Measurement of $\mathcal{CP}$ Observables in $B_s^0 \to J/\psi\phi$ decays at LHCb experiment

Varvara Batozskaya

Supervisor: **dr hab. Krzysztof Kurek, prof. NCBJ**

Supporting supervisor: **dr Konrad Klimaszewski**

National Centre for Nuclear Research

Thesis is submitted for the degree of
Doctor of Philosophy

Warsaw February 2019

To my beloved mother and grandmother,
who raised me

Acknowledgements

I wish to express my sincerely gratitude to my supervisor, Prof. Krzysztof Kurek, for providing me an opportunity to be a PhD student at National Centre for Nuclear Research and to my team leader, Prof. Wojciech Wiślicki, for possibility to be a part of the LHCb Warsaw group and a support of my research work.

I sincerely thank to Dr Konrad Klimaszewski for his guidance and encouragement in carrying out this project work. I very appreciate the time when we worked together as one harmonious mechanism. I am very grateful to LHCb Warsaw group, Dr Wojciech Krzemień, Dr Dmytro Melnichuk, Dr Adam Szabelski and Dr Artur Ukleja, who rendered their help and fruitful discussions during the period of my project work.

I thank my colleagues from the LHCb collaboration for their professional and the great working atmosphere.

Finally, I would like to acknowledge with gratitude, the support and love of my friends and family, especially my grandmother, Galina, and my mother, Valeriya. A special word of gratitude is addressed to my life friend Valery, who never stopped challenging me, and helping develop my ideas. They all kept me going, and this thesis would not have been possible without them.

This work was supported by the National Science Centre (NCN) in Poland under the grant UMO-2015/17/N/ST2/04056.

Abstract

The symmetries and conservation laws have played an important role in the development of the Standard Model (SM). However, not all symmetries are strictly conserved. In particular, the symmetry with respect to the charge-conjugation-parity operator ($\mathcal{CP}$) is not conserved in weak interactions. This leads to differences in behaviour of particles and anti-particles. In the SM the $\mathcal{CP}$ violation arises through a single phase in the Cabibbo-Kobayashi-Moskawa quark mixing matrix, describing the interactions among quarks and the charged carrier of the weak force. In neutral B meson decays to a final state which is accessible to both B and $\bar{B}$ mesons, the interference between the amplitude for the direct decay and the amplitude for decay after oscillation, leads to a time dependent $\mathcal{CP}$-violating asymmetry between the decay time distributions of B and $\bar{B}$ mesons. The decay $B_s^0 \to J/\psi\phi$ allows for a measurement of such an asymmetry, which can be expressed in terms of the decay width difference $\Delta\Gamma_s$ and a single phase ϕ_s.

The objective of this thesis is the measurement of the $\mathcal{CP}$-violating parameters in the $B_s^0 \to J/\psi\phi$ decays with e^+e^- final state for J/ψ meson decay and K^+K^- pair for ϕ meson decay. Using a dataset corresponding to an integrated luminosity of 3.0 fb^{-1} recorded at a center-of-mass energies of $\sqrt{s} = 7$ and 8 TeV collected with the LHCb detector at CERN Large Hadron Collider during 2011 and 2012, a tagged time dependent angular analysis of $B_s^0 \to J/\psi\phi$ decays is performed. The selected sample corresponds to 12195$\pm$497 decay candidates. The $\mathcal{CP}$-violating phase of the B_s^0 system is measured to be $\phi_s = -0.18^{+0.37}_{-0.39}(\text{stat.}) \pm 0.25(\text{syst.})$ rad that is consistent with the SM predictions and measurement results of other channels. This is the first time that the $\mathcal{CP}$-violating parameters have been measured in a decay with electron final state. The result allows to increase the data statistics and to improve the knowledge of a possible systematic effects.

In addition to physics results, the alignment studies of the LHCb tracking stations are presented. The investigations are focused on the variation of the detector positions dependence on magnetic field polarizations and lead to improved alignment of tracking detectors. The study results are discussed in the thesis.

Streszczenie

Symietrie i prawa zachowania odegrały ogromną rolę w rozwoju Modelu Standartowego oddziaływań elementarnych. Jednakże nie wszystkie symetrie są ściśle zachowane. W szczególności symetria względem operatora parzystości przestrzenno-ładunkowej ($\mathcal{CP}$) nie jest zachowana w oddziaływaniach słabych. Prowadzi to do różnic w zachowaniu się cząstek i anty-cząstek. W ramach Modelu Standartowego łamanie symetrii $\mathcal{CP}$ opisane jest przez pojedynczą fazę w macierzy mieszania Cabibbo-Kobayashi-Maskawa opisującej interakcje kwarków z nośnikiem oddziaływań słabych. W rozpadach neutralnych mezonów B, do stanu dostępnego zarówno dla B i $\bar{B}$, zachodzi interferencja pomiędzy amplitudami rozpadów bezpośrednich i rozpadów z mieszaniem. Prowadzi to do zależnej od czasu asymetrii między rozkładami mezonów B oraz $\bar{B}$. Rozpad $B_s^0 \to J/\psi\phi$ pozwala na pomiar takiej asymetrii, która może zostać wyrażona przez różnicę szerokości rozpadów $\Delta\Gamma_s$ oraz pojedynczą fazy ϕ_s.

Celem niniejszej pracy jest pomiar łamania $\mathcal{CP}$ w rozpadach $B_s^0 \to J/\psi\phi$ do stanu końcowego e^+e^- w rozpadzie mezonu J/ψ oraz pary K^+K^- w rozpadzie mezonu ϕ. Do badania wykorzystano dane zabrane przez detektor LHCb przy zderzaczu LHC pochodzące z zderzeń proton-proton przy energii w środku masy 7 i 8 TeV, w latach 2011-2012, odpowiadające scałkowanej świetlności 3.0 fb^{-1}. W wyniku selekcji próbki danych otrzymano 12195±497 kandydatów na rozpad $B_s^0 \to J/\psi\phi$ w kanale elektronowym. Wartości parametrów łamania $\mathcal{CP}$ wyznaczono z dopasowania teoretycznych przewidywań szybkości zaniku dla kanału $B_s^0 \to J/\psi\phi$ do eksperymentalnych rozkładów kątowych oraz czasu życia. Wynik pomiaru fazy ϕ_s w układzie mezonu B_s^0 wynosi $\phi_s = -0.18^{+0.37}_{-0.39}(\text{stat.})\pm 0.25(\text{syst.})$ rad, jest on zgodny z przewidywaniami Modelu Standartowego i wynikami pomiarów z innych kanałów. Otrzymany wynik jest pierwszym pomiarem parametrów łamania $\mathcal{CP}$ w kanale rozpadu z elektronami w stanie końcowym. Rezultat pozwala na zwiększenie dostępnej próbki danych i poprawę wiedzy o możliwych efektach systematycznych.

Poza wynikami fizycznymi, przedstawiono badania współosiowania stacji śladowych eksperymentu LHCb. Badania koncentrują się na zmienności pozycji detektorów w zależności od polaryzacji pola magnetycznego i prowadzą do lepszego pozycjonowania detektorów śladowych. Wyniki badań zostały omówione w niniejszej pracy.

Table of contents

Chapter 1

$\mathcal{CP}$ Violation in Neutral Mesons

The chapter gives an overview of the Standard Model focusing on the quark flavour sector. The CKM mechanism and the phenomenon of $\mathcal{CP}$ violation caused by a non-trivial CKM phase are discussed in details. The mixing of neutral B mesons is introduced and a formalism to describe both mixing and decay of B mesons is presented. The $B_s^0 \to J/\psi\phi$ decay is introduced as an example of time dependent $\mathcal{CP}$ violation. In this decay, the $\mathcal{CP}$-violating phase ϕ_s arises due to the interference between the direct decay and the decay after $B_s^0 - \bar{B}_s^0$ mixing. Finally, the current experimental status of the phase ϕ_s is presented and the effect of possible contributions from processes beyond the Standard Model on ϕ_s is discussed.

1.1 Introduction to the Standard Model

The Standard Model (SM) of the particle physics is a renormalizable field theory describing the fundamental particles and their interactions [1, 2]. The particles that construct the known matter are described by fermion fields with half-integer spin. They can be divided into quarks and leptons and are grouped in three generations as shown in Table 1.1. Each generation contains a lepton with an electric charge (e, μ, τ), one lepton without electric charge, called neutrino $(\nu_e, \nu_\mu, \nu_\tau)$, one up-type quark (u, c, t) and one down-type quark (d, s, b). For each fermion f one can define its counterpart, an anti-fermion $\bar{f}$ with opposite quantum numbers.

Table 1.1 Fermions in the SM, grouped in three generations. The values are taken from [3].

	quarks		leptons	
generation	type	mass	type	mass
1	u	1.8 - 2.8 MeV	ν_e	<2 eV
	d	4.3 - 5.2 MeV	e	0.511 MeV
2	c	1.28 GeV	ν_μ	<2 eV
	s	92 - 104 MeV	μ	105.7 MeV
3	t	173 GeV	ν_τ	<2 eV
	b	4.18 GeV	τ	1.78 GeV

The interactions between the fermions are mediated by gauge bosons with integer spin. They are introduced in the SM by transforming the global symmetries of the SM Lagrangian to local gauge symmetries that are separately valid at each space-time point. The gauge bosons, listed in Table 1.2, carry the three fundamental forces: the massless photon γ mediates the electromagnetic force, the massless gluons g carry the "charge" of the strong force and the massive $W^{\pm}$ and Z^0 bosons mediate the weak force. The gravitational force is not described in the SM. The Higgs boson H is not related to a fundamental force and its existence is needed for the mass generation mechanism in the SM as discussed later. In the following the different fundamental interactions and their relation to the SM gauge symmetry group $\mathrm{SU(3)_C}\otimes\mathrm{SU(2)_L}\otimes\mathrm{U(1)_Y}$ are discussed in detail.

Table 1.2 Bosons and fundamental forces in the SM. The values are taken from [3].

interaction	gauge boson	mass
electromagnetic	γ	$<1{\cdot}10^{-18}$ eV
strong	g	0
weak	$W^{\pm}$	80.4 GeV
	Z^0	91.2 GeV
-	H	125.1 GeV

Strong interaction

The theory of strong interaction, quantum chromodynamics (QCD), is based on the $\mathrm{SU(3)_C}$ gauge group and defines the interactions and couplings of the quarks. C denotes a quantum number related to the symmetry, called colour. The strong force acts on all particles with colour quantum number. The quarks can have the following colour quantum numbers: red, green or blue and the corresponding anti-colours. The gauge bosons of the strong interaction, gluons, have a colour quantum number as well and can therefore self-interact. The leptons do not have a colour and, thus, do not take part in strong interactions. A special feature of the strong interaction is that quarks can only exist in bound states, a phenomenon referred to as a confinement. The bound states are called hadrons. There are two most common types of hadrons: mesons consisting of a quark-anti-quark pair and baryons consisting of three quarks of different type.

Electroweak interaction

In the SM the electromagnetic and weak interactions are unified to a common theoretical description, the electroweak interaction with gauge group $\mathrm{SU(2)_L}\otimes\mathrm{U(1)_Y}$. The $\mathrm{SU(2)_L}$ symmetry group introduces three gauge bosons W_i. The related conserved quantum number is a weak isospin T with its projection T_3. Only left-handed fermions and right-handed anti-fermions have $T \neq 0$ and couple to the W_i bosons. The symmetry of the $\mathrm{U(1)_Y}$ gauge group implies the hypercharge quantum number Y and introduces a gauge boson A. The

hypercharge of a fermion is given by $Y = 2(Q - T_3)$ where Q is the electric charge. The exchange bosons of the electromagnetic and weak interactions are given by linear combinations of the gauge bosons W_i and A.

For the electromagnetic interaction the exchange boson is photon, a combination of the W_3 and A gauge bosons. It couples to the electric charge Q of fermions. All fermions except the neutrinos have an electric charge and therefore take part in electromagnetic interaction processes. The weak interaction processes can be classified by two different types. First, the neutral current (NC) is mediated by the Z^0 boson that is similar to the photon a combination of the W_3 and A gauge bosons. Due to charge conservation, it couples only to fermion-anti-fermion pairs. The charged current (CC) is carried by the charged $W^{\pm}$ bosons that are linear combinations of the W_1 and W_2 gauge bosons. The CC is the only process in the SM where fermions of different generations can take part.

The experiments have demonstrated that the $W^{\pm}$ and Z^0 gauge bosons have masses. Separate mass terms in the SM Lagrangian, however, would violate the local gauge symmetry. By introducing a scalar Higgs field with non-zero vacuum expectation value, the electroweak symmetry is spontaneously broken and the bosons gain masses. A consequence of the Higgs-mechanism is the existence of an additional scalar Higgs boson H. A first observation of a Standard Model H-like particle was made by the ATLAS [4] and CMS [5] collaborations. Similar to the boson masses, the masses of the fermions are introduced by their coupling to the Higgs field, the Yukawa coupling.

1.2 Flavour Sector of the Standard Model

1.2.1 CKM matrix

The charged weak interaction is the only interaction in the SM that can change flavour. The charged current part of the SM Lagrangian which governs the charged weak interaction in the quark sector is given by

$$\mathcal{L}_{\mathrm{CC}} = -\frac{g}{2\sqrt{2}} \left[\bar{u}_i W_\mu^+ \gamma^\mu (1-\gamma^5) d_i' + \bar{d}'_i W_\mu^- \gamma^\mu (1-\gamma^5) u_i \right], \tag{1.1}$$

where the (dashed) quark fields $\overset{(-)}{u}_i$ and $\overset{(-)}{d}'_i$ denote the weak eigenstates of up- and down-type (anti-)quarks, respectively [6]. The weak eigenstates are constructed by rotating the mass eigenstates with Cabbibo-Kobayashi-Maskawa (CKM) matrix V_{CKM} [7, 8], a complex unitary 3×3 matrix connecting down-type quarks to up-type quarks. In particular, the off-diagonal elements of the CKM matrix allow for transitions between quarks of different generations. It is customary to only rotate the down-type quarks (d, s, b) to

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.2}$$

As V_{CKM} is unitary, $V_{\mathrm{CKM}} V_{\mathrm{CKM}}^{\dagger} = 1$, the 18 parameters which describe a complex 3×3 matrix reduce to nine. Five parameters can be absorbed as unobservable quark phases which leaves four independent parameters. The standard parametrization of V_{CKM}, Chau-Keung parametrization [9], is characterized by three Euler angles $\theta_{12}, \theta_{13}, \theta_{23}$ and a phase δ. Using the shortcuts $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$, V_{CKM} is given by

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \tag{1.3}$$

A common parametrization of V_{CKM} which reflects the hierarchy of matrix elements is the Wolfenstein parametrization [10]. It uses the real parameters λ, A, ρ and η, with η being responsible for the imaginary part of the entries in the V_{CKM}. It can be obtained by using the definitions

$$s_{12} = \lambda, \qquad s_{23} = A\lambda^2, \qquad s_{13}e^{-i\delta} = A\lambda^3(\rho - i\eta)$$

in the standard parametrization. The parameter $\lambda \approx 0.22$ has the role of an expansion parameter which simplifies the estimation of the size of the V_{CKM} elements. The parametrization up to order λ^3 is [11]

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4).$$

From this parametrization it is immediately clear that the diagonal elements are ∼1, whereas the off-diagonal elements which are responsible for transitions between different generations are smaller with V_{us}, $V_{cd} \sim \lambda$, V_{cb}, $V_{ts} \sim \lambda^2$ and V_{ub}, $V_{td} \sim \lambda^3$. The imaginary part relative to the magnitude of the matrix element is largest for V_{ub}. Up to order λ^3 only V_{ub} and V_{td} have an imaginary component.

1.2.2 $\mathcal{CP}$ violation

Prior to 1957, it was thought that all physical laws were respected in quantum systems before and after a parity transformation, $\mathcal{P}$. The parity transformation is such that the spacial

coordinates of a system are transformed to their opposite values:

$$\mathcal{P}(x, y, z) = (-x, -y, -z). \tag{1.4}$$

In 1957 C. S. Wu and co-workers [12] made observations of Co^{60} decays at the request of T. D. Lee and C. N. Yang [13] and confirmed their prediction of P operations violation in β-decays. The result was confirmed by R. L. Garwin and co-workers [14] in parallel and in collaboration in cyclotron experiments. After this result it was thought that there would still be symmetry conservation if the $\mathcal{P}$ transformation was combined with charge conjugation, $\mathcal{C}$. Under a charge conjugation a particle is replaced with its anti-particle and positive charge is changed to negative charge:

$$\mathcal{C}(e^-) = e^+, \qquad \mathcal{C}(K^0) = \bar{K}^0. \tag{1.5}$$

Applying the $\mathcal{CP}$ transformation to the charged current part of the SM Lagrangian given by

$$\mathcal{L}_{\mathrm{CC}} = -\frac{g}{2\sqrt{2}}\left[(V_{\mathrm{CKM}})_{ij}\bar{u}_i W^+_\mu \gamma^\mu(1-\gamma^5)d_j + (V^*_{\mathrm{CKM}})_{ij}\bar{d}_j W^-_\mu \gamma^\mu(1-\gamma^5)u_i\right]$$

results in the $\mathcal{CP}$ conjugated Lagrangian [11]

$$\mathcal{L}^{\mathcal{CP}}_{\mathrm{CC}} = -\frac{g}{2\sqrt{2}}\left[(V_{\mathrm{CKM}})_{ij}\bar{d}_j W^-_\mu \gamma^\mu(1-\gamma^5)u_i + (V^*_{\mathrm{CKM}})_{ij}\bar{u}_i W^+_\mu \gamma^\mu(1-\gamma^5)d_j\right].$$

Both expressions are identical if $(V^*_{\mathrm{CKM}})_{ij} = (V_{\mathrm{CKM}})_{ij}$, i.e. all CKM matrix elements are real. A non-trivial phase δ (Eq. 1.3) inducing complex CKM matrix elements can, therefore, lead to $\mathcal{CP}$ violation.

In nature the $\mathcal{CP}$ symmetry is indeed violated. The $\mathcal{CP}$ violation was first discovered in 1964 in neutral K decays [15]. More recently the $\mathcal{CP}$ violation has also been observed in the B meson sector by the BaBar and Belle collaborations [16, 17].

Unitarity triangles

The unitarity condition $V_{\mathrm{CKM}}V^\dagger_{\mathrm{CKM}} = 1$ results in six equations for the off-diagonal elements[1]:

$$V_{ud}V^*_{us} + V_{cd}V^*_{cs} + V_{td}V^*_{ts} = 0, \tag{1.6}$$

$$V_{us}V^*_{ub} + V_{cs}V^*_{cb} + V_{ts}V^*_{tb} = 0, \tag{1.7}$$

$$V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0. \tag{1.8}$$

[1]The other three off-diagonal elements in the unitarity equation $V_{\mathrm{CKM}}V^\dagger_{\mathrm{CKM}} = 1$ result in three equations which are just the complex conjugates of Eqs. 1.6-1.8. Keep in mind that each one of the Eqs. 1.6-1.8 gives two conditions, one for the real part and one for the imaginary part.

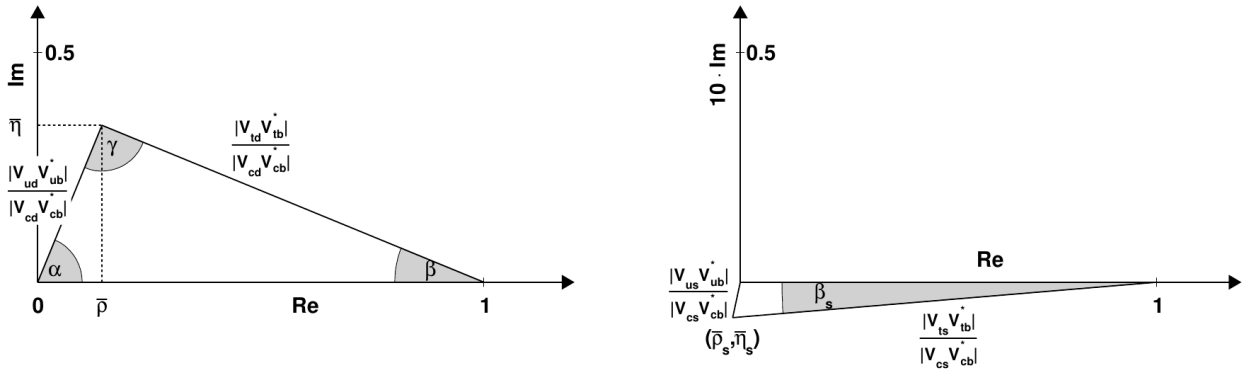
Since on the left-hand side of the equations there are sums of three complex numbers, the conditions can be visualized as triangles in the complex plane [9]. Usually, one side of the three triangles is normalized to coincide with the real axis which results in the equations:

$$\frac{V_{ud}V_{us}^*}{V_{cd}V_{cs}^*} + 1 + \frac{V_{td}V_{ts}^*}{V_{cd}V_{cs}^*} = 0, \tag{1.9}$$

$$\frac{V_{us}V_{ub}^*}{V_{cs}V_{cb}^*} + 1 + \frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*} = 0, \tag{1.10}$$

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0. \tag{1.11}$$

An illustration of the unitarity triangle resulting from Eq. 1.11 is given in Fig. 1.1(a). It is



B_d^0 triangle B_s^0 triangle

Fig. 1.1 Illustration of (left) the B_d^0 triangle and (right) the B_s^0 triangle in the complex plane. The B_s^0 triangle is very flat because the angle β_s is very small. Note that the imaginary axis for the right plot is scaled by a factor 10.

also called the "B_d^0 unitarity triangle", since its sides and angles are accessible through B_d^0 decays. The vertices of the triangle are, due to the normalization, (0,0), (0,1) and $(\bar{\rho}, \bar{\eta})$. The angles at the vertices are defined as:

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{1.12}$$

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \tag{1.13}$$

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right). \tag{1.14}$$

Fig. 1.1(b) gives the "B_s^0 unitarity triangle" resulting from Eq. 1.10. The apex of the triangle is located at $(\bar{\rho}_s, \bar{\eta}_s)$. This unitarity triangle is nearly flat with the small angle β_s at the

vertex (1,0) given by

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right). \tag{1.15}$$

Experimental status of the unitarity triangles determination

The SM does not predict the values of the CKM matrix elements. A central aim of flavour physics is to overconstrain the CKM triangles to test the unitarity conditions. Fig. 1.2 shows the unitarity triangles resulting from a global analysis of measurements determining the CKM matrix parameters in the framework of the SM and some of its extensions [18]. The lengths of the sides of the unitarity triangles are determined by the absolute values of CKM matrix elements:

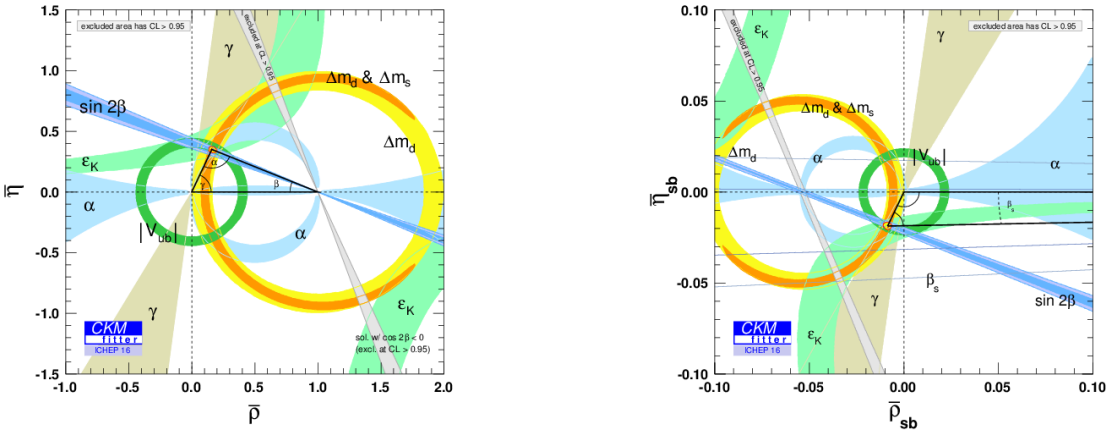


Fig. 1.2 (left) B_d^0 triangle and (right) B_s^0 triangle in the complex plane from a global CKM fit. The figures are taken from [18].

- $|V_{ub}|$ and $|V_{cb}|$ determine the length of the left side of the B_d^0 triangle. Both quantities can be accessed in semileptonic B decays. $|V_{ub}|$ can be extracted from semileptonic decays to light mesons, e.g. $B \to \pi l\nu$. $|V_{cb}|$ can be accessed using semileptonic B decays to charm, e.g. $B \to Dl\nu$.

- The right side of the B_d^0 triangle is determined by $|V_{td}|$ and $|V_{tb}|$ which can be constrained using $B_d^0 - \bar{B}_d^0$ and $B_s^0 - \bar{B}_s^0$ mixing.

The angles of the unitarity triangles are directly related to complex CKM matrix elements. They can, therefore, be accessed by precision measurements of $\mathcal{CP}$ violation.

- The angle β appears in B_d^0 meson mixing. The precise measurement of $\sin 2\beta$ was one of the main goal of the BaBar [16] and Belle [17] experiments and confirmed the existence of $\mathcal{CP}$ violation in the B sector for the first time. Both experiments determine β using

the "gold-plated" decay channel $B^0_d \to J/\psi K^0_S$. The most recent combined result [3] including LHCb [19] and other measurements [20] is

$$\sin 2\beta = 0.691 \pm 0.017.$$

This combination represents the most precise angular constraint entering the fit. It is given by the dark blue band in Fig. 1.2.

- The angle β_s appears in B^0_s meson mixing. The current experimental situation is discussed in detail in Sec. 1.3.5. The SM determination of $\phi_s = -2\beta_s$ using the $B^0_s \to J/\psi\phi$ decay channel is one of the main physics goals of the LHCb experiment and the topic of this thesis.

1.2.3 Mixing phenomenology

The overview of the phenomenology of B meson mixing and decay broadly follows more detailed review articles [11, 21, 22]. Even though the formalism developed below describes neutral B mesons B^0_q (with $q = d, s$), it is also valid for K and D mesons.

The decays of neutral mesons B^0_q and their $\mathcal{CP}$ conjugates $\bar{B}^0_q$ into final states f and $\bar{f}$ are described by the decay amplitudes

$$A_f = \langle f|\mathcal{H}|B^0_q\rangle, \qquad \bar{A}_f = \langle f|\mathcal{H}|\bar{B}^0_q\rangle,$$

$$A_{\bar{f}} = \langle \bar{f}|\mathcal{H}|B^0_q\rangle, \qquad \bar{A}_{\bar{f}} = \langle \bar{f}|\mathcal{H}|\bar{B}^0_q\rangle.$$

The neutral B mesons transform under $\mathcal{CP}$ operation according to

$$\mathcal{CP}|B^0_q\rangle = -|\bar{B}^0_q\rangle \quad \text{and} \quad \mathcal{CP}|\bar{B}^0_q\rangle = -|B^0_q\rangle$$

where an arbitrary non-physical phase factor has been omitted. The final state f and its $\mathcal{CP}$ conjugate state $\bar{f}$ are connected via the $\mathcal{CP}$ operation according to

$$\mathcal{CP}|f\rangle = \eta_f|\bar{f}\rangle \quad \text{and} \quad \mathcal{CP}|\bar{f}\rangle = \eta_f|f\rangle$$

with $\eta_f = \pm 1$ denoting the phase factors appearing in the $\mathcal{CP}$ transformation of the states.

The time development of the flavour eigenstates $|B^0_q\rangle$ and $|\bar{B}^0_q\rangle$ is given by the phenomenological Schrödinger equation,

$$i\frac{\partial}{\partial t}\begin{pmatrix}|B^0_q\rangle\\ |\bar{B}^0_q\rangle\end{pmatrix} = \left(M - \frac{i}{2}\Gamma\right)\begin{pmatrix}|B^0_q\rangle\\ |\bar{B}^0_q\rangle\end{pmatrix} = \begin{pmatrix}M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12}\\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22}\end{pmatrix}\begin{pmatrix}|B^0_q\rangle\\ |\bar{B}^0_q\rangle\end{pmatrix},$$

where the Hamiltonian is constructed from two hermitian matrices: mass M and decay width Γ matrix. The mass matrix describes $B^0_q - \bar{B}^0_q$ meson mixing.

In the SM the off-diagonal terms M_{12} and M_{12}^* occur from the flavour changing $\Delta b = 2$ processes given in Fig. 1.3. Due to the hermeticity of M and Γ the off-diagonal elements

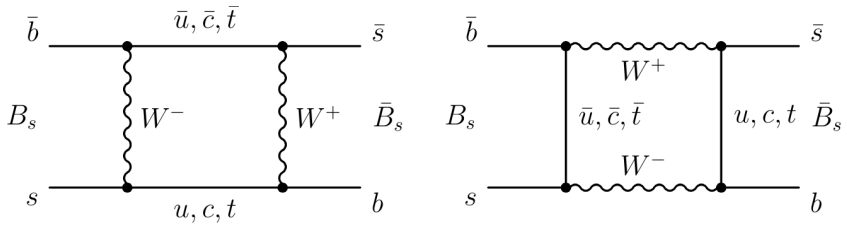


Fig. 1.3 B_s^0 meson mixing in the Standard Model.

are complex conjugates, $M_{21} = M_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$. Additionally, the $\mathcal{CPT}$ invariance theorem [23] gives $\Gamma = \Gamma_{11} = \Gamma_{22}$ and $M = M_{11} = M_{22}$, so, that the Hamiltonian simplifies to

$$i\frac{\partial}{\partial t}\begin{pmatrix}|B_q^0\rangle\\|\bar{B}_q^0\rangle\end{pmatrix} = \begin{pmatrix}M - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12}\\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M - \frac{i}{2}\Gamma\end{pmatrix}\begin{pmatrix}|B_q^0\rangle\\|\bar{B}_q^0\rangle\end{pmatrix}. \tag{1.16}$$

Diagonalization of the Hamiltonian in Eq. 1.16 leads to the mass eigenstates $|B_L\rangle$ and $|B_H\rangle$

$$|B_L\rangle = p|B_q^0\rangle + q|\bar{B}_q^0\rangle,$$

$$|B_H\rangle = p|B_q^0\rangle - q|\bar{B}_q^0\rangle$$

where p and q are the complex coefficients and $|p|^2 + |q|^2 = 1$. The $|B_{L/H}\rangle$ denotes the lighter and heavier mass eigenstate, respectively. The mass eigenstates develop in time with

$$|B_L(t)\rangle = e^{-iM_L t}e^{-\frac{\Gamma_L}{2}t}|B_L\rangle,$$

$$|B_H(t)\rangle = e^{-iM_H t}e^{-\frac{\Gamma_H}{2}t}|B_H\rangle$$

where M_L and M_H denote the masses and Γ_L and Γ_H are the decay widths of $|B_L\rangle$ and $|B_H\rangle$. The diagonalization procedure relates $M_{L/H}$ and $\Gamma_{L/H}$ to the elements of the Hamiltonian in Eq. 1.16

$$\begin{aligned} M_{L/H} - \frac{i}{2}\Gamma_{L/H} &= M - \frac{i}{2}\Gamma \mp \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}\\ &= M - \frac{i}{2}\Gamma \mp \sqrt{|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2 - i|M_{12}||\Gamma_{12}|\cos(\phi_\Gamma - \phi_M)} \end{aligned} \tag{1.17}$$

with the phases $\phi_\Gamma = \arg(\Gamma_{12})$ and $\phi_M = \arg(M_{12})$. The mass and decay width of the mass eigenstates are related to Γ and M via

$$\Gamma = \frac{\Gamma_L + \Gamma_H}{2}, \qquad M = \frac{M_L + M_H}{2}.$$

In the following the mass difference Δm and decay width difference $\Delta\Gamma$ are defined as

$$\Delta m = M_H - M_L, \qquad \Delta\Gamma = \Gamma_L - \Gamma_H.$$

Using Eq. 1.17 the following relations for Δm and $\Delta\Gamma$ can be derived:

$$\Delta m^2 - \frac{1}{4}\Delta\Gamma^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2, \tag{1.18}$$

$$\Delta m \Delta\Gamma = -4|M_{12}||\Gamma_{12}|\cos(\phi_\Gamma - \phi_M). \tag{1.19}$$

In both the B_d^0 and B_s^0 system experimental evidence shows $\Delta m \gg \Delta\Gamma$ [3] which translates to $|M_{12}| \gg |\Gamma_{12}|$. Neglecting $\Delta\Gamma$ and $|\Gamma_{12}|^2$ in Eq. 1.18 results in

$$\Delta m = 2|M_{12}|. \tag{1.20}$$

Inserting this in Eq. 1.19 gives

$$\Delta\Gamma = -2|\Gamma_{12}|\cos(\phi_\Gamma - \phi_M). \tag{1.21}$$

The coefficients of the eigenstates are determined to be

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

When studying the B_d^0 and the B_s^0 system this expression can be expanded in $|\Gamma_{12}|/|M_{12}|$ which gives

$$\frac{q}{p} = -\sqrt{e^{-2i\phi_M}\frac{1 - \frac{i}{2}\frac{|\Gamma_{12}|}{|M_{12}|}e^{-i\phi_\Gamma + i\phi_M}}{1 - \frac{i}{2}\frac{|\Gamma_{12}|}{|M_{12}|}e^{+i\phi_\Gamma - i\phi_M}}}$$

$$= -e^{-i\phi_M}\left[1 - \frac{1}{2}\sin(\phi_\Gamma - \phi_M)\frac{|\Gamma_{12}|}{|M_{12}|} + \mathcal{O}\left(\frac{|\Gamma_{12}|^2}{|M_{12}|^2}\right)\right].$$

The time development for $|B_q^0\rangle$ and $|\bar{B}_q^0\rangle$ is given by

$$|B_q^0(t)\rangle = \frac{1}{2p}(|B_L(t)\rangle + |B_H(t)\rangle),$$

$$|\bar{B}_q^0(t)\rangle = \frac{1}{2q}(|B_L(t)\rangle - |B_H(t)\rangle)$$

inserting $|B_L(t)\rangle$ and $|B_H(t)\rangle$ results in

$$|B_q^0(t)\rangle = g_+(t)|B_q^0\rangle + \frac{q}{p}g_-(t)|\bar{B}_q^0\rangle,$$

$$|\bar{B}_q^0(t)\rangle = \frac{p}{q} g_-(t)|B_q^0\rangle + g_+(t)|\bar{B}_q^0\rangle$$

with

$$g_\pm(t) = \frac{1}{2}\left(e^{-iM_L t}e^{-\frac{\Gamma_L}{2}t} \pm e^{-iM_H t}e^{-\frac{\Gamma_H}{2}t}\right).$$

For the determination of the time dependent decay rates in the following, it is convenient to explicitly calculate some combinations of these terms

$$|g_+(t)|^2 = \frac{1}{2}e^{-\Gamma t}\left(\cosh\frac{\Delta\Gamma}{2}t + \cos\Delta m t\right),$$

$$|g_-(t)|^2 = \frac{1}{2}e^{-\Gamma t}\left(\cosh\frac{\Delta\Gamma}{2}t - \cos\Delta m t\right),$$

$$g_+(t)g_-^*(t) = \frac{1}{2}e^{-\Gamma t}\left(-\sinh\frac{\Delta\Gamma}{2}t - i\sin\Delta m t\right),$$

$$g_+^*(t)g_-(t) = \frac{1}{2}e^{-\Gamma t}\left(-\sinh\frac{\Delta\Gamma}{2}t + i\sin\Delta m t\right).$$

To simplify the expressions, it is further useful to define one central quantity for $\mathcal{CP}$ violation,

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}. \tag{1.22}$$

While the phases of both $\bar{A}_f/A_f$ and q/p are convention dependent, the phase of λ_f is a measurable physical quantity. The time dependent decay rate for the decay of a produced B_q^0 to the final state f is

$$\frac{\mathrm{d}\Gamma(B_q^0 \to f)}{\mathrm{d}t\mathcal{N}_f} = |\langle f|B_q^0(t)\rangle|^2 = \left|g_+(t)A_f + \frac{q}{p}g_-(t)\bar{A}_f\right|^2$$

$$= \frac{1}{2}|A_f|^2 e^{-\Gamma t}\left[(1+|\lambda_f|^2)\cosh\frac{\Delta\Gamma}{2}t + (1-|\lambda_f|^2)\cos\Delta m t - 2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\Re(\lambda_f) - 2\sin(\Delta m t)\Im(\lambda_f)\right]. \tag{1.23}$$

Similarly the decay rate for a $\bar{B}_q^0$ to decay to f is given by

$$\frac{\mathrm{d}\Gamma(\bar{B}_q^0 \to f)}{\mathrm{d}t\mathcal{N}_f} = |\langle f|\bar{B}_q^0(t)\rangle|^2 = \left|\frac{p}{q}g_-(t)A_f + g_+(t)\bar{A}_f\right|^2$$

$$= \frac{1}{2}\left|\frac{p}{q}\right|^2 |A_f|^2 e^{-\Gamma t}\left[(1+|\lambda_f|^2)\cosh\frac{\Delta\Gamma}{2}t - (1-|\lambda_f|^2)\cos\Delta m t - 2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\Re(\lambda_f) + 2\sin(\Delta m t)\Im(\lambda_f)\right] \tag{1.24}$$

where $\mathcal{N}_f$ is a normalization factor. The decay rates to the $\mathcal{CP}$ conjugated final state $\bar{f}$ are given by the above expressions when substituting $A_f \to A_{\bar{f}}$, $\bar{A}_f \to \bar{A}_{\bar{f}}$ and $\lambda_f \to \lambda_{\bar{f}} = \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_{\bar{f}}}$ (App. A).

1.2.4 Types of $\mathcal{CP}$ violation

In the previous section, the formalism for meson decay and mixing has been developed. Assuming $\mathcal{CPT}$ invariance [23], all $\mathcal{CP}$-violating effects in a $B_q^0 - \bar{B}_q^0$ meson system are governed by the Schrödinger-like equation developed in Eq. 1.16. The resulting time dependent decay rates are expressed using the decay amplitudes A_f, $A_{\bar{f}}$, $\bar{A}_f$, $\bar{A}_{\bar{f}}$ and the mixing parameters q and p. However, to arrive at non-vanishing $\mathcal{CP}$-violating observables, specific conditions for the transition amplitudes and their phases need to be met.

In this context, two types of phases need to be distinguished: "strong" and "weak" phases. The strong phases arise in final state interaction scattering from intermediate on-shell states, e.g. through strong or electromagnetic interactions. As these interactions are $\mathcal{CP}$ invariant, the strong phases are equal for two $\mathcal{CP}$ conjugate states. Weak phases originate from complex couplings in the Lagrangian. In the case of the SM, these arise in the $W^{\pm}$ couplings to the quarks via the complex CKM matrix elements. The phases of this type change sign under $\mathcal{CP}$ conjugation. Accordingly, in the course of a further investigation of the CKM mechanism, it is of highest interest to cleanly measure weak phases in meson decays.

It is important to note that although the existence of weak and strong phases is physically motivated, their absolute values are convention-dependent. In contrast, relative strong or weak phases between different terms in transition amplitudes are convention-independent, and hence are physically meaningful. Depending on the origin and interplay of the transitions that introduce weak phases, $\mathcal{CP}$-violating effects can be categorized as $\mathcal{CP}$ violation in the decay (direct $\mathcal{CP}$ violation), $\mathcal{CP}$ violation in the mixing (indirect $\mathcal{CP}$ violation), or $\mathcal{CP}$ violation in the interference of mixing and decay.

- **$\mathcal{CP}$ violation in decay**

  In the case of meson decays, multiple contributions a_i, each with different weak phases ϕ_i and strong phases δ_i, can contribute to the decay amplitudes. In the simple case of two possible transitions with amplitudes a_1 and a_2, the resulting decay amplitudes are

$$A_f = |a_1| e^{i(\delta_1+\phi_1)} + |a_2| e^{i(\delta_2+\phi_2)},$$

$$\bar{A}_{\bar{f}} = |a_1| e^{i(\delta_1-\phi_1)} + |a_2| e^{i(\delta_2-\phi_2)}.$$

  This directly implies

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| \neq 1, \tag{1.25}$$

  leading to $\mathcal{CP}$ violation in the decay. This is the only type of $\mathcal{CP}$ violation that can occur in charged meson systems B^+ and B^-. It leads to a time independent asymmetry of the form

$$\mathcal{A}_{\mathcal{CP}}^{dir} = \frac{\Gamma(B^- \to f^-) - \Gamma(B^+ \to f^+)}{\Gamma(B^- \to f^-) + \Gamma(B^+ \to f^+)} = \frac{|\bar{A}_{\bar{f}}|^2 - |A_f|^2}{|\bar{A}_{\bar{f}}|^2 + |A_f|^2} = \frac{|\bar{A}_{\bar{f}}/A_f|^2 - 1}{|\bar{A}_{\bar{f}}/A_f|^2 + 1}. \tag{1.26}$$

An example for a such kind of the $\mathcal{CP}$ asymmetry is the $B_d^0 \to K^+\pi^-$ decay [24]. In this decay the tree level amplitude interferes with penguin amplitudes leading to the $\mathcal{CP}$ asymmetry of $\sim$10%.

- **$\mathcal{CP}$ violation in mixing**
  The $\mathcal{CP}$ violation in the $B_q^0 - \bar{B}_q^0$ meson mixing occurs when $|q/p| \neq 1$. In this case, the probability for a B_q^0 to transition into a $\bar{B}_q^0$, $\mathcal{P}(B_q^0 \to \bar{B}_q^0)$, differs from the probability for the $\mathcal{CP}$ conjugated process, $\mathcal{P}(\bar{B}_q^0 \to B_q^0)$. The resulting asymmetry assuming no direct $\mathcal{CP}$ violation, i. e. $A_f = \bar{A}_{\bar{f}}$ and $A_{\bar{f}} = \bar{A}_f = 0$, is given by

$$\mathcal{A}_{\mathcal{CP}}^{mix} = \frac{\Gamma(\bar{B}_q^0 \to f) - \Gamma(B_q^0 \to \bar{f})}{\Gamma(\bar{B}_q^0 \to f) + \Gamma(B_q^0 \to \bar{f})} = \frac{|\frac{p}{q} g_-(t) A_f|^2 - |\frac{q}{p} g_-(t) \bar{A}_{\bar{f}}|^2}{|\frac{p}{q} g_-(t) A_f|^2 + |\frac{q}{p} g_-(t) \bar{A}_{\bar{f}}|^2} = \frac{1 - \left|\frac{q}{p}\right|^4}{1 + \left|\frac{q}{p}\right|^4}. \tag{1.27}$$

  The $\mathcal{CP}$ violation in B meson mixing can be studied using semileptonic $B_q^0 \to X l^+$ decays. The events, where a B_q^0 meson has mixed before decaying semileptonically, result in "wrong sign" decays containing a l^- in the final state. Under the assumption of equal production of B_q^0 and $\bar{B}_q^0$ mesons, $\mathcal{P}(B_q^0 \to \bar{B}_q^0) \neq \mathcal{P}(\bar{B}_q^0 \to B_q^0)$ will lead to an asymmetry in the observed number of l^+ and l^- given by $\mathcal{A}_{\mathcal{CP}}^{mix}$. The measurement results of CP asymmetry in the mixing of B_q^0 mesons are compatible with zero and are in agreement with the SM expectations [25, 26].

- **Mixing induced $\mathcal{CP}$ violation**
  The B meson mixing induced $\mathcal{CP}$ violation can occur when the direct decay $B_q^0 \to f$ interferes with mixing from B_q^0 to $\bar{B}_q^0$ followed by the decay $\bar{B}_q^0 \to f$. This situation is illustrated in Fig. 1.4. If the term λ_f in Eqs. 1.23 and 1.24 has a non-trivial phase, i. e. $\Im(\lambda_f) = \Im(\frac{q}{p}\frac{\bar{A}_f}{A_f}) \neq 0$, this gives rise to the time dependent $\mathcal{CP}$ asymmetry

$$\mathcal{A}_{\mathcal{CP}}(t) = \frac{\Gamma(\bar{B}_q^0 \to f)(t) - \Gamma(B_q^0 \to f)(t)}{\Gamma(\bar{B}_q^0 \to f)(t) + \Gamma(B_q^0 \to f)(t)} = \frac{(1 - |\lambda_f|^2)\cos(\Delta m t) - 2\sin(\Delta m t)\Im(\lambda_f)}{(1 + |\lambda_f|^2)\cosh(\frac{\Delta\Gamma}{2}t) + 2\sinh(\frac{\Delta\Gamma}{2}t)\Re(\lambda_f)}. \tag{1.28}$$

  In case of B_q^0 candidates $\Delta\Gamma$=0 [3] and assuming that the transition is dominated by only one amplitude, $|A_f| = |\bar{A}_f|$ or $|\lambda_f| = 1$, the asymmetry simplifies to

$$\mathcal{A}_{\mathcal{CP}}(t) = -\sin(\Delta m t)\Im(\lambda_f). \tag{1.29}$$

  An example for this type of $\mathcal{CP}$ violation is the measurement of $\sin 2\beta_d$ in the $B_d^0 \to J/\psi K_S^0$ decay [19]. The analogous "golden mode" in the B_s^0 meson system is given by the $B_s^0 \to J/\psi\phi$ decay which is the focus of this thesis.

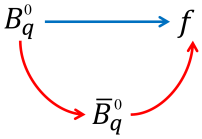
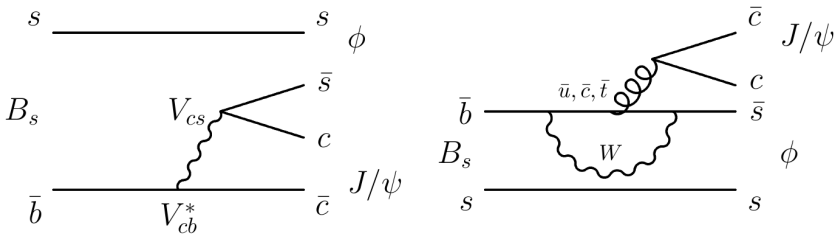


Fig. 1.4 Interference between (blue) decay and (red) decay after $B_q^0 - \bar{B}_q^0$ meson mixing.

1.3 $\mathcal{CP}$-violating phase ϕ_s in $B_s^0 \to J/\psi\phi$

The $B_s^0 \to J/\psi\phi$ decay implies a $b \to c\bar{c}s$ quark transition as shown in Fig. 1.5. The total



(a) Tree contribution (b) Penguin contribution

Fig. 1.5 Both (a) tree and (b) penguin processes can contribute to the decay $B_s^0 \to J/\psi\phi$ but the tree decay amplitude dominates.

amplitude $A_{J/\psi\phi}$ for the transition of a B_s^0 meson into the final state $J/\psi\phi$ is given by a sum

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

$$= V_{cs}V_{cb}^*(T + P_c - P_t) + V_{us}V_{ub}^*(P_u - P_t)$$

where T is the tree amplitude (Fig. 1.5a) and P_q with $q = u, c, t$ denotes the corresponding penguin amplitudes (Fig. 1.5b). $V_{ts}V_{tb}^*$ is expressed through $V_{cs}V_{cb}^*$ and $V_{us}V_{ub}^*$ using "B_s^0 unitarity triangle" (Eq. 1.10). The decay amplitude A_f of $B_q^0 \to f$ can be written as

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

where ϕ_D is a weak phase that enters the decay amplitude and its $\mathcal{CP}$ conjugate with different signs. The δ_f is a strong phase that does not change sign under a $\mathcal{CP}$ transformation. The decay amplitude of $\bar{B}_q^0 \to f$ can be related to $\bar{A}_f$ using the $\mathcal{CP}$ condition $|f\rangle = \eta_{J/\psi\phi}|\bar{f}\rangle$, and assuming no $\mathcal{CP}$ violation in decay, as

$$\bar{A}_f = \eta_{J/\psi\phi}\bar{A}_{\bar{f}} = \eta_{J/\psi\phi}|A_f|e^{i(\delta_f - \phi_D)}$$

where the $\mathcal{CP}$-parity of the final state is denoted by $\eta_{J/\psi\phi}$.

Neglecting the terms proportional to $V_{us}V_{ub}^*$ (Eq. 1.10) in the total amplitude $A_{J/\psi\phi}$, the amplitude ratio $\bar{A}_f/A_f$ is given by

$$\frac{\bar{A}_f}{A_f} = -\eta_{J/\psi\phi}\frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}} = -\eta_{J/\psi\phi}e^{2i\phi_D}$$

with the decay phase, $\phi_D = \arg(V_{cb}V_{cs}^*)$. The same final state $J/\psi\phi$ can be reached if the B_s^0 meson, first, mixes to a $\bar{B}_s^0$ (Fig. 1.3) and then decays. In such case the ratio q/p can be expressed as

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

where the mixing phase, $\phi_M = -2\arg(V_{ts}V_{tb}^*)$. Combining $\bar{A}_f/A_f$ and q/p into λ_f (Eq. 1.22) gives

$$\lambda_f = \eta_{J/\psi\phi}\frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}}\frac{V_{ts}V_{tb}^*}{V_{ts}^*V_{tb}} = \eta_{J/\psi\phi}e^{+2i\phi_D - i\phi_M} = \eta_{J/\psi\phi}e^{-i\phi_s} \tag{1.30}$$

with $\phi_s = -\arg(\eta_{J/\psi\phi}\lambda_f)$. The phase ϕ_s can be related to the angle β_s in the B_s^0 unitarity triangle (Fig. 1.1b) via the relation $\phi_s = -2\beta_s$ with $\beta_s = \arg(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*})$ as defined in Eq. 1.15. While the mixing and decay phases depend on phase conventions and are not observable, the phase ϕ_s is a measurable physical quantity.

Looking at the Wolfenstein parametrization (Sec. 1.2.1) it is obvious that the angle β_s will be much smaller than β since V_{ts} is real up to $\mathcal{O}(\lambda^3)$ whereas V_{td} is already complex at $\mathcal{O}(\lambda^3)$. In the SM the phase ϕ_s is very small and very well predicted and is equal to $-0.0367^{+0.0007}_{-0.0008}$ rad [18].

In contrast to the B_d^0 system, the B_s^0 system possess a significant decay width difference $\Delta\Gamma_s$. The SM predicts $\Delta\Gamma_s = 0.088 \pm 0.020$ ps^{-1} [27], which means that the decay width difference may not be neglected.

In summary of the decay phenomenology discussed above, the time dependent mixing induced $\mathcal{CP}$ asymmetry (Sec. 1.2.4) for the signal $B_s^0 \to J/\psi\phi$ decay is given by

$$\mathcal{A}_{\mathcal{CP}}(t) = \frac{-\eta_{J/\psi\phi}\sin(\phi_s)\sin(\Delta m_s t)}{\cosh(\frac{\Delta\Gamma_s}{2}t) - \eta_{J/\psi\phi}\cos(\phi_s)\sinh(\frac{\Delta\Gamma_s}{2}t)}. \tag{1.31}$$

1.3.1 Angular analysis

Since both J/ψ and ϕ are vector mesons, the decay $B_s^0 \to J/\psi\phi$ is a pseudoscalar to vector-vector transition ($P \to VV$). This allows the final state mesons to have relative angular momentum l which leads to different $\mathcal{CP}$ eigenvalues of the final state depending on l,

$$\eta_{J/\psi\phi} = \eta_{J/\psi}^{\mathcal{CP}}\eta_{\phi}^{\mathcal{CP}}(-1)^l = (-1)^l.$$

To decouple the $\mathcal{CP}$-odd ($\eta_{J/\psi\phi}$ = -1) and $\mathcal{CP}$-even ($\eta_{J/\psi\phi}$ = +1) components statistically, an analysis of the angular distributions is necessary. The measured physics observables are independent from the choice of the basis but the basis determines the power to separate three components. Two bases can be used for this kind of measurement:

- The transversity basis with three transversity angles $\Omega_{\text{tr}} = \{\psi, \theta, \varphi\}$ is defined in Fig. 1.6. The angles θ and φ are defined in the rest frame of the J/ψ meson. The $x-y$ plane is defined by the K^+K^- plane, the x-axis by the flight direction of the ϕ meson and the y-axis such that $p_y(K^+) > 0$. The angle θ is then defined as the angle between the l^+ lepton flight direction and the z-axis. The angle φ is the angle between x-axis and the projection of the l^+ lepton flight direction on the $x-y$ plane. The angle ψ is defined as the angle between x-axis and the K^+ flight direction in the K^+K^- rest frame.

- The helicity basis with three helicity angles $\Omega_{\text{hel}} = \{\theta_l, \theta_K, \varphi_h\}$ is defined in Fig. 1.7. The θ_l is the angle between the l^+ and the opposite B_s^0 flight direction in the l^+l^- center-of-mass frame. The angle θ_K is similarly defined as the angle between the K^+ and the opposite B_s^0 flight direction in the K^+K^- center-of-mass frame. The relative rotation of the two decay planes is defined by φ_h as the angle between the K^- side of the K^+K^- plane and the l^+ side of the l^+l^- plane.

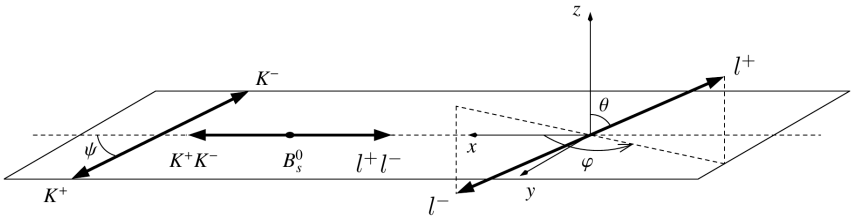


Fig. 1.6 Definition of transversity angle basis. The figure is taken from [28].

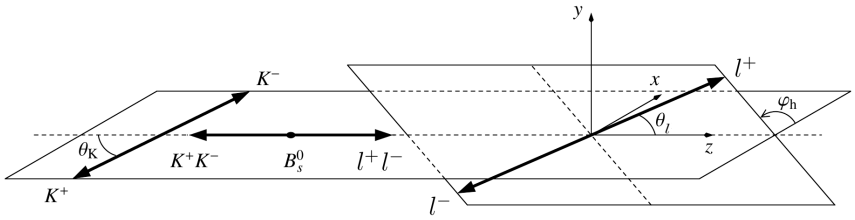


Fig. 1.7 Definition of helicity angle basis. The figure is taken from [28].

Both angular bases provide the same separation power for the three angular momentum states of the $B_s^0 \to J/\psi\phi$ decay and could both be chosen. This analysis is performed with the helicity basis ($\Omega_{\text{hel}} \equiv \Omega$) which has advantages in the description of background and detector acceptances.

1.3.2 Differential decay rate

The time and angular dependent differential decay rate for a B_s^0 meson decay to the $J/\psi\phi$ final state produced at t =0 is given by

$$\frac{\mathrm{d}\Gamma(B_s^0 \to J/\psi\phi)}{\mathrm{d}t\mathrm{d}\Omega} \propto \sum_{i=1}^{6} A_i(t) \cdot f_i(\cos\theta_K, \cos\theta_l, \varphi_h) \tag{1.32}$$

where A_i are six combinations of the three complex polarization amplitudes $A_0(t)$, $A_\perp(t)$ and $A_\parallel(t)$ [29] that correspond to different linear polarization states of the vector mesons J/ψ and ϕ with respect to their direction of motion (Fig. 1.8). The amplitude $A_0(t)$ corresponds to the longitudinal polarization of the vector mesons. The amplitudes $A_\parallel(t)$ and $A_\perp(t)$ correspond to transverse polarization of the vector mesons, in the former case the polarization states of J/ψ and ϕ are parallel, in the latter case, perpendicular to each other.

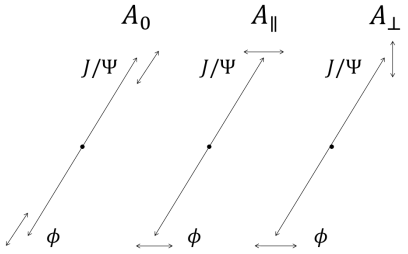


Fig. 1.8 Illustration of the three polarization amplitudes of the $B_s^0 \to J/\psi\phi$ decay. The $A_0(t)$ corresponds to the longitudinal polarization of the J/ψ and ϕ mesons, $A_\parallel(t)$ and $A_\perp(t)$ to a parallel and perpendicular transverse polarization.

The relative phases of the transversity amplitudes at t =0 are denoted as $\delta_0 = \arg(A_0(0))$, $\delta_\parallel = \arg(A_\parallel(0))$ and $\delta_\perp = \arg(A_\perp(0))$. The phases arise from strong final state interactions and are difficult to predict in theory. As they are invariant under $\mathcal{CP}$ transformation, they are also called "strong phases". One of these phases can be chosen freely since only the relative phases are important, in this thesis the convention $\delta_0 = 0$ is used. The magnitudes of the polarization amplitudes fulfill the normalization condition $|A_0(0)|^2 + |A_\perp(0)|^2 + |A_\parallel(0)|^2$=1. The angular dependent terms $f_i(\cos\theta_K, \cos\theta_l, \varphi_h)$ are given in Ref. [30] and summarized in Table 1.3. The time dependent amplitude combinations $A_i(t)$ for a B_s^0 meson decay are defined as [30]

$$
\begin{aligned}
|A_0(t)|^2 &= |A_0(0)|^2 e^{-\Gamma_s t}\left[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \sin\phi_s \sin(\Delta m_s t)\right],\\
|A_\parallel(t)|^2 &= |A_\parallel(0)|^2 e^{-\Gamma_s t}\left[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \sin\phi_s \sin(\Delta m_s t)\right],\\
|A_\perp(t)|^2 &= |A_\perp(0)|^2 e^{-\Gamma_s t}\left[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \sin\phi_s \sin(\Delta m_s t)\right],\\
\Im\{A_\parallel^*(t)A_\perp(t)\} &= |A_\parallel(0)||A_\perp(0)|e^{-\Gamma_s t}[-\cos(\delta_\perp - \delta_\parallel)\sin\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad + \sin(\delta_\perp - \delta_\parallel)\cos(\Delta m_s t) - \cos(\delta_\perp - \delta_\parallel)\cos\phi_s \sin(\Delta m_s t)],\\
\Re\{A_0^*(t)A_\parallel(t)\} &= |A_0(0)||A_\parallel(0)|e^{-\Gamma_s t}\cos(\delta_\parallel - \delta_0)[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad + \sin\phi_s \sin(\Delta m_s t)],\\
\Im\{A_0^*(t)A_\perp(t)\} &= |A_0(0)||A_\perp(0)|e^{-\Gamma_s t}[-\cos(\delta_\perp - \delta_0)\sin\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad + \sin(\delta_\perp - \delta_0)\cos(\Delta m_s t) - \cos(\delta_\perp - \delta_0)\cos\phi_s \sin(\Delta m_s t)].
\end{aligned}
$$

The prescription used for produced $\bar{B}_s^0$ mesons at t =0 is very similar and uses $\bar{A}_0(t)$, $\bar{A}_\perp(t)$ and $\bar{A}_\parallel(t)$

$$\frac{\mathrm{d}\Gamma(\bar{B}_s^0 \to J/\psi\phi)}{\mathrm{d}t\mathrm{d}\Omega} \propto \sum_{i=1}^{6} \bar{A}_i(t) \cdot f_i(\cos\theta_K, \cos\theta_l, \varphi_h) \tag{1.33}$$

with the same angular dependent terms $f_i(\cos\theta_K, \cos\theta_l, \varphi_h)$. The time dependent amplitudes for a $\bar{B}_s^0$ meson decay are found by changing the signs in front of all terms proportional to $\sin(\Delta m_s t)$ or $\cos(\Delta m_s t)$ which results in

$$
\begin{aligned}
|\bar{A}_0(t)|^2 &= |\bar{A}_0(0)|^2 e^{-\Gamma_s t}\left[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \sin\phi_s \sin(\Delta m_s t)\right],\\
|\bar{A}_\parallel(t)|^2 &= |\bar{A}_\parallel(0)|^2 e^{-\Gamma_s t}\left[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \sin\phi_s \sin(\Delta m_s t)\right],\\
|\bar{A}_\perp(t)|^2 &= |\bar{A}_\perp(0)|^2 e^{-\Gamma_s t}\left[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \sin\phi_s \sin(\Delta m_s t)\right],\\
\Im\{\bar{A}_\parallel^*(t)\bar{A}_\perp(t)\} &= |\bar{A}_\parallel(0)||\bar{A}_\perp(0)| e^{-\Gamma_s t}[-\cos(\delta_\perp - \delta_\parallel)\sin\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad - \sin(\delta_\perp - \delta_\parallel)\cos(\Delta m_s t) + \cos(\delta_\perp - \delta_\parallel)\cos\phi_s \sin(\Delta m_s t)],\\
\Re\{\bar{A}_0^*(t)\bar{A}_\parallel(t)\} &= |\bar{A}_0(0)||\bar{A}_\parallel(0)| e^{-\Gamma_s t}\cos(\delta_\parallel - \delta_0)[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad - \sin\phi_s \sin(\Delta m_s t)],\\
\Im\{\bar{A}_0^*(t)\bar{A}_\perp(t)\} &= |\bar{A}_0(0)||\bar{A}_\perp(0)| e^{-\Gamma_s t}[-\cos(\delta_\perp - \delta_0)\sin\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad - \sin(\delta_\perp - \delta_0)\cos(\Delta m_s t) + \cos(\delta_\perp - \delta_0)\cos\phi_s \sin(\Delta m_s t)].
\end{aligned}
$$

Table 1.3 Angular dependent terms $f_i(\cos\theta_K, \cos\theta_l, \varphi_h)$ [30].

i	$A_i(t)$	$\bar{A}_i(t)$	$f_i(\cos\theta_K, \cos\theta_l, \varphi_h)$
1	$\lvert A_0(t)\rvert^2$	$\lvert\bar{A}_0(t)\rvert^2$	$2\cos^2\theta_K \sin^2\theta_l$
2	$\lvert A_\parallel(t)\rvert^2$	$\lvert\bar{A}_\parallel(t)\rvert^2$	$\sin^2\theta_K(1 - \sin^2\theta_l \cos^2\varphi_h)$
3	$\lvert A_\perp(t)\rvert^2$	$\lvert\bar{A}_\perp(t)\rvert^2$	$\sin^2\theta_K(1 - \sin^2\theta_l \sin^2\varphi_h)$
4	$\Im\{A_\parallel^*(t)A_\perp(t)\}$	$\Im\{\bar{A}_\parallel^*(t)\bar{A}_\perp(t)\}$	$\sin^2\theta_K \sin^2\theta_l \sin 2\varphi_h$
5	$\Re\{A_0^*(t)A_\parallel(t)\}$	$\Re\{\bar{A}_0^*(t)\bar{A}_\parallel(t)\}$	$\frac{1}{2}\sqrt{2}\sin 2\theta_K \sin 2\theta_l \cos\varphi_h$
6	$\Im\{A_0^*(t)A_\perp(t)\}$	$\Im\{\bar{A}_0^*(t)\bar{A}_\perp(t)\}$	$-\frac{1}{2}\sqrt{2}\sin 2\theta_K \sin 2\theta_l \sin\varphi_h$

The differential decay rate (Eqs. 1.32-1.33) is invariant under the transformation

$$(\phi_s, \Delta\Gamma_s, \delta_\parallel, \delta_\perp, \delta_0) \to (\pi - \phi_s, -\Delta\Gamma_s, -\delta_\parallel, \pi - \delta_\perp, -\delta_0) \tag{1.34}$$

and thus gives rise to a twofold ambiguity in the measured results. The possible of resolving this ambiguity is discussed in the next section.

1.3.3 S-wave contribution

The differential decay rates given in the previous section assume that the detected K^+K^- system in the final state originates solely from the decay of the ϕ resonance. Since the ϕ is a vector meson, the K^+K^- system is in a P-wave configuration. In addition to this contribution, there is, however, the possibility that the detected K^+K^- results from a non-resonant ($l = 0$) contribution or the $f_0(980)$ decay which is a scalar meson [31]. In both cases the K^+K^- system would be in a S-wave configuration. These decays represent an irreducible contribution to the final state $l^+l^-K^+K^-$ of the signal decay. The contribution to the differential decay rates can be described by the introduction of the S-wave amplitude $A_S(t)$ with phase δ_S. The S-wave amplitude can also interfere with the P-wave amplitudes. Hence, the Eqs. 1.32 and 1.33 need to be modified to

$$\frac{\mathrm{d}\Gamma(B_s^0 \to J/\psi K^+K^-)}{\mathrm{d}t\mathrm{d}\Omega} \propto \sum_{i=1}^{10} A_i(t) \cdot f_i(\cos\theta_K, \cos\theta_l, \varphi_h), \tag{1.35}$$

$$\frac{\mathrm{d}\Gamma(\bar{B}_s^0 \to J/\psi K^+K^-)}{\mathrm{d}t\mathrm{d}\Omega} \propto \sum_{i=1}^{10} \bar{A}_i(t) \cdot f_i(\cos\theta_K, \cos\theta_l, \varphi_h). \tag{1.36}$$

The additional amplitude combinations $A_i(t) = \{7...10\}$ are given by [32] as

$$\begin{aligned}
|A_S(t)|^2 &= |A_S(0)|^2 e^{-\Gamma_s t}\left[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right) - \sin\phi_s \sin(\Delta m_s t)\right],\\
\Re\{A_S^*(t)A_\parallel(t)\} &= |A_S(0)||A_\parallel(0)|e^{-\Gamma_s t}[-\sin(\delta_\parallel - \delta_S)\sin\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad + \cos(\delta_\parallel - \delta_S)\cos(\Delta m_s t) - \sin(\delta_\parallel - \delta_S)\cos\phi_s \sin(\Delta m_s t)],\\
\Im\{A_S^*(t)A_\perp(t)\} &= |A_S(0)||A_\perp(0)|e^{-\Gamma_s t}\sin(\delta_\perp - \delta_S)[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \cos\phi_s \sin(\Delta m_s t)\\
&\quad - \sin\phi_s \sin(\Delta m_s t)],\\
\Re\{A_S^*(t)A_0(t)\} &= |A_S(0)||A_0(0)|e^{-\Gamma_s t}[-\sin(\delta_0 - \delta_S)\sin\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad + \cos(\delta_0 - \delta_S)\cos(\Delta m_s t) - \sin(\delta_0 - \delta_S)\cos\phi_s \sin(\Delta m_s t)]
\end{aligned}$$

and the combinations $\bar{A}_i(t) = \{7...10\}$ as

$$\begin{aligned}
|\bar{A}_S(t)|^2 &= |\bar{A}_S(0)|^2 e^{-\Gamma_s t}\left[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \cos\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \sin\phi_s \sin(\Delta m_s t)\right],\\
\Re\{\bar{A}_S^*(t)\bar{A}_\parallel(t)\} &= |\bar{A}_S(0)||\bar{A}_\parallel(0)|e^{-\Gamma_s t}[-\sin(\delta_\parallel - \delta_S)\sin\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad - \cos(\delta_\parallel - \delta_S)\cos(\Delta m_s t) + \sin(\delta_\parallel - \delta_S)\cos\phi_s \sin(\Delta m_s t)],\\
\Im\{\bar{A}_S^*(t)\bar{A}_\perp(t)\} &= |\bar{A}_S(0)||\bar{A}_\perp(0)|e^{-\Gamma_s t}\sin(\delta_\perp - \delta_S)[\cosh\left(\tfrac{\Delta\Gamma_s}{2}t\right) + \cos\phi_s \sin(\Delta m_s t)\\
&\quad + \sin\phi_s \sin(\Delta m_s t)],\\
\Re\{\bar{A}_S^*(t)\bar{A}_0(t)\} &= |\bar{A}_S(0)||\bar{A}_0(0)|e^{-\Gamma_s t}[-\sin(\delta_0 - \delta_S)\sin\phi_s \sinh\left(\tfrac{\Delta\Gamma_s}{2}t\right)\\
&\quad - \cos(\delta_0 - \delta_S)\cos(\Delta m_s t) - \sin(\delta_0 - \delta_S)\cos\phi_s \sin(\Delta m_s t)].
\end{aligned}$$

In presence of a S-wave contribution the magnitude of the amplitudes needs to fulfill a new normalization condition $|A_0(0)|^2 + |A_\perp(0)|^2 + |A_\parallel(0)|^2 + |A_S(0)|^2$=1. The angular dependent terms $f_i = \{7...10\}$ are given in Table 1.4. The S-wave strong phase is also invariant under the transformation (Eq. 1.34): $\delta_S \to -\delta_S$. The different $m(K^+K^-)$ dependence of P- and S-wave amplitudes would single out one physically correct solution of Eq. 1.34 and thus resolves the twofold ambiguity in the differential decay rate [32]. The correct parameter set has been found in the $B_s^0 \to J/\psi(\mu^+\mu^-)K^+K^-$ decay analysis [28] and this solution is implemented in the presented thesis.

Table 1.4 Angular dependent terms $f_i(\cos\theta_K, \cos\theta_l, \varphi_h)$ for the S-wave [32].

i	$A_i(t)$	$\bar{A}_i(t)$	$f_i(\cos\theta_K, \cos\theta_l, \varphi_h)$
7	$\lvert A_S(t)\rvert^2$	$\lvert\bar{A}_S(t)\rvert^2$	$\frac{2}{3}\sin^2\theta_l$
8	$\Re\{A_S^*(t)A_\parallel(t)\}$	$\Re\{\bar{A}_S^*(t)\bar{A}_\parallel(t)\}$	$\frac{1}{3}\sqrt{6}\sin\theta_K \sin 2\theta_l \cos\varphi_h$
9	$\Im\{A_S^*(t)A_\perp(t)\}$	$\Im\{\bar{A}_S^*(t)\bar{A}_\perp(t)\}$	$-\frac{1}{3}\sqrt{6}\sin\theta_K \sin 2\theta_l \sin\varphi_h$
10	$\Re\{A_S^*(t)A_0(t)\}$	$\Re\{\bar{A}_S^*(t)\bar{A}_0(t)\}$	$\frac{4}{3}\sqrt{3}\cos\theta_K \sin^2\theta_l$

1.3.4 Beyond the SM contribution to B mixing

Physical processes beyond the SM (New Physics) can manifest themselves by introducing additional contributions to $B_q^0 - \bar{B}_q^0$ meson mixing ($q = d, s$), for example through heavy virtual particles contributing to the mixing diagram. The effect on matrix element M_{12} in

the Hamiltonian (Eq. 1.16) can be parametrized in a model-independent way by introducing complex parameters Δ_q [33, 34]

$$M_{12}^q = M_{12}^{q,\mathrm{SM}} \Delta_q \quad \text{with} \quad \Delta_q = |\Delta_q| e^{i\phi_q^\Delta}.$$

The SM hypothesis is $\Delta_d = \Delta_s = 1$. If the phase ϕ_s^Δ is nonzero this would directly modify the phase of λ_f leading to the measurement of

$$\phi_s = \phi_M - 2\phi_D \to \phi_s^{\mathrm{SM}} + \phi_s^\Delta. \tag{1.37}$$

Assuming the independence of Δ_d and Δ_s a global fit of the CKM matrix elements (Sec. 1.2.1) and the parameters Δ_d and Δ_s describing New Physics (NP) contributions in the B sector was performed in Ref. [34]. Fig. 1.9 shows the results of the fit for the parameters Δ_d and Δ_s in the complex plane. The fit disfavours the SM prediction for B_d^0 meson mixing, $\Delta_d = 1$, by 3.0 standard deviations. For B_s^0 meson mixing, no deviation from the SM value $\Delta_s = 1$ is observed. The combined probability of $\Delta_d = \Delta_s = 1$ is found to be equivalent to 2.4σ [34]. These hints at possible problems of the SM description of B meson mixing represent a strong motivation for a more precise determination of ϕ_s.

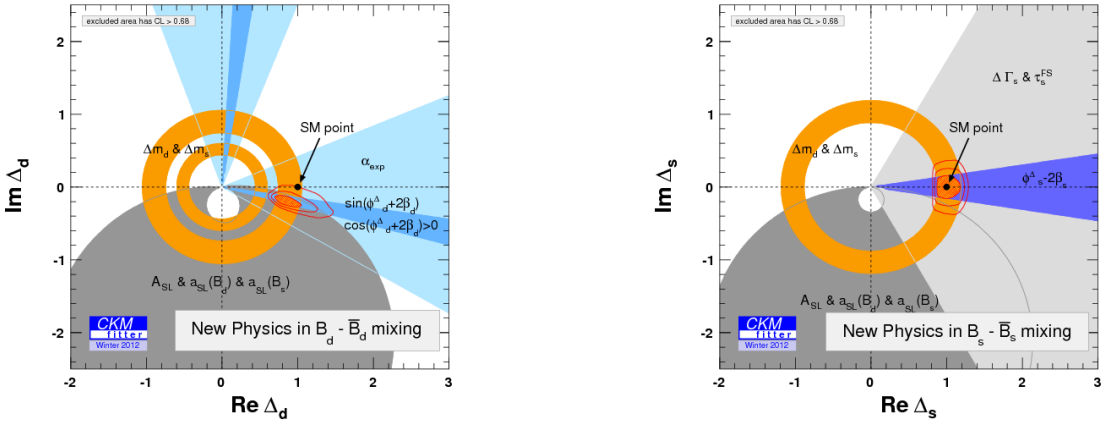


Fig. 1.9 Global fit of possible New Physics contributions to (left) $B_d^0 - \bar{B}_d^0$ meson mixing and (right) $B_s^0 - \bar{B}_s^0$ meson mixing. The figures are taken from [34].

1.3.5 Current experimental status of ϕ_s

The phase ϕ_s has been measured in eight analyses: four $B_s^0 \to J/\psi\phi$ analyses from CDF [35], D$\emptyset$ [36], ATLAS [37] and CMS [38], and four analyses from LHCb of the decay modes $B_s^0 \to J/\psi K^+K^-$ [39, 40], $B_s^0 \to J/\psi\pi^+\pi^-$ [41], $B_s^0 \to \psi(2S)\phi$ [42] and $B_s^0 \to D_s^+ D_s^-$ [43]. A combined two-dimensional fit of ϕ_s and $\Delta\Gamma_s$, without external assumption, yields two symmetric solutions related through $\phi_s \leftrightarrow \pi - \phi_s$ and $\Delta\Gamma_s \leftrightarrow -\Delta\Gamma_s$.

Fig. 1.10 shows, in the (ϕ_s, $\Delta\Gamma_s$) plane, the individual 68% confidence level contours of ATLAS, CMS, CDF, DØ and LHCb, their combined contour (white solid line and shaded area), as well as the SM predictions (very thin black rectangle). The prediction for ϕ_s is taken as the indirect determination of $-2\beta_s$ via a global fit to experimental data within the Standard Model, $-2\beta_s = -0.0367^{+0.0007}_{-0.0008}$ rad [18], while the SM prediction for $\Delta\Gamma_s = 0.088 \pm 0.020$ ps^{-1} [27]. The combined result, $-2\beta_s = -0.021 \pm 0.031$ rad and $\Delta\Gamma_s = 0.085 \pm 0.006$ ps^{-1}, is consistent with these predictions. The uncertainties of the single measurements are still large, allowing for possible contributions of NP to the phase ϕ_s.

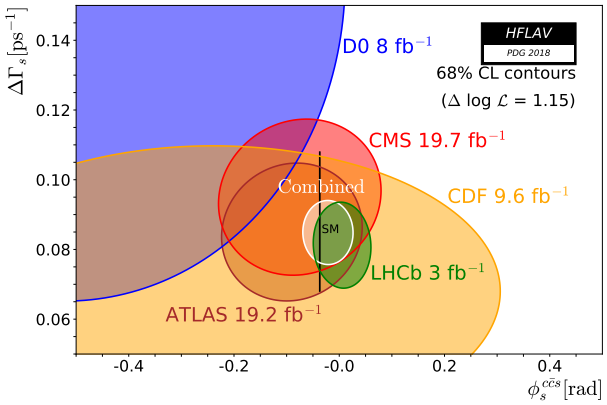


Fig. 1.10 68% confidence level regions in $\Delta\Gamma_s$ and ϕ_s plane obtained from individual contours of CDF [35], DØ [36], ATLAS [37], CMS [38] and LHCb [39] measurements and the combined contour (solid line and shaded area). The expectation from the SM [18] is shown as a black thin rectangle. The figure is taken from [20].

Chapter 2

LHCb experiment

The LHCb (Large Hadron Collider beauty) detector is one of the four large particle physics experiments at the Large Hadron Collider. It is especially designed to measure decays of b- and c-hadrons in search for $\mathcal{CP}$ violation and rare B meson decays. In the following chapter the running conditions of the Large Hadron Collider and the production mechanism of b-quarks in proton-proton collisions are briefly discussed. The working principles and technical details of the various LHCb detector components are summarized. At the end of the chapter, the measured and simulated event samples used in the thesis are introduced.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [44] is a proton-proton (pp) collider located at CERN (Geneva, Switzerland). In a 27 km long underground tunnel, two proton beams are accelerated and brought to collision at four interaction points which house the four major LHC experiments: ATLAS, CMS, ALICE and LHCb. Fig. 2.1 shows the main accelerator ring with the four experiments. ATLAS and CMS are general purpose detectors, whereas ALICE is specialized for heavy ion collisions and LHCb is an experiment dedicated to measurements of b- and c-hadron decays.

The design centre-of-mass collision energy of the LHC is $\sqrt{s}$ = 14 TeV. At a design luminosity of $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$, the proton beams are separated into 2808 bunches, each containing $\sim 10^{11}$ protons. With a bunch spacing of 25 ns, this gives an interaction rate of 40 MHz. For the LHCb experiment the design luminosity is $\mathcal{L} = (2\text{-}5)\times 10^{32}\text{cm}^{-2}\text{s}^{-1}$ in order to limit the number of interactions per bunch-crossing.

In the 2011 and 2012 runs of LHC, the protons were collided with an energy of $\sqrt{s} = 7$ and 8 TeV, respectively. The instantaneous luminosity delivered to LHCb is $\mathcal{L} = 4\times 10^{32}\text{cm}^{-2}\text{s}^{-1}$. The integrated luminosity recorded by the LHCb experiment is $\mathcal{L}_{\text{int}} = 1.0\ \text{fb}^{-1}$ in 2011 and $\mathcal{L}_{\text{int}} = 2.0\ \text{fb}^{-1}$ with the higher energy in 2012 that consist of total $\mathcal{L}_{\text{int}} = 3.0\ \text{fb}^{-1}$ for two years, called Run 1. After long shutdown in 2013-2014, the colliding energy of the protons has been increased to $\sqrt{s}$= 14 TeV for upgraded LHC experiments. The total integrated

luminosity recorded by the LHCb during 2015-2018 is $\mathcal{L}_{\text{int}} = 5.7\ \text{fb}^{-1}$ called Run 2 data. The presented thesis is focused on the data collected during 2011-2012 runs.

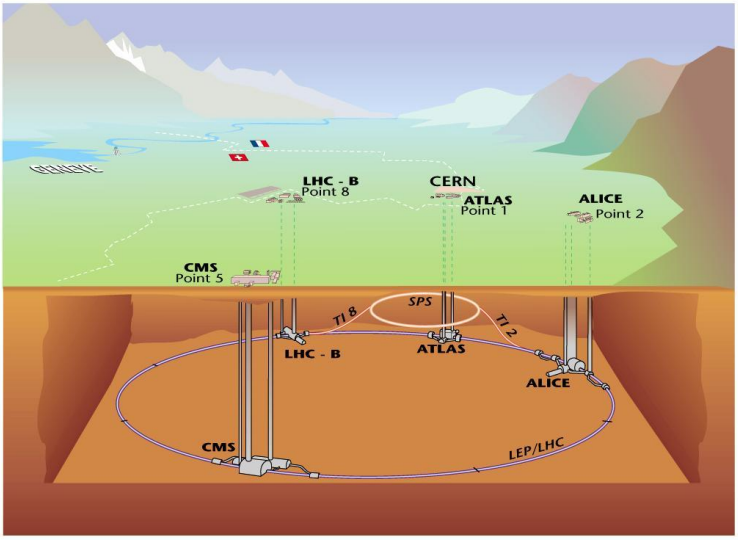


Fig. 2.1 Schematic overview of the LHC accelerator ring and the location of the four particle detectors ATLAS, CMS, ALICE and LHCb. The figure is taken from [45].

2.2 b-hadron production at the LHC

At the LHC, b-quarks are predominantly produced in inelastic pp collisions in pairs of quark and anti-quark ($b\bar{b}$). The dominant process is the fusion of two gluons (gg) or two quarks ($q\bar{q}$). Fig. 2.2 shows the leading-order Feynman diagrams of the production processes.

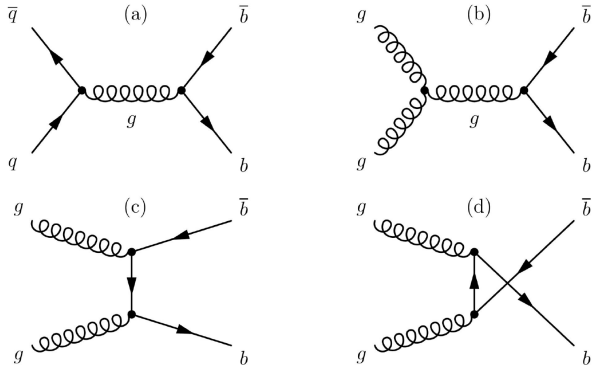


Fig. 2.2 Feynman diagrams of the dominant $b\bar{b}$ production processes in pp interactions at the LHC. The figure is taken from [46].

The production cross section for $b\bar{b}$ pairs depends on the polar angle to the beam axis. The reason for this is, compared with the beam energy, the low mass of the $b\bar{b}$ system in conjunction with the parton distribution functions of quarks and gluons in the colliding protons. Since the quark and anti-quark (or two gluons) which produce the $b\bar{b}$ system can

carry very different momentum fractions of the proton, the $b\bar{b}$ system will likely be boosted in either forward or backward direction. Fig. 2.3 shows the polar angle distribution of produced $b\bar{b}$ pairs, which clearly peaks at small and high polar angles. To maximize the $b\bar{b}$ production in the detector acceptance, the LHCb detector is built as a single-arm forward spectrometer. The details of the detector are discussed in the next sections.

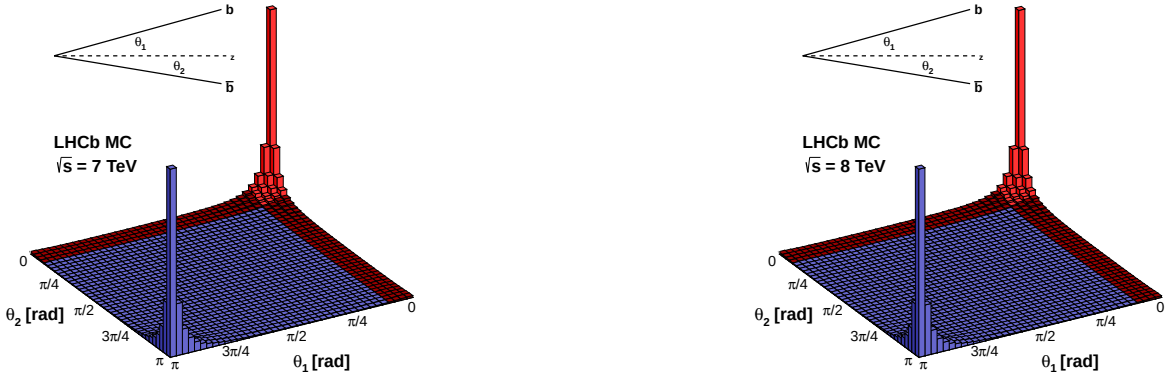


Fig. 2.3 Simulated two-dimensional polar angle distribution of produced $b\bar{b}$ pairs in inelastic pp collisions at a center-of-mass energy of (left) $\sqrt{s} = 7$ TeV and (right) 8 TeV. Both b and $\bar{b}$ quark are either produced in forward or in backward direction. The red region marks the LHCb acceptance. The figures are taken from [47].

The total $b\bar{b}$ cross section in inelastic pp collisions at a center-of-mass energy of $\sqrt{s} =$ 7 and 8 TeV was measured at the LHCb using J/ψ meson from B hadron decays to be $\sigma(pp \to b\bar{b}X) = 288\pm4\pm48$ μb [48] and $\sigma(pp \to b\bar{b}X) = 298\pm2\pm36$ μb [49], respectively. With a recorded integrated luminosity $\mathcal{L}_{\text{int}}^{2011} = 1.0$ fb^{-1} and $\mathcal{L}_{\text{int}}^{2012} = 2.0$ fb^{-1} the number of produced $b\bar{b}$-pairs can be calculated to be $N_{b\bar{b}} = \sigma(pp \to b\bar{b}X) \cdot \mathcal{L}_{\text{int}} \approx 3{\times}10^{11}$ in 2011 run and two times larger in 2012 run of the LHC. Approximately 25% of them are produced within the LHCb detector acceptance.

2.3 LHCb detector

The LHCb detector [50] is designed as a single-arm forward spectrometer to cover the optimal acceptance for b-hadron decays. The layout of the detector is shown in Fig. 2.4. The coordinate system is defined by the z-axis along the beam line and the y-axis in the vertical direction. The acceptance coverage is 10-300 mrad in the bending plane (x-direction) and 10-250 mrad in the non-bending plane (y-direction) of the dipole magnet. The different subdetectors are, starting at the left closest to the interaction point:

- The Vertex Locator (VELO) is the LHCb's silicon vertex tracker which is built around the pp interaction point. The VELO resolves the primary interaction vertex and decay vertices of long-lived particles with high precision;

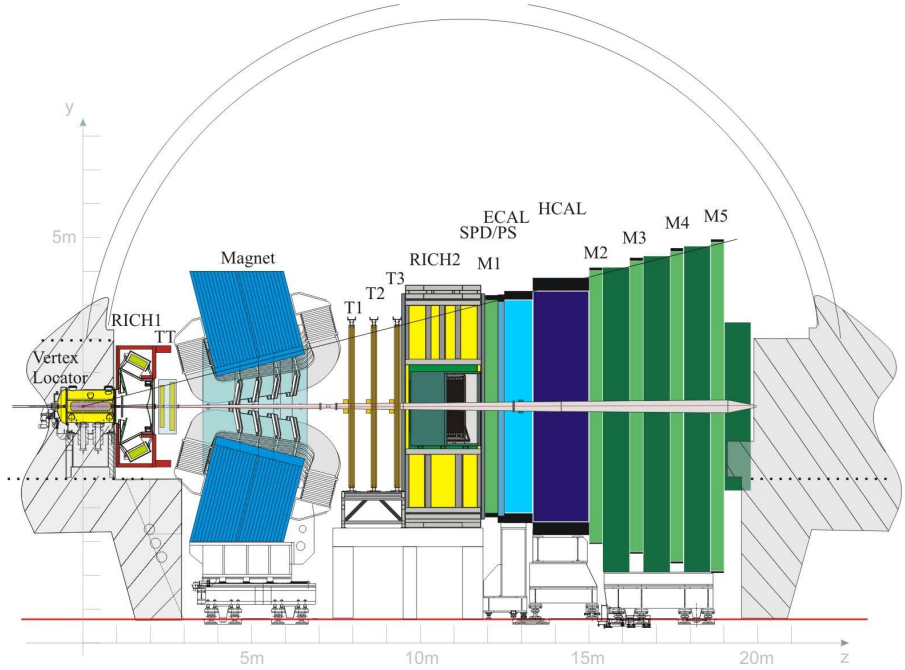


Fig. 2.4 A side view of the LHCb detector. The *pp* collisions occur to the very left in LHCb's silicon vertex tracker, VELO. Downstream of the VELO, the RICH1 and TT detectors are placed in front of the magnet. The main tracking stations, consisting of IT and OT are located behind the magnet. Finally, particles traverse the RICH2, ECAL and HCAL detectors and muon chambers.

- The first ring imaging Cherenkov detector (RICH1) for particle identification follows behind the VELO;
- The Tracker Turicensis (TT), the first tracking station, is located in front of the magnet;
- The main tracker consists of three tracking stations located behind the magnet. The Outer Tracker (OT) covers the largest fraction of the acceptance. The Inner Tracker (IT) covers the high occupancy region close to the beam pipe;
- The RICH2 detector, the second ring imaging Cherenkov detector, is located behind the main tracker and is used for particle identification;
- The calorimetry consists of two subsystems, the electromagnetic (ECAL) and hadronic (HCAL) calorimeter. The calorimeters measure the energy deposited by electromagnetic and hadronic showers, respectively;
- The rightmost subsystem is the muon chambers called M1-M5 in Fig. 2.4. The muon chambers are used for muon identification and are essential for the muon trigger system.

All subsystems are briefly reviewed in the next sections. A more detailed overview of the LHCb detector is given in Refs. [50, 51].

2.3.1 Tracking System

The main purpose of the tracking system is the momentum measurement of charged particles whose trajectories are bent in the magnetic field of the dipole magnet. Their trajectories are measured before the magnet with the VELO and TT and after the magnet with three main tracking stations of the IT and OT. The momentum can be determined from the deflection of the particle trajectory. The magnetic field is designed to be almost homogeneous, with a large component in y-direction and small components in x- and z-direction. The particles moving in z-direction are therefore bent mainly in the $x-z$ (horizontal) plane. Fig. 2.5 shows the field strength as a function of the z-coordinate. Integrated over a length of $l = 10$ m the field strength is $\int B\mathrm{d}l = 4$ Tm. As indicated in Fig. 2.5, the orientation of the magnetic field can be switched, giving a powerful tool to measure charge dependent detector asymmetries.

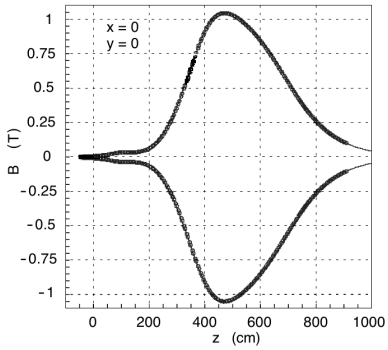
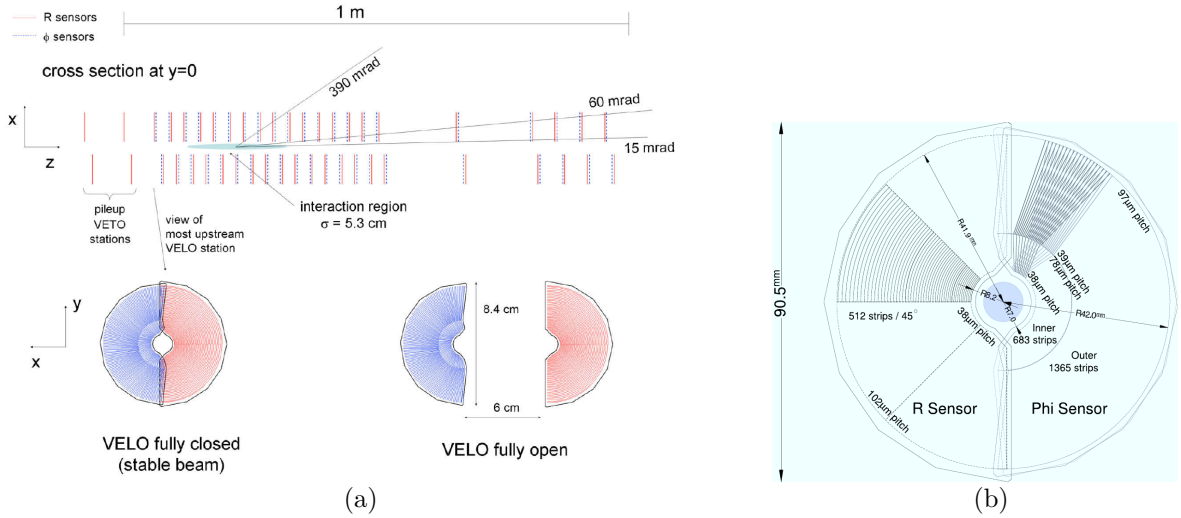


Fig. 2.5 Field map of the LHCb dipole magnet. The polarity of the magnetic field can be switched from down (positive curve) to up (negative curve). The figure is taken from [50].

Vertex detector

The Vertex Locator is a silicon strip detector positioned closely around the pp interaction point to measure precisely the positions of the primary interaction and secondary displaced vertices. It is built of 21 stations (Fig. 2.6a) with each containing a pair of silicon modules with half disk shape. The modules have an overlap in the x-direction and are mounted in 2 mm intervals in the z-direction. They consist of two different types of 300 μm thick sensors mounted back to back. The r-sensors measure the radial distance of a particle track to the beam axis and are made of circular strips around the beam. The ϕ-sensors are made of straight radial strips and measure the polar angle of the tracks. Fig. 2.6b illustrates the $r-\phi$ geometry of the sensors. Additional two pile-up stations are located upstream of the VELO consisting of two r-sensor modules. They are used in the hardware trigger to detect beam-gas interactions.

The precision of the reconstructed vertices depends on the extrapolation of the measured track positions. Therefore, the VELO modules are placed close to the interaction point. The sensitive regions of the modules starts in 8 mm distance from the beam line. To protect the



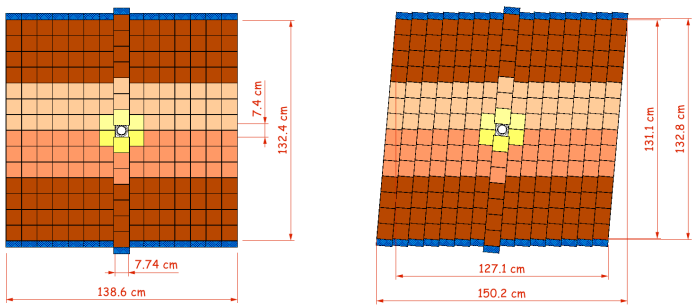
(a) (b)

Fig. 2.6 (a) Cross section of the VELO in $x-z$ plane showing the 21 stations with r- and ϕ-sensors. The pile-up stations are only used in the trigger. (b) Illustration of the r- and ϕ-sensors of the VELO modules. Both figures are taken from [50].

sensors from excessive radiation damage during unstable beam conditions, the half modules can be moved away from the beam line.

Tracker Turicensis

The Tracker Turicensis is a silicon strip detector located upstream of the dipole magnet. It consists of two 150 cm wide and 130 cm high stations with a detection area of about 8.4 m^2. Each station has two detection layers, where the strips are arranged in (x,u)-layers for the first station and in (v,x)-layers for the second station. The strips in the x-layers are vertical, while they are rotated in the u- and v-layers by a stereo angle of $\pm 5°$ (Fig. 2.7). This allows for a three-dimensional track reconstruction with the best precision in the horizontal bending plane of the magnet. The strips are 500 μm thick with a pitch of 183 μm. The single hit resolution of the TT is about 50 μm.



(a) x-layer of TT (b) u-layer of TT

Fig. 2.7 Layout of different TT detection layers. The different colours correspond to different readout sectors. The figure is taken from [52].

Inner Tracker

The Inner Tracker is a silicon strip detector located in the center of three tracking stations after the dipole magnet. It covers a 120 cm wide and 40 cm high cross shaped region around the beam line with an active area of about 4 m^2 (Fig. 2.8). Each station is built of four layers where the strips are arranged in a (x, u, v, x) geometry. Similar to the TT, the strips in the x-layers are vertical but rotated by $\pm 5°$ in the u- and v-layers. The detector is arranged in four boxes: left (ASide) and right (CSide) boxes, top (Top) and bottom (Bottom) boxes as shown in Fig. 2.8. The strip geometry was chosen to limit the maximum hit occupancy per sensor to a few percents. The pitch between the sensors is about 200 μm leading to a single hit resolution of 50 μm, similar to the TT.

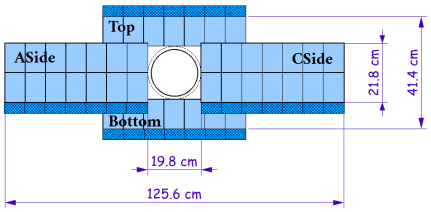


Fig. 2.8 Front view of an IT x-layer (station 2) indicating four different detector boxes. The dark blue regions mark the readout electronics. The figure is taken from [52].

Outer Tracker

The Outer Tracker is a straw tube drift-time detector consisting of three stations after the dipole magnet. The three stations are located to cover the large area outside the IT acceptance as shown in Fig. 2.9. The stations are of equal size with the outer boundary corresponding to an acceptance of 300 mrad in the horizontal and 250 mrad in the vertical plane. They consist of four layers in the same (x, u, v, x) layout as the IT. The layers are built as arrays of modules, each module containing a double layer of straw tubes. The inner diameter of the straws is 4.9 mm, the pitch between them is 5.25 mm (Fig. 2.9). The spatial resolution of single straws is 200 μm. The light materials in the OT result in a radiation length below 10%.

2.3.2 Ring Imaging Cherenkov Detectors

The LHCb experiment uses ring imaging Cherenkov counters to distinguish between different charged particle species. When charged particles traverse material with refraction index n, with a higher velocity than the velocity of light in the material ($c' = c/n$), they emit photons under the Cherenkov angle θ_{Ch} with

$$\cos\theta_{\mathrm{Ch}} = \frac{1}{n\beta},$$

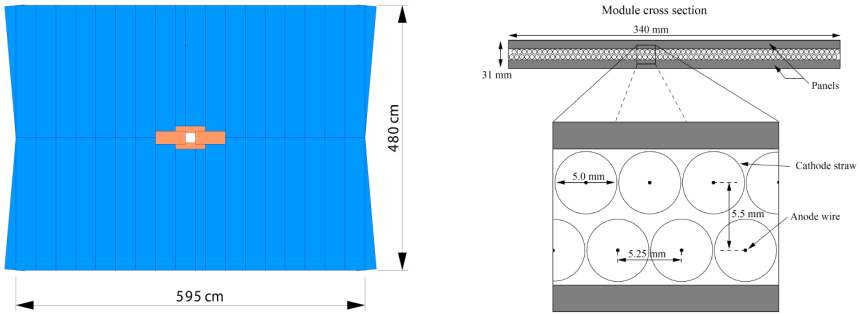


Fig. 2.9 (left) Front-view of an OT layout where the inner (orange) part is the IT. (right) Cross section of an OT module and straw tubes. The figures are taken from [50].

where $\beta = v/c$ is the velocity of the charged particle. Together with the momentum measurement of the particle performed by the tracking system this allows to calculate the particle mass and therefore to determine the particle type. Fig. 2.10 shows the Cherenkov angle depending on particle momentum for different particle species and different radiators (silica aerogel, C_4F_{10} gas, CF_4 gas).

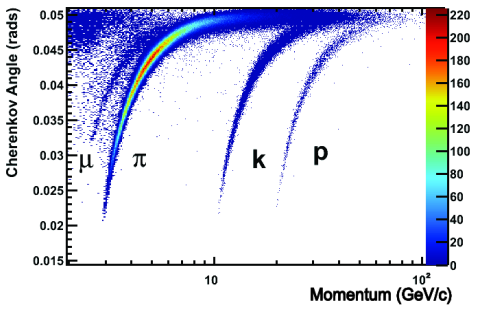


Fig. 2.10 Reconstructed Cherenkov angle as a function of particle momentum for different particle species. The figure is taken from [50].

To guarantee a good $K-\pi$ separation over the full momentum range from 2 to 100 GeV/c, the LHCb uses three different radiators in two Cherenkov detectors. One of the two detectors, RICH1 (Fig. 2.11), is located before the magnet. It uses two radiators, silica aerogel (n=1.03) and C_4F_{10} gas (n=1.0014), and is responsible for particle identification at low momenta from ~1 to 60 GeV/c.

The second RICH detector, RICH2 (Fig. 2.11), is located behind the magnet and tracking stations and is responsible for the identification of particles with momenta from 15 GeV/c up to 100 GeV/c and beyond. It contains CF_4 (n=1.0005) as a gas radiator. Whereas the RICH1 detector covers the full LHCb acceptance of 25-300 mrad (25-250 mrad) in the horizontal (vertical) plane, RICH2 covers only an angular acceptance of 15-120 mrad (15-100 mrad) horizontally (vertically).

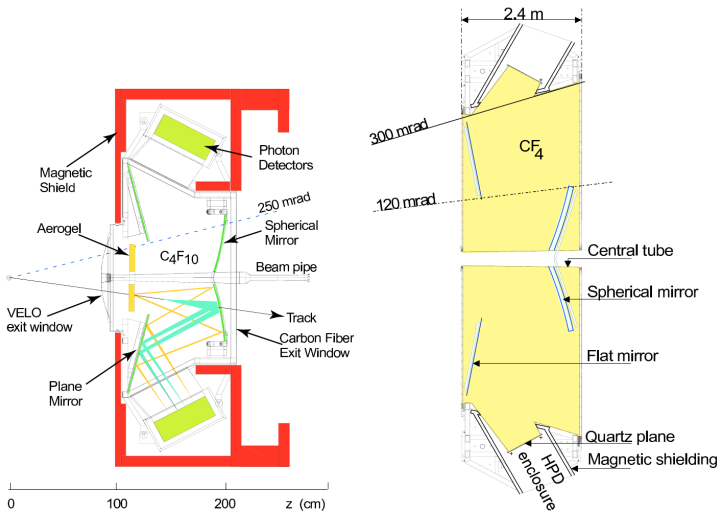


Fig. 2.11 Schematic view of (left) RICH1 and (right) RICH2. The figures are taken from [50].

Charged hadron identification

A likelihood approach [53] is used to identify different types of particles from the measured Cherenkov photons. For every track the Cherenkov angle for all the possible particle species is calculated (π, K, p, etc.) using the known track momentum. A likelihood is created by calculating the probability for every detected Cherenkov photon to have been emitted under the reconstructed angle given the specific track and the particle hypothesis. Of particular interest in many analyses is the difference between the logarithm of the likelihoods for different particle species, an example would be the variable

$$\Delta\ln\mathcal{L}_{K\pi} = \ln\mathcal{L}_K - \ln\mathcal{L}_\pi$$

which is used to separate pions from kaons. This variable is important for the study of final states with charged kaons, where, without particle identification, one would be overwhelmed by the pion background.

The RICH detectors provide an excellent $K-\pi$ separation over the full momentum range of 2-100 GeV/c with an average kaon identification efficiency of 95% and a π-as-K misidentification rate of 5% [50].

2.3.3 Calorimeter System

The LHCb calorimeter system measures the size and the position of energy depositions and performs electron, photon and hadron identification. Furthermore, the calorimetry is used in the first level trigger (Sec. 2.3.5). This puts stringent requirements on the readout speed of the calorimeters, since the L0 decision needs to occur after only 4 μs.

Charged incident particles and photons can deposit energy in the electromagnetic calorimeter by producing electromagnetic showers via bremsstrahlung and pair production. Hadrons

can produce hadronic showers in the hadronic calorimeter. The particles of the shower produced in the absorber material traverse active scintillating material which they are able to excite. Photons emitted during the de-excitation process can then be read out via wavelength shifting fibers and detected with photomultiplier tubes (PMTs). The LHCb's calorimeter system consists of the following subsystems:

- Scintillating Pad Detector (SPD) used for the separation of electrons and photons;
- Preshower (PS) follows the SPD after 15 mm of lead absorber. It is used to distinguish between electrons and charged pions;
- Electromagnetic calorimeter (ECAL) which measures the energy and position of showers from electrons and photons;
- Hadronic calorimeter (HCAL) which measures the energy and shower shape of hadrons.

Scintillating pad detector and preshower detectors

The SPD and PS are located behind the first muon station (M1) on the two opposite sides of a 15 mm thick lead absorber (Fig. 2.4).

The purpose of the SPD detector is to separate electrons from photons. The principle behind the electron-photon separation is that while electrons will give a signal in the SPD the electrically neutral photons will not. The photon as electron misidentification rate is found to be below 3%.

The PS detector is built for electron-pion separation. The deposited energy for 50 GeV electrons and pions is shown in Fig. 2.12. In tests with e/π beams the PS showed pion rejection rates of 99.6%, 99.6%, and 99.7% with electron retention rates of 91%, 92% and 97% for particles with momentum 10, 20 and 50 GeV/c, respectively.

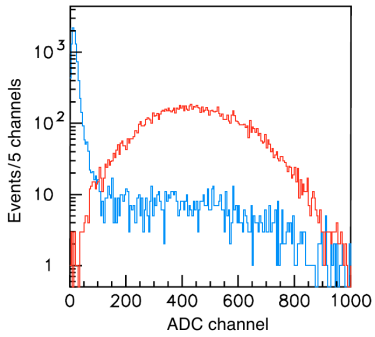


Fig. 2.12 Energy deposition (in arbitrary units) in the PS detector for (red line) 50 GeV electrons and (blue line) 50 GeV pions. The figure is taken from [50].

Electromagnetic calorimeter

The electromagnetic calorimeter is located behind the PS and built as a so called "shashlyk" calorimeter. It is a sampling structure built from alternating layers of 2 mm lead absorber and 4 mm scintillator material read out via wavelength shifting fibers. The ECAL has a length corresponding to 25 radiation lengths X_0 and 1.1 hadronic interaction lengths λ_I. The energy resolution of the ECAL is given by

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

where the resolution term, on the right-hand side, should be added in quadrature.

Hadronic calorimeter

The HCAL is located after the ECAL and uses an iron/scintillator sampling structure. In total the HCAL has a depth corresponding to 5.6 hadronic interaction lengths λ_I. The energy resolution of the HCAL is given by

$$\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E(\text{GeV})}} \oplus 10\%.$$

2.3.4 Muon System

The muon chambers are essential for muon identification and triggering of B meson decays containing muons in the final state. The LHCb muon system consists of five muon stations (M1-M5) as shown in Fig. 2.13. The M1 is located upstream of the calorimeters to improve the transverse momentum resolution in the muon trigger by minimizing uncertainties caused by multiple scattering in the calorimeter material. The 80 cm thick iron absorbers are placed between the muon stations M2-M5 located after the calorimeter to filter out all particles except muons. To traverse all muon chambers and interleaved absorbers, muons need to have a minimum momentum of 6 GeV/c. The muon stations are built of multiwire proportional chambers (MWPC) with an exception for the region of high particle flux around the beam pipe (R1) of station M1 where gas electron multiplier (GEM) detectors are used. Both detector types are able to collect the signal in less than 20 ns with an efficiency larger than 95%.

Muon identification

A track which is considered as a muon candidate is extrapolated into the muon system. To be confirmed as muon candidate, a certain number of hits in the muon chambers need to be found in a field of interest around the track. In addition, a likelihood [54] is created using the

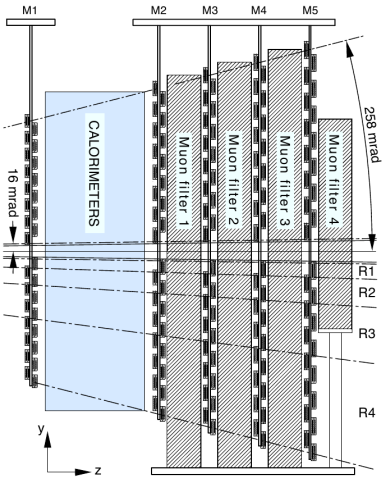


Fig. 2.13 Overview of the six muon stations. The figure is taken from [50].

distance of the hits from the extrapolated track under both the muon and pion hypothesis

$$\Delta \ln \mathcal{L}_{\mu\pi} = \ln \mathcal{L}_{\mu} - \ln \mathcal{L}_{\pi}.$$

It can be used to differentiate between the two particle types. The LHCb muon system exhibits a high muon identification efficiency of 97% and a π-as-μ misidentification rate of (1-3)% [50].

2.3.5 Trigger System

The LHCb uses a multi-level trigger system to reduce the bunch crossing rate of 40 MHz to a data rate of about 2.2 kHz which is written to tape for later analyses. It therefore needs to reject events that are of no interest in the LHCb physics program but at the same time retain events containing decays of B and D mesons. The LHCb trigger system consists of three levels: Level 0 (L0), High Level Trigger 1 (HLT1) and High Level Trigger 2 (HLT2) [51, 55] as overviewed in Fig. 2.14. The trigger system is designed to be very configurable to be

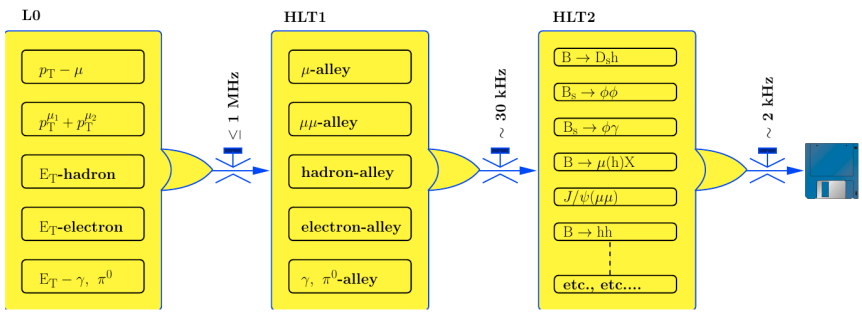


Fig. 2.14 Three trigger levels of the LHCb trigger system. The figure is taken from [50].

able to adapt to different running conditions. A unique Trigger Configuration Key (TCK) is

assigned to identify every trigger configuration.

Level 0 Trigger

The L0 trigger is completely implemented in hardware to be able to cope with the high interaction rate. It takes advantage of the somewhat higher transverse momenta of particles from B meson decays relative to QCD events which in general show a "softer" p_{T} distribution. The L0 trigger consists of three subsystems:

- The *calorimeter trigger* sums up the transverse energy E_{T} of an array of 2×2 calorimeter cells to form calorimeter clusters. A particle hypothesis (e, γ or hadron) is assigned to the clusters using information from the SPD, PS, ECAL and HCAL (Sec. 2.3.3). The trigger then selects the hadron, electron and photon cluster with the largest transverse energy in the event.

- The *muon trigger* tries to reconstruct, for each quadrant of the detector (Sec. 2.3.4), two muons with the largest transverse momentum p_{T}. Assuming the muon tracks originate from the interaction point, the p_{T} of the candidates can be estimated using the track slope in the first two muon stations[1]. The muon stand-alone reconstruction exhibits a momentum resolution of ∼25%.

- The *pileup system* upstream of the VELO (Sec. 2.3.1) can be used to veto events with high backwards activity indicating multiple interactions.

The L0 decision unit combines the different trigger subsystems and gives the final trigger decision. The L0 trigger reduces the rate to 1 MHz with which the detectors are read out.

High Level Trigger

The HLT further reduces the event rate from 1 MHz to 2.2 kHz which is written to disk. The HLT consists of two stages:

- The HLT1 performs a partial event reconstruction. It tries to confirm the L0 decision by reconstructing tracks in the VELO and the tracking stations corresponding to the candidates found by the L0 trigger. This L0 confirmation reduces the event rate to ∼30 kHz.

- The HLT2 fully reconstructs tracks in the event in a manner very similar to the offline procedure. It performs various inclusive and exclusive selections. The inclusive selections try to partially reconstruct the final states of B meson decays, an example would be the reconstruction of b-hadron decays to (J/ψ + X). The exclusive selections are used for B →hadron decays where the final state is fully reconstructed. The HLT2 reduces the event rate to 2.2 kHz.

[1] The effect of the magnetic field is described by an effective bending plane.

2.3.6 LHCb Software

Several software packages are needed to perform an analysis with the data taken by the LHCb detector:

- Data that was triggered and written to disk by the event filter farm is available as so called raw data files. The first step in the analysis is to reconstruct physical quantities from the measurements of the detector subsystems. The **Brunel** software package [56] performs this reconstruction step. Brunel performs track finding and fitting. It also links tracks to the available particle identification information extracted from the calorimeter, muon and RICH subdetectors. Furthermore, Brunel reconstructs the energy of electromagnetic and hadronic showers using the calorimetry. At the end of the reconstruction step Data Summary Tape (DST) files are produced on which all further analyses is based.

- **DaVinci** [57] is the analysis software package of LHCb built on the Gaudi framework [58] used for reconstruction of the signal decays and data selection. To make the large amounts of data more manageable for physics analyses, the preselections are performed on the reconstructed data. This analysis step is also called "stripping". Similar preselections are combined to form so called streams for different physics topics. Examples for streams provided by the stripping are the hadron stream which selects hadronic decays of B mesons, the charm stream which selects decays of D mesons and the dimuon stream which is optimized for B decays containing two muons in the final state. The offline analysis is performed on the appropriate physics stream, again using the DaVinci analysis software. The full selection is applied using the signal candidates reconstructed in the stripping stage.

  Due to the large quantities of data that need to be processed at LHCb these tasks are performed on the Grid services [59] using the **Dirac** workload and data management system [60].

2.3.7 LHCb Monte Carlo Simulation Framework

To asses the feasibility of physics studies and to devise and test appropriate analysis strategies, the simulation of the physical processes occurring at the LHCb is essential. The method of Monte Carlo (MC) simulation is used to model both the particles emerging from the pp collision and their interaction with material. The full MC simulation is performed using the LHCb simulation framework which is based on the Gaudi framework [58] and consists of several steps:

- The MC simulated events are generated using the **Gauss** software package [61, 62]. For the simulation of the pp collision Gauss uses the `Pythia6` [63] or `Pythia8` [64] program,

a general purpose MC event generator. The `Pythia` contains routines for the hard processes (examples are the b production processes given in Sec. 1.3) in the pp collision as well as initial and final state radiation and subsequent hadronization. The decay of heavy B mesons is performed using the `EvtGen` [65] package, which is specialized on the physics of B meson decays, and `Photos++` [66] package describing a final state radiation. The particles produced in the pp collision finally decay to long-lived particles which traverse the detector. The simulation of their electromagnetic and hadronic interaction with material is performed by the `Geant4` toolkit [67].

- The output of the detector simulation stage is used as an input for **Boole** software package [68] which simulates the detector response to hits deposited in the detector by traversing particles.

- The **Moore** software package [69] is the high level trigger application of the LHCb. It can be run on MC simulation in practically identical fashion as on the event filter farm, accurately replicating the behavior of the trigger on real data.

- The reconstruction of the events is performed the same way as for real data using the **Brunel** software package (Sec. 2.3.6). For simulated data Brunel additionally performs the associations (matching) of reconstructed tracks and calorