tau_{eff}^{K^0_s}} \otimes G(t|\sigma_t^{K^0_s})}{e^{-t/\tau_{eff}^{K^{*0}}} \otimes G(t|\sigma_t^{K^{*0}})} \times \frac{acc(t)^{K^0_s}}{acc(t)^{K^{*0}}}. \tag{6.6}$$

Even though this approach has the advantage to rely on the simulated samples only for the very small correction introduced by the ratio of the acceptances, the $\Delta\Gamma_d$ measured value has been found to strongly depend on the choice of the binning scheme, as it can be seen in Fig. 6.20, and it shows always a positive bias in the obtained $\Delta\Gamma_d$ value.
Thus, a different approach has been used. Considering that the total bias introduced by the reconstruction and selection procedure is small, the acceptances extracted from the corrected and truth-matched simulated samples can be directly used to correct the decay-time distribution of each channel. Following this method, a second fitter has been developed. It is a simultaneous unbinned maximum likelihood fit of the two decay-time distributions, where the $\Delta\Gamma_d$ parameter is in common in the definition of the probability density functions for the two channels. Each function is defined as in the following

$$P(t) = e^{-t/\tau_{eff}} \otimes G(t|\sigma_t) \times acc(t), \tag{6.7}$$

where the σ_t and the acceptance coefficients are fixed and vary accordingly to the considered channel. Since the fit is made simultaneously, it is still sensitive to the differences between the two channels, so mainly on the difference between the two acceptances and on the different dependence of the effective lifetime from $\Delta\Gamma_d$. The simultaneous fit is validated on simulated samples and the results are shown in Fig. 6.21, on top for the $B^0_d \to J/\psi K^0_s$ channel and on the bottom for the $B^0_d \to J/\psi K^{*0}$ one. The obtained value is $\Delta\Gamma_d = -0.0012 \pm 0.0022$ ps^{-1} and it is compatible with zero, giving a first evidence of the validity of the fitter.
A preliminary measurement of $\Delta\Gamma_d$ in the background subtracted 2016 dataset is then performed,

using the simultaneous approach discussed so far. The result of the fit is shown in Fig. 6.22. Since it is just a preliminary result, the central value of the measurement is blinded, in order to perform the analysis without any bias, thus only the precision of the measurement is reported, $\sigma_{stat}(\Delta\Gamma_d) = 0.008$ ps^{-1}. From this quantity, we can estimate the expected precision reachable with the full Run 2 dataset, corresponding to the data recorded from the 2015 to the 2018 by the LHCb detector, simply scaling the precision achieved in 2016 by the ratio between the squared root of the full Run 2 and 2016 only luminosities. The expected precision of the measurement is 0.004 ps^{-1} and a very significant improvement with respect to the current combined value, $\Delta\Gamma_d^{current} = 0.0007 \pm 0.0066$ ps^{-1} [2], can be reached, considering that the former results from a single measurement and not from a world average. From preliminary studies, the systematic uncertainties seem not limiting the precision of the measurement.

Concluding remarks

This thesis presented the ongoing measurements of *CP* violating and mixing parameters in the decay channels $B_s^0 \to J/\psi\phi$ and $B_s^0 \to J/\psi\pi^+\pi^-$, namely the determination of the *CP* violating phase ϕ_s, of the decay width difference $\Delta\Gamma_s$ and of the decay widths Γ_s and Γ_H. A flavour-tagged time- and angular-dependent analysis of the corresponding decay rates has been carried out, using proton-proton collision data collected by the LHCb experiment from 2015 to 2018 during the so-called Run 2 and corresponding to an integrated luminosity of 5.9 fb^{-1}. These analyses are among the flagship results of the LHCb Collaboration and being rather complex they involve many scientists and requires careful checks. Moreover, the main physics results are kept blind until the end of the analysis in order to avoid the possibility of biasing the results toward previously known measurements or predictions. In this thesis, the full strategy developed has been reviewed, with a particular focus on the parts of the analyses that have been carried out by the author of this thesis. For the $B_s^0 \to J/\psi\phi$ decay mode, the full analysis is presented together with the achievable statistical uncertainty using the Run 2 data. In particular, the precision reached on the ϕ_s measurement is about 0.022 rad and thus, once concluded, it will be the most precise single measurement of ϕ_s in the world, with the potential to finally observe *CP* violation in the interference between a decay with and without mixing. However, its SM prediction, taken as the indirect determination of $-2\beta_s$ via a global fit to experimental data within the SM and corresponding to $-2\beta_s = -0.0370^{+0.0007}_{-0.0008}$ rad [3], will still remain more precise than the experimental value, thus highlighting the necessity to perform again this measurement using future Runs of the LHC. Indeed, at the end of the Run 5 period of data acquisition of LHC, a dataset corresponding to a total integrated luminosity of about 300 fb^{-1} will be available, allowing to reach a statistical precision on the ϕ_s measurement of 4 mrad from $B_s^0 \to J/\psi\phi$ only [125], as reported in Fig. 6.23. Nevertheless, the current experimental precision on ϕ_s achievable in Run 2 clearly underline the importance of more precise estimations of the impact of higher order penguin diagrams, as explained in Sec. 1.3.3.

This thesis, additionally, presented the ongoing measurement of the mixing parameter $\Delta\Gamma_d$ of the B_d^0 meson system, that will be performed exploiting the decay channels $B_d^0 \to J/\psi K^*(892)^0$ and $B_d^0 \to J/\psi K_s^0$, using proton-proton collision data collected by the LHCb experiment from 2015 to 2018. Here, the preliminary results obtained using 2016 data have been presented. A new strategy with respect to the previous measurement [1] performed using only 2011 data has

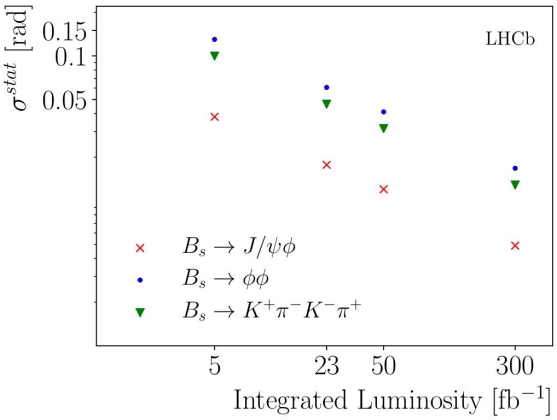

Figure 6.23: Statistical uncertainty on ϕ_s as a function of the amount of data collected by LHCb for different B_s^0 decay modes, represented by different coloured solid lines as shown in the legend. The Standard Model prediction is indicated by the dashed line. The combined luminosity of Run 1 and Run 2 corresponds to 9 fb^{-1}, and a total of 300 fb^{-1} is expected after Run 5. The current results are scaled using the expected running conditions and assuming current detector and flavour tagging performances. Figure taken from Ref. [125].

been fully developed and tested, showing that a good control of the different distorting effects applied to the decay rates by the track reconstruction and the selection procedure is possible. A preliminary measurement of $\Delta\Gamma_d$ in 2016 data showed a statistical precision of 0.008 ps^{-1} and an achievable precision with full Run 2 dataset of 0.004 ps^{-1}. This measurement represents the second determination of $\Delta\Gamma_d$ performed by the LHCb collaboration and the final result will supersede the precision reached in the previous LHCb analysis [1] and of the current world average value, $\Delta\Gamma_d = 0.0007 \pm 0.0066$ ps^{-1} [2], representing the major contribution to the future world average combination.

APPENDIX A

Analysis tools

In this chapter some tools and techniques used during the data analysis are discussed, such as the description of the two most important probability density functions used in the analyses (Crystal Ball and Ipatia), the parameters estimation of probability density functions using the Maximum Likelihood (ML) method, the statistical background subtraction using the unfolding method called *sPlot* technique, and finally the multivariate methods to classify signal and background.

A.1 Crystal Ball and Ipatia functions

In order to properly describe the mass peaks commonly observed in high energy physics experiments, often a single Gaussian is not sufficient, because of the non-Gaussian mass resolution of the detector or because of non-Gaussian physical processes, such as the radiative energy loss in the decay.

A simple PDF is the so-called Crystal Ball (single-side) [117],

$$\mathcal{CB}(x) = \mathcal{N} \begin{cases} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) & \text{if } \frac{x-\mu}{\sigma} > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{\alpha^2}{2}\right) \cdot \left(\frac{n-\alpha^2}{|\alpha|} - \frac{x-\mu}{\sigma}\right) & \text{if } \frac{x-\mu}{\sigma} \leq -\alpha. \end{cases} \tag{A.1}$$

that is a function with a Gaussian core, with mean μ and width σ, and an exponential tail at low values in order to take into account the radiative energy loss of the final state particles. The parameter α determines the point in which the exponential tail begins, while the parameter n quantifies the slope at that point. The function and its first derivative are both continuous at the threshold point. In the equation, $\mathcal{N}$ represents a normalisation factor.

A generalisation of the single-sided Crystal Ball is the so-called Ipatia, which allows to obtain an excellent description of mass resolution with non-Gaussian tails. The single-side Ipatia has the following form [116]

$$I(x;\,\mu,\sigma,\lambda,\xi,\beta,a,n) \propto \begin{cases} A/(B+x-\mu) & \text{if } x-\mu < -a\sigma, \\ ((x-\mu)^2+\delta^2)^{\frac{1}{2}\lambda-\frac{1}{4}}\, e^{\beta(x-\mu)}\, K_{\lambda-\frac{1}{2}}(\alpha\sqrt{(x-\mu)^2+\delta^2}) & \text{otherwise.} \end{cases}$$

where K_ν is the modified Bessel function of the second type, the parameter δ and α are defined as

$$\delta = \sigma\sqrt{\xi K_\lambda(\xi)/K_{\lambda+1}(\xi)},$$
$$\alpha = \sigma\sqrt{\xi K_{\lambda+1}(\xi)/K_\lambda(\xi)}/\sigma,$$

and the parameters A and B are obtained by imposing the continuity of the function itself and its first derivative at the connection point, i.e. $x = \mu - a\sigma$. The parameter μ describes the most probable value of the distribution core and σ describes the mass resolution. Similarly to the $\mathcal{CB}$ function, the left tail describes the radiative energy loss or badly reconstructed events, but in this case the non-Gaussian core is able to describe higher order resolution effects, for instance resolution where the variance varies event-by-event. The parameters ξ and β are generally fixed to zero.

Both the Crystal Ball and Ipatia functions can be generalised to their double-side versions, just by introducing the additional tail parameters for the right tail and by applying the continuity conditions at the connection point.

A.2 Maximum Likelihood Methods

The ML method is commonly used in high-energy physics in order to perform parameter estimation of the probability density functions, which describe distributions of observed experimental data. Since in most cases the parameters are directly related to physical quantities (mass, decay time, CP-violating phases, amplitudes etc.) the ML method is then a powerful method to perform the physical measurements of such quantities.

Consider a set of N measurements of a generic observable x, indicated as $\vec{x} = (x_1, x_2, ..., x_N)$, and the corresponding probability density function describing the probability of the experimental outcome of x, $f(x, \vec{\theta})$. We indicate with $\vec{\theta} = (\theta_1, \theta_2, ..., \theta_M)$ a set of M parameters which the probability function depends on. Then the probability of the N observations can be written as the product of each single probability, following the compound probability theorem, as

$$P(\vec{x}|\vec{\theta}) = \prod_i^N f(x_i, \vec{\theta}). \tag{A.2}$$

where the probability density function is normalized in the whole domain D,

$$\int_D f(x,\vec{\theta})dx = 1. \tag{A.3}$$

In the case of unbinned datasets, each event i enters in the compound probability expression; while if the datasets is binned, where the events are grouped in bins of the x observable, the number of events in each bin enters in the expression. In the following discussion we consider unbinned datasets, but the same conclusions are valid for the binned case.

The value of the compound probability depends on the specific set of parameters $\vec{\theta}$. In the case the parameters are variable and unknown, i.e. their values need to be estimated, the compound probability is called Likelihood function, $\mathcal{L}(\vec{\theta})$. The ML method consist to consider the best estimators of the parameters, $\vec{\theta}_{\rm best}$, if the corresponding Likelihood $\mathcal{L}(\vec{\theta}_{\rm best})$ has its global maximum.

In order to perform the computation of the global maximum in a more convenient way, the negative logarithm of the likelihood function is used, because the expression becomes a sum of terms and the best estimators can be obtained by minimising the expression, thus allowing to use common minimisation software tools,

$$-\ln\mathcal{L}(\vec{\theta}) = -\sum_i^N \ln f(x_i,\vec{\theta}). \tag{A.4}$$

In the asymptotic limit of very large number of observations, $N \to \infty$, the estimated parameters $\vec{\theta}_{\rm best}$ are unbiased, i.e. they converge to the true values $\vec{\theta}_{\rm true}$, and reach the minimum variance. Moreover, it can be demonstrated that, in the asymptotic limit and for any pdf $f(\vec{x},\vec{\theta})$, the Likelihood converges to a multivariate Gaussian distribution of the parameters $\vec{\theta}$,

$$\mathcal{L} \sim e^{-\frac{1}{2}(\vec{\theta}-\vec{\theta}_{\rm best})^{\intercal}\,\mathcal{H}\,(\vec{\theta}-\vec{\theta}_{\rm best})}. \tag{A.5}$$

where $\mathcal{H}$ is a $M \times M$ matrix describing the width of the multi-dimensional Gaussian, thus it is related to the covariance matrix of the parameters, $\mathcal{V}_{ij}(\vec{\theta}_{\rm best})$,

$$\mathcal{V}_{ij}(\vec{\theta}_{\rm best}) = \mathcal{H}^{-1} = \left[-\left.\frac{\partial^2\mathcal{L}}{\partial\theta_i\partial\theta_k}\right|_{\vec{\theta}=\vec{\theta}_{\rm best}} \right]^{-1}. \tag{A.6}$$

In the case of very low statistics, the ML method leads to biased parameter estimators, thus other methods should be preferred.

Weighted Maximum Likelihood The original dataset $\vec{x}$ can be weighted, i.e. a weight w_i is associated to each event, for example in the case of the reweighting procedure to equalise the kinematical distributions of two decay modes, or in the case of acceptance correction, etc. The ML method must take into account the weights properly.

In literature there are different definitions of a weighted Likelihood, depending on the particular application. In high energy physics the commonly used definition is the following, a slightly different expression of Eq. A.2,

$$\mathcal{L}(\vec{\theta}) = \prod_i^N f(x_i, \vec{\theta})^{w_i}, \tag{A.7}$$

where w_i is the weight of the i-th event. Since the probability density function of each event enters in the expression in a multiplicative way, each term is raised to the corresponding weight. With this definition, the expression to calculate the covariance matrix of the parameters, Eq. A.6, is no longer valid. The correct way to calculate the covariance matrix, taking into account the proper scaling of the uncertainties in the presence of weights, is given by the following expression,

$$\mathcal{V}_{ij}(\vec{\theta}_{\text{best}}) = \mathcal{H}^{-1}\mathcal{F}\mathcal{H}^{-1}, \tag{A.8}$$

where $\mathcal{H}$ and $\mathcal{F}$ are matrices which contain the weight information. In particular $\mathcal{H}$ depends on w_i, while $\mathcal{F}$ depends on w_i^2. In this way the resulting covariance matrix elements are properly scaled by the factor $\sum_i w_i^2 / \sum_i w_i$, which quantifies the reduction in the statical power due to the weights. The explicit expression of the matrices are

$$\mathcal{H}_{ij} = -\frac{\partial^2 \sum_k w_k \ln f(x_k, \vec{\theta})}{\partial\theta_i \partial\theta_j}\bigg|_{\vec{\theta}=\vec{\theta}_{\text{best}}}, \tag{A.9}$$

$$\mathcal{F}_{ij} = -\frac{\partial^2 \sum_k w_k^2 \ln f(x_k, \vec{\theta})}{\partial\theta_i \partial\theta_j}\bigg|_{\vec{\theta}=\vec{\theta}_{\text{best}}}. \tag{A.10}$$

Extended Maximum Likelihood The ML method described above is used to obtain the best estimators of the parameters $\vec{\theta}$, where the number N of observed events is fixed. However, N is not a good estimator of the number of expected events, since N is itself a random variable distributed according to a Poisson distribution with expected value $\mathcal{N}$. In order to obtain the best estimate of $\mathcal{N}$ the ML method must be slightly modified by relaxing the normalisation condition in Eq. A.3, in particular

$$\int_D f(x, \vec{\theta}) dx = \mathcal{N}(\theta). \tag{A.11}$$

The Likelihood function can be then extended in order to describe also the fluctuation of the number of events, by multiplying the expression with the Poisson term,

$$\mathcal{L}(\vec{\theta},\mathcal{N}) = e^{-\mathcal{N}}\frac{\mathcal{N}^N}{N!}\prod_i^N f(x_i,\vec{\theta}). \tag{A.12}$$

The parameter $\mathcal{N}$ can be then estimated like any other parameter θ.

A.3 Multivariate Analysis

In order to retain the signal candidates within a dataset and to reject the background components, the analyst must perform selections and range restrictions on the variables used to describe the physical event (masses, momenta, angles, etc.), in order to identify confidence regions in which the probability to find signal candidates is highest. Typically this selection can be done independently for each variable (a method commonly called "rectangular cut"), with the advantage of clear physical meaning and simple systematic uncertainties determination, but with the drawback of neglecting all the information present in the variable correlations. Multivariate analysis methods are used instead to perform event selection in a multidimensional approach, in order to exploit variable correlations and maximise signal efficiency and background rejection.

Boosted Decision Tree One of the main multivariate method used in this thesis is the Boosted Decision Tree (BDT), which allows to classify the events in two categories (signal and background in this context) by using a tree of selections on single variables per node. In each tree node the selection is actually a rectangular cut, however the different selection paths of the tree lead to a series of several sub-selections in the multidimensional parameter space. The Decision Tree is then able to divide the parameter space into several hyper-boxes, therefore adaptable to the distribution of the signal sample, unlike the rectangular cuts in which only one hyper-box can be defined.

To define the decision tree selections, a procedure called *training* must be followed: two datasets, corresponding to a pure signal and pure background components, are passed to the BDT training algorithm, which optimises all the cuts and tree selection paths, since the algorithm knows a priori which of the two categories the events belong to. A second procedure called *testing* is used to test the BDT performance: two datasets of pure signal and pure background components (different and statistical independent with respect to the ones used in the training procedure) are mixed together and the actual BDT classifier is applied on the full dataset, in order to evaluate the classification performance afterwards.

If the BDT is trained with few degrees of freedom, i.e. too many parameters with respect to the number of events, a problem called *over-training* can arise. In this case the BDT adapts the selections cuts in order to follow excessively well the specific signal distributions used in

the training procedure, thus leading to worse behaviour when the same selections are applied to an independent dataset. The BDT method highly suffers from this problem with respect to other multivariate methods, since the decision paths can be changed with high adaptability. The over-training can be studied during the testing procedure, by comparing the classification performance when using the same dataset used during the training. A way to take the over-training under control is to reduce the number of degrees of freedom of the BDT algorithm, such as the number of trees, the depth of the trees, etc.

BDT Training In this paragraph the algorithm to build the decision tree is discussed in more detail. Let us consider a sample of N events, each depending on a set of variables x_j, for example kinematical or geometrical variables describing the event. If the events can be classified in two categories only (signal and background), it follows that $N = N_{sig} + N_{bkg}$, where N_{sig} is the number of signal events and N_{bkg} the number of background events. In the whole sample, or in any sub-sample, there is a mix of signal and background and the purity of the sample can be defined as

$$P = \frac{N_{sig}}{N}, \tag{A.13}$$

which goes from 0 to 1, in the case of a pure background or pure signal sample, respectively. If each event has a weight w_i, the weighted purity can be defined in the following way,

$$P_w = \frac{\sum_{i,sig} w_{i,sig}}{\sum_{i,sig} w_{i,sig} + \sum_{j,bkg} w_{j,bkg}}, \tag{A.14}$$

where the sums are calculated only on signal or background events. Let us consider now the first step of the training algorithm: starting from the root node (the original sample) the algorithm scans all the possible cuts on all the variables defined by the user, with a given scan granularity. Each cut allows to identify two sub-samples ($x_j > \alpha$ and $x_j \leq \alpha$); in the original sample and in the two sub-samples a quantity called Gini index is calculated,

$$I_G = \sum_{i=1}^{n} w_i P_w (1 - P_w), \tag{A.15}$$

where n is the number of events in the sample considered and P_w the weighted purity. The Gini index is 0 if the sample is completely pure (either signal or background) while it reaches its maximum value if the same number of signal and background events are present. The variable x_j and the corresponding cut α are chosen by the algorithm when the separation induced by the cut is maximum, that means when the difference of Gini indexes of the parent sample and the two daughter samples, $I_G^{parent} - I_G^{daughter1} - I_G^{daughter2}$, is maximised. The same procedure is then applied to each sub-samples, and so on. The procedure stops if one the sub-samples reaches a purity close to 0 or 1, depending on a given configurable threshold. All the events in the final samples are tagged as signal or background. Since the threshold must be set in order to not get

over-training, a certain number of events can be mis-classified, i.e. signal events can be tagged as background and vice-versa. When the training finished and the selection cut is defined, the tree itself can be used to classify new events: depending on the specific values of their variables these events will follow a decision tree path and will fall in one signal or background final node.

BDT Boosting The training procedure described above allows to build a single decision tree, but it has been demonstrated that multiple decision trees can increase drastically the classification performance. However the same algorithm cannot be executed multiple times, since it will lead to exactly the same decision tree. Thus, a procedure called *boosting* is implemented: this procedure is the core of the BDT and it allows the algorithm to learn from the mis-classification of the events.

Let us suppose to have already created M decision trees, then we can introduce three new variables, for each event:

- a two-values variable y_i which flags the true event category: $+1$ if the event is signal and -1 if it is background,
- a two-values variable t_{mi} which flags the classification: $+1$ if in the m-th tree the event is classified as signal, -1 as background.
- a binary variable I_{mi} which flags the misclassification: if $y_i = t_{mi}$ the event is classified correctly and $I_{mi} = 0$, otherwise $I_{mi} = 1$.

Thus, it is possible to quantify how much the m-th tree is globally subject to wrong classification. In the AdaBoost algorithm, that is one of the simplest boosting algorithm available, the weighted average of the I_{mi} variables is used,

$$\epsilon_m = \frac{\sum_{i=1}^{N} w_i I_{mi}}{\sum_{i=1}^{N} w_i}. \tag{A.16}$$

A new set of event weights, w_i', can be chosen in order to take into account either the average error of the m-th tree and the single i-th event mis-classification,

$$w_i' = w_i \, e^{\beta \, I_{mi}} \, \frac{1 - \epsilon_m}{\epsilon_m}, \tag{A.17}$$

where β is a constant coefficient depending on the particular algorithm used. Thus the BDT is able to learn how to better classify the events in the two categories because, through the boosting procedure, a bigger weight is given to events wrongly classified in the previous trees. With the different weights w_i' a new training procedure will produce in general a different decision tree with respect to the one obtained with the old weights, thus it will contain new information that can be used for the overall classification. For each event i-th the final decision

over M decision trees can be taken as a majority vote of all the trees, or, in more quantitative way, as a weighted average of the values t_{mi},

$$BDT_i = \sum_{m=1}^{M} \gamma_m t_{mi}, \tag{A.18}$$

where γ_m is a global weight associated to each tree. In the AdaBoost algorithm it can be expressed as $\gamma_m = \beta \ln(\frac{1-\epsilon_m}{\epsilon_m})$.

A.4 *sPlot* technique

Even if the multivariate methods are powerful tools to classify the candidates and perform a very efficient background rejection, often they are not sufficient to obtain the required dataset purity. Another technique, used in this thesis, to unfold signal and background components, in particular where their probability density functions are not known, is called *sPlot* technique.

Let us consider, for simplicity, a dataset with N events described by two observables, the invariant mass m (discriminating variable) and decay time τ (control variable), and let us suppose that only two physical components are present (signal and background). The aim of the study is to obtain the decay time distribution for the signal only component, thus to subtract the contribution of background. If the probability density functions of signal and background are know only in the discriminating variable m (for example by performing a ML fit procedure) the *sPlot* technique can be used to extract per-candidate weights, called s-Weights, ${}_sW$, for each component, that are related to the probability that an event belongs to the signal or background category. By reweighting the dataset using the signal weights, ${}_sW_s$, the corresponding reweighted decay time distribution matches the signal decay time distribution.

In a more general way, for a specific component n, the s-Weight for the candidate i can be calculated as

$${}_sW_n(m_i) = \frac{\sum_{j=1}^{M} \mathcal{V}_{nj}\, f_j(m_i)}{\sum_{k=1}^{M} N_k\, f_j(m_i)}, \tag{A.19}$$

where M is the number of categories (2 in the example discussed above), N_k is the yield of the k-th component, $\mathcal{V}$ is the covariance matrix obtained by the ML fit for the yields only, f_j are the probability density functions of the components, and m_i is the discriminating variable (the invariant mass in the example discussed above).

Intuitively, when the ML fit is performed, in each small region of the discriminating variable m, there is a fixed proportions of the candidate components and the events can be weighted according to these proportions. If, for example, the region is dominated only by the background, then all the events in that region will have higher background s-Weights and low signal s-Weights, according to the proportions of the yields.

For a specific component n, the sum of the corresponding s-Weights must be equal to the total yield of that component, $\mathcal{N}_n$,

$$\sum_i sW_n(m_i) = \mathcal{N}_n, \tag{A.20}$$

while, for each event, the sum of the weights over the components must be equal to 1. An important assumption, ensuring the correct statistical subtraction, is that the discriminating and control variables must be uncorrelated, within each component, in order that all the remaining correlations can be associated only to the proportions of the different components.

APPENDIX B

Supplementary material for $\Delta\Gamma_d$ analysis

B.1 Legend of the lifetime scan as a function of the analysis steps

In this Appendix the different analysis steps used in Fig. 6.16 are explained. From step 4, all the following steps contain also the previous requirements.

- 0. All particles within LHCb geometrical acceptance.
- 1. Reconstruction of the final state hadrons (MC generated sample). All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 2. Reconstruction of the final state muon (MC generated sample)s. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 3. Reconstruction of all the final state particles (MC generated sample). All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 4. Reconstruction of all the final state particles (MC reconstructed sample). All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 5. Stripping selection, with biasing requirement of J/ψ $|FDS| > 3$. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.

- 6. `HLT1DiMuonHighMass(TOS)` selection. All final state particles: $2 < \eta < 5$, $|z(PV)| <$ 100 mm.
- 7. `HLT2DiMuonDetachedJPsi(TOS)` selection, with biasing requirement of J/ψ $FDS > 3$. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 8. track $\chi^2/nDoF < 4$ for all the hadrons in the final state and π^- $PIDk < 0$. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 9. track $\chi^2/nDoF < 4$ for all the muons in the final state. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 10. Transverse momentum of the K^0_s/K^{*0} greater than 1200 MeV/c. K^0_s invariant mass n the range $(422 - 572)$ MeV/c^2. K^{*0} invariant mass n the range $(826 - 966)$ MeV/c^2. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 11. J/ψ decay vertex $\chi^2/nDoF < 16$. J/ψ invariant mass in the range (3030−3150) MeV/c^2. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 12. B^0_d invariant mass in the range $(5200 - 5350)$ MeV/c^2. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 13. B^0_d $\chi^2_{DTF}/nDoF$ $(shifted) < 5$. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100mm$.
- 14. B^0_d $\chi^2(IP) < 25$. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.
- 15. B^0_d $\chi^2(IP)_{\text{next best}} > 50$ or $== -1$. All final state particles: $2 < \eta < 5$, $|z(PV)| <$ 100 mm.
- 16. B^0_d BDT selection. All final state particles: $2 < \eta < 5$, $|z(PV)| < 100$ mm.

List of Tables

List of Figures

Figure **Page**

References

[1] F. Dordei, *Lifetime measurements of beauty hadrons at the LHCb experiment*, PhD thesis, Heidelberg U., 11, 2014, doi: 10.11588/heidok.00018801. ↪ cited on pages i, 12, 19, 58, 72, 185, 186, 200, 201, 207, 208, 223, 224, 225, 231

[2] HFLAV, Y. S. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*, Eur. Phys. J. **C81** (2021) 226, arXiv:1909.12524, updated results and plots available at https://hflav.web.cern.ch/. ↪ cited on pages i, iv, 70, 71, 86, 136, 183, 204, 206, 208, 225, 227

[3] J. Charles *et al.*, *Predictions of selected flavour observables within the Standard Model (update with http://ckmfitter.in2p3.fr/)*, Phys. Rev. D **84** (2011) 033005, arXiv:1106.4041. ↪ cited on pages iv, 18, 71, 207

[4] A. Lenz and G. Tetlalmatzi-Xolocotzi, *Model-independent bounds on new physics effects in non-leptonic tree-level decays of b-mesons*, Journal of High Energy Physics **2020** (2020) . ↪ cited on pages iv, 28, 71

[5] LHCb Collaboration, R. Aaij *et al.*, *Updated measurement of time-dependent CP-violating observables in $B_s^0 \to J/\psi K^+K^-$ decays*, Eur. Phys. J. C **79** (2019) 706. 42 p, arXiv:1906.08356. ↪ cited on pages v, 70, 130, 131

[6] R. Aaij *et al.*, *Measurement of the cp-violating phase ϕ from $b_s^0 \to j/\psi\pi^+\pi^-$ decays in 13 tev pp collisions*, Physics Letters B **797** (2019) 134789. ↪ cited on pages v, 23, 70, 136

[7] *Measurement of $B_s^0 \to J/\psi K^+K^-$ with the full Run 2 data*, , Once ready, it will be available at https://cds.cern.ch/collection/LHCb ↪ cited on pages v, 69

[8] *CP violation in $B_s^0 \to J/\psi\pi^+\pi^-$ decays with Run 2 data*, , Once ready, it will be available at https://cds.cern.ch/collection/LHCb ↪ cited on pages v, 135

[9] E. Noether, *Invariant variation problems*, Transport Theory and Statistical Physics **1** (1971) 186–207. ↪ cited on page 1

[10] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508. ↪ cited on page 2

[11] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321. ↪ [see prev. ref.]

[12] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585. ↪ cited on page 2

[13] P. D. Group *et al.*, *Review of Particle Physics*, Pro-

gress of Theoretical and Experimental Physics **2020** (2020) , arXiv:https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf, 083C01. ↪ cited on pages 3, 7, 16, 23, 61, 72, 74, 75, 77, 80, 81, 83, 92, 93, 97, 120, 140, 158, 167, 174, 184, 186, 189, 193, 194,

[14] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413. ↪ cited on page 4

[15] T. D. Lee and C. N. Yang, *Question of parity conservation in weak interactions*, Phys. Rev. **104** (1956) 254. ↪ cited on page 4

[16] R. L. Garwin, L. M. Lederman, and M. Weinrich, *Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon*, Phys. Rev. **105** (1957) 1415. ↪ cited on page 4

[17] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *Decay of the* K_2^0 *Meson*, Phys. Rev. Lett. **13** (1964) 138. ↪ cited on page 5

[18] S. Descotes-Genon and P. Koppenburg, *The ckm parameters*, Annual Review of Nuclear and Particle Science **67** (2017) 97–127. ↪ cited on page 7

[19] J. Charles *et al.*, *CP violation and the CKM matrix: assessing the impact of the asymmetric B factories*, The European Physical Journal C - Particles and Fields **41** (2005) 1. ↪ [see prev. ref.]

[20] G. Eigen, G. Dubois-Felsmann, D. G. Hitlin, and F. C. Porter, *Global CKM Fits with the Scan Method*, Phys. Rev. **D89** (2014), no. 3 033004, arXiv:1301.5867. ↪ [see prev. ref.]

[21] A. Hocker, H. Lacker, S. Laplace, and F. Le Diberder, *A New approach to a global fit of the CKM matrix*, Eur. Phys. J. **C21** (2001) 225, arXiv:hep-ph/0104062. ↪ [see prev. ref.]

[22] M. Ciuchini *et al.*, *2000 CKM triangle analysis: A Critical review with updated experimental inputs and theoretical parameters*, JHEP **07** (2001) 013, arXiv:hep-ph/0012308. ↪ cited on page 7

[23] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945. ↪ cited on page 7

[24] R. Aaij *et al.*, *Measurement of the semileptonic cp asymmetry in* b^0 - $\overline{B}^0$ *mixing*, Physical Review Letters **114** (2015) . ↪ cited on page 15

[25] R. Aaij *et al.*, *Measurement of the cp asymmetry in* b_s^0 - $\overline{B}_s^0$ *mixing*, Physical Review Letters **117** (2016) . ↪ cited on page 15

[26] L. Zhang and S. Stone, *Time-dependent dalitz-plot formalism for* $b_q \to j/\psi h^+ h^-$, Physics Letters B **719** (2013) 383–387. ↪ cited on page 22

[27] D. Brundu, *Radiation hardness of the upgraded LHCb muon detector electronics and prospects for a full angular analysis in multi-body rare charm decays*, Jan, 2020. Presented 04 Feb 2020. ↪ cited on page 23

[28] R. Aaij *et al.*, *Measurement of resonant and cp components in* $b_s^0 \to j/\psi\pi^+\pi^-$ *decays*, Physical Review D **89** (2014) . ↪ cited on page 23

[29] BES collaboration, M. Ablikim *et al.*, *Resonances in* $J/\psi \to \phi\pi^+\pi^-$ *and* $\phi K^+ K^-$, Phys.

Lett. **B607** (2005) 243, arXiv:hep-ex/0411001. ↪ cited on page 23

[30] LHCb collaboration, R. Aaij *et al.*, *Resonances and CP-violation in $\overline{B}{}^0_s$ and $B^0_s \to J/\psi K^+K^-$ decays in the mass region above the $\phi(1020)$*, JHEP **08** (2017) 037, arXiv:1704.08217. ↪ cited on page 23

[31] R. Aaij *et al.*, *Analysis of the resonant components in $\overline{B}{}^0_s$ to$j/\psi\pi^+\pi^-$*, Physical Review D **86** (2012) . ↪ cited on pages 25, 135

[32] K. De Bruyn and R. Fleischer, *A roadmap to control penguin effects in $b^0_d \to j/\psi k^0_s$ and $b^0_s \to j/\psi\phi$*, Journal of High Energy Physics **2015** (2015) . ↪ cited on page 25

[33] S. Faller, R. Fleischer, and T. Mannel, *Precision physics with $b^0_s \to j/\psi\phi$ at the lhc: The quest for new physics*, Physical Review D **79** (2009) . ↪ cited on page 25

[34] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation parameters and polarisation fractions in $B^0_s \to J/\psi\overline{K}{}^{*0}$ decays*, JHEP **11** (2015) 082. 39 p, arXiv:1509.00400, Comments: 39 pages, 7 tables, 8 figures. ↪ cited on page 26

[35] R. Aaij *et al.*, *Measurement of the cp -violating phase beta in $b^0 \to j/\psi\pi^+\pi^-$ decays and limits on penguin effects*, Physics Letters B **742** (2015) 38–49. ↪ cited on page 26

[36] R. Fleischer and R. Knegjens, *Effective lifetimes of b_s decays and their constraints on the b^0_s–$\bar{B}^0_s$ mixing parameters*, The European Physical Journal C **71** (2011) . ↪ cited on page 29

[37] T. Gershon, *$\Delta\Gamma_d$: A Forgotten Null Test of the Standard Model*, J. Phys. G **38** (2011) 015007, arXiv:1007.5135. ↪ cited on page 29

[38] J. Charles *et al.*, *Current status of the standard model ckm fit and constraints on $\delta f = 2$ new physics*, Physical Review D **91** (2015) . ↪ cited on pages 31, 32, 223

[39] A. Dighe, A. Kundu, and S. Nandi, *Possibility of large lifetime differences in neutral b meson systems*, Physical Review D **76** (2007) . ↪ cited on page 32

[40] D0 Collaboration, V. M. Abazov and A. et al. *Evidence for an anomalous like-sign dimuon charge asymmetry*, Phys. Rev. D **82** (2010) 032001. ↪ cited on page 32

[41] D0 Collaboration, V. M. Abazov and A. et al. *Evidence for an anomalous like-sign dimuon charge asymmetry*, Phys. Rev. Lett. **105** (2010) 081801. ↪ cited on page 32

[42] The D0 Collaboration, V. M. Abazov and A. et al. *Measurement of the anomalous like-sign dimuon charge asymmetry with* 9 fb^{-1} *of $p\overline{p}$ collisions*, Phys. Rev. D **84** (2011) 052007. ↪ cited on page 32

[43] D0 Collaboration, V. M. Abazov and A. et al. *Study of cp-violating charge asymmetries of single muons and like-sign dimuons in $p\overline{p}$ collisions*, Phys. Rev. D **89** (2014) 012002. ↪ cited on page 32

[44] A. J. Lenz, *Selected topics in heavy flavour physics*, Journal of Physics G: Nuclear and Particle Physics **41** (2014) 103001. ↪ cited on page 32

[45] G. Borissov and B. Hoeneisen, *Understanding the like-sign dimuon charge asymmetry in $p\overline{p}$ collisions*, Physical Review D **87** (2013) . ↪ cited on page 32

[46] A. Lenz and G. Tetlalmatzi-Xolocotzi, *Model-independent bounds on new physics effects*

in non-leptonic tree-level decays of b-mesons, Journal of High Energy Physics **2020** (2020) . ↪ cited on pages 32, 183

[47] A. A. Alves *et al.*, *The LHCb detector at the LHC*, Journal of Instrumentation **3** (2008) S08005. ↪ cited on pages 33, 34

[48] L. Evans and P. Bryant, *LHC machine*, Journal of Instrumentation **3** (2008) S08001. ↪ cited on page 33

[49] C. De Melis, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, OPEN-PHO-ACCEL-2016-001 (2016), General Photo. ↪ cited on pages 34, 223

[50] T. A. Collaboration, *The ATLAS experiment at the CERN large hadron collider*, Journal of Instrumentation **3** (2008) S08003. ↪ cited on page 34

[51] T. C. Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08004. ↪ cited on page 34

[52] T. A. Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08002. ↪ cited on page 34

[53] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820. ↪ cited on pages 36, 223

[54] LHCb collaboration, C. Elsässer, $\bar{b}b$ *production angle plots*, . ↪ cited on pages 36, 223

[55] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions. Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2016) 052002. 11 p, arXiv:1612.05140, This version incorporates the Errata and presents only the updated measurements. Erratum added on Sept. 1, 2017. Please use for 13 TeV cross-sections and cross-section ratios. All figures and tables, along with any supplementary material and additional information, are available at https://lhcbproject.web.cern.ch/lhcbproject/Publications/LHCbProjectPublic/LHCb-PAPER-2016-031.html. ↪ cited on pages 35, 36

[56] Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and*

τ

-lepton properties as of summer 2016, The European Physical Journal C **77** (2017) . ↪ cited on pages 36, 221

[57] H. Burkhardt and P. Grafström, *Absolute Luminosity from Machine Parameters*, Tech. Rep. LHC-PROJECT-Report-1019. CERN-LHC-PROJECT-Report-1019, 2007. ↪ cited on page 37

[58] F. Follin and D. Jacquet, *Implementation and experience with luminosity levelling with offset beam*, in *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*, pp. 183–187, 2014. arXiv:1410.3667. [,183(2014)], doi: 10.5170/CERN-2014-004.183. ↪ cited on page 38

[59] LHCb Collaboration, S. Amato *et al.*, *LHCb magnet: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000. ↪ cited on page 38

[60] LHCb Collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb VELO (VErtex LOcator): Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001. ↪ cited on page 39

[61] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808. ↪ cited on page 40

[62] LHCb Collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb inner tracker: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2002. revised version number 1 submitted on 2002-11-13 14:14:34. ↪ cited on page 40

[63] LHCb Collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb outer tracker: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001. ↪ cited on page 42

[64] LHCb Collaboration, S. Amato *et al.*, *LHCb RICH: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000. ↪ cited on page 43

[65] LHCb RICH Group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759. ↪ cited on pages 43, 44, 224

[66] LHCb Collaboration, S. Amato *et al.*, *LHCb calorimeters: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000. ↪ cited on page 44

[67] LHCb Collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb muon system: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001. ↪ cited on page 46

[68] LHCb Collaboration, *LHCb muon system: addendum to the Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2003. ↪ [see prev. ref.]

[69] LHCb Collaboration, *LHCb muon system: second addendum to the Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2005. Submitted on 9 Apr 2005. ↪ cited on page 46

[70] LHCb Collaboration, *Nu, Mu and Pile-Up. The LHCb definitions of what we see and what we don't see.*, https://twiki.cern.ch/twiki/bin/view/LHCb/NuMuPileUp. ↪ cited on page 46

[71] LHCb, F. Dordei, *LHCb detector and trigger performance in Run II*, EPJ Web Conf. **164** (2017) 01016. 10 p. ↪ cited on page 46

[72] LHCb Collaboration, R. Antunes-Nobrega *et al.*, *LHCb trigger system: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2003. revised version number 1 submitted on 2003-09-24 12:12:22. ↪ cited on page 46

[73] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055. ↪ cited on page 46

[74] LHCb Collaboration, *LHCb Tracking Strategies*, https://twiki.cern.ch/twiki/bin/view/LHCb/LHCbTrackingStrategies. ↪ cited on pages 48, 52

[75] R. Frühwirth, *Application of kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987), no. 2 444 . ↪ cited on pages 48, 53

[76] T. Likhomanenko *et al.*, *LHCb Topological Trigger Reoptimization*, J. Phys. Conf. Ser.

664 (2015), no. 8 082025, arXiv:1510.00572. ↪ cited on page 49

[77] M. Clemencic *et al.*, *Recent developments in the LHCb software framework Gaudi*, J. Phys. Conf. Ser. **219** (2010) 042006. ↪ cited on page 50

[78] LHCb Collaboration, *The Brunel project*, http://cern.ch/lhcbdoc/brunel/. ↪ cited on pages 50, 74

[79] LHCb Collaboration, *The DaVinci project*, http://cern.ch/lhcbdoc/davinci/. ↪ cited on page 50

[80] LHCb Collaboration, *Decay Tree Fitter*, https://twiki.cern.ch/twiki/bin/view/LHCb/DecayTreeFitter. ↪ cited on pages 50, 59

[81] LHCb Collaboration, *The Moore project*, http://cern.ch/lhcbdoc/moore/. ↪ cited on page 50

[82] LHCb Collaboration, *The Gauss project*, http://cern.ch/lhcbdoc/gauss. ↪ cited on page 50

[83] T. Sjöstrand *et al.*, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv:1410.3012. ↪ cited on page 50

[84] A. Ryd *et al.*, *EvtGen: A Monte Carlo Generator for B-Physics*, . ↪ cited on page 50

[85] GEANT4, S. Agostinelli *et al.*, *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250. ↪ cited on page 50

[86] LHCb Collaboration, *The Boole project*, http://cern.ch/lhcbdoc/boole/. ↪ cited on page 50

[87] LHCb Collaboration, *The LHCb Stripping*, https://twiki.cern.ch/twiki/bin/view/LHCb/LHCbStripping. ↪ cited on page 51

[88] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, Tech. Rep. LHCb-PUB-2014-039. CERN-LHCb-PUB-2014-039, CERN, Geneva, May, 2014. ↪ cited on page 51

[89] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the Velo*, tech. rep., CERN, Geneva, Jan, 2011. LHCb. ↪ cited on page 52

[90] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, tech. rep., CERN, Geneva, May, 2007. ↪ cited on page 53

[91] E. Rodrigues, *Dealing with clones in the tracking*, tech. rep., CERN, Geneva, Nov, 2006. ↪ cited on page 54

[92] H. G. Moser and A. Roussarie, *Mathematical methods for B^0 anti-B^0 oscillation analyses*, Nucl. Instrum. Meth. **A384** (1997) 491. ↪ cited on page 59

[93] M. Veronesi, *Quantum oscillations of neutral particles with the beauty quark*, , Presented 15 Nov 2021. ↪ cited on pages 64, 224

[94] M. Veronesi, *Quantum oscillations of neutral particles with the beauty quark*, PhD thesis, Vrije Universiteit Amsterdam, Nov., 2021. ↪ cited on page 64

[95] R. Aaij *et al.*, *A new algorithm for identifying the flavour of b_s^0 mesons at lhcb*, Journal of

Instrumentation **11** (2016) P05010–P05010. ↪ cited on page 65

[96] R. Aaij *et al.*, *Opposite-side flavour tagging of b mesons at the lhcb experiment*, The European Physical Journal C **72** (2012) . ↪ cited on page 65

[97] T. L. collaboration, *B flavour tagging using charm decays at the lhcb experiment*, Journal of Instrumentation **10** (2015) P10005–P10005. ↪ cited on page 65

[98] LHCb, D. Fazzini, *Flavour Tagging in the LHCb experiment*, PoS **LHCP2018** (2018) 230. 7 p. ↪ cited on pages 65, 224

[99] Q. Fuhring, *Flavour Tagging at the LHCb Experiment*, Feb, 2020. ↪ cited on pages 67, 224

[100] LHCb Collaboration, R. Aaij and A. et al. *Precision measurement of cp violation in* $B_s^0 \to j/\Psi K^+K^-$ *decays*, Phys. Rev. Lett. **114** (2015) 041801. ↪ cited on pages 70, 116, 122

[101] R. Aaij *et al.*, *Resonances and cp violation in* b_s^0 *and*

$$\overline{b}_s^0 \to j/\psi K^+K^-$$

decays in the mass region above the $\phi(1020)$, Journal of High Energy Physics **2017** (2017) . ↪ cited on page 70

[102] R. Aaij, B. Adeva, and M. A. et al. *First study of the cp-violating phase and decay-width difference in* b_s^0 *to*$\psi(2s)\phi$ *decays*, Physics Letters B **762** (2016) 253. ↪ cited on page 70

[103] R. A. et al. *Measurement of the cp-violating phase* ϕ_s *in* $\overline{B}_s^0 \to j/\psi\pi^+\pi^-$ *decays*, Physics Letters B **736** (2014) 186. ↪ cited on pages 70, 136

[104] LHCb Collaboration, R. Aaij and A. B. et al. *Measurement of the cp-violating phase* ϕ_s *in* $\overline{b}_s^0 \to D_s^+D_s^-$ *decays*, Phys. Rev. Lett. **113** (2014) 211801. ↪ cited on pages 70, 136

[105] CDF Collaboration, T. Aaltonen and A. G. et al. *Measurement of the bottom-strange meson mixing phase in the full cdf data set*, Phys. Rev. Lett. **109** (2012) 171802. ↪ cited on page 70

[106] The D0 Collaboration, V. M. Abazov and A. et al. *Measurement of the cp-violating phase* $\phi_s^{J/\psi\phi}$ *using the flavor-tagged decay* $B_s^0 \to j/\psi\phi$ *in* 8 fb^{-1} *of* $p\overline{p}$ *collisions*, Phys. Rev. D **85** (2012) 032006. ↪ cited on page 70

[107] ATLAS Collaboration, G. Aad and A. et al. *Flavor tagged time-dependent angular analysis of the* $B_s^0 \to j/\psi\phi$ *decay and extraction of* $\Delta\Gamma_s$ *and the weak phase* ϕ_s *in atlas*, Phys. Rev. D **90** (2014) 052007. ↪ cited on page 70

[108] G. Aad *et al.*, *Measurement of cp-violating phase* ϕ_s *and the* b_s^0 *meson decay width difference with* $b_s^0 \to j/\psi\phi$ *decays at atlas*, Journal of High Energy Physics **2016** (2016) . ↪ cited on page 70

[109] ATLAS, *Measurement of the CP violation phase* ϕ_s *in* $B_s \to J/\psi\phi$ *decays in ATLAS at 13 TeV*, . ↪ cited on page 70

[110] V. K. et al. *Measurement of the cp-violating weak phase* ϕ_s *and the decay width difference* $\delta\gamma_s$ *using the* $b_s^0 \to j/\psi\phi(1020)$ *decay channel in pp collisions at s=8 tev*, Physics Letters B **757** (2016) 97. ↪ cited on page 70

[111] A. M. S. et al. *Measurement of the cp-violating phase* ϕ_s *in the* $b_s^0 \to j/\psi\phi(1020) \to$

$\mu^+\mu^-k^+k^-$ *channel in proton-proton collisions at s=13tev*, Physics Letters B **816** (2021) 136188. ↪ cited on page 70

[112] A. Lenz, *Lifetimes and hqe*, 2019. ↪ cited on page 71

[113] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016. ↪ cited on pages 77, 187

[114] P. Speckmayer, A. Höcker, J. Stelzer, and H. Voss, *The toolkit for multivariate data analysis, TMVA 4*, . ↪ cited on page 78

[115] Y. Xie, *sfit: a method for background subtraction in maximum likelihood fit*, 2009. ↪ cited on page 79

[116] D. Martinez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. A **764** (2014) 150, arXiv:1312.5000. ↪ cited on pages 81, 88, 210

[117] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, PhD thesis, Cracow, INP, 1986. ↪ cited on pages 83, 209

[118] A. Rogozhnikov, *Reweighting with boosted decision trees*, Journal of Physics: Conference Series **762** (2016) 012036. ↪ cited on page 89

[119] A. Bertolin *et al.*, *Precision measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency with the $B_s^0 \to D_s^- \pi^+$ decays*, . ↪ cited on page 115

[120] LHCb Collaboration, R. Aaij and A. B. et al. *Measurement of the polarization amplitudes in $B^0 \to j/\psi K^*(892)^0$ decays*, Phys. Rev. D **88** (2013) 052002. ↪ cited on pages 122, 125, 222

[121] T. du Pree, *Search for a Strange Phase in Beautiful Oscillations*, PhD thesis, 2010, Presented on 22 Oct 2010. ↪ cited on page 124

[122] LHCb Collaboration, R. Aaij and A. B. et al. *Measurement of the $\bar{B}_s^0$ effective lifetime in the $J/\psi f_0(980)$ final state*, Phys. Rev. Lett. **109** (2012) 152002. 14 p, arXiv:1207.0878, Comments: 14 pages, 5 figures. ↪ cited on page 136

[123] LHCb collaboration, R. Aaij and A. B. et al. *Measurement of the $B_s^0 \to J/\psi\eta$ lifetime*, Phys. Lett. B **762** (2016) 484, arXiv:1607.06314, All figures and tables, along with any supplementary material and additional information, are available at https://lhcbproject.web.cern.ch/lhcbproject/Publications/LHCbProjectPublic/LHCb-PAPER-2016-017.html. ↪ cited on page 136

[124] J. Stevens and M. Williams, *uboost: a boosting method for producing uniform selection efficiencies from multivariate classifiers*, Journal of Instrumentation **8** (2013) P12013–P12013. ↪ cited on page 149

[125] R. Aaij and A. B. et al. *Physics case for an lhcb upgrade ii - opportunities in flavour physics, and beyond, in the hl-lhc era*, 2019. ↪ cited on pages 207, 208, 232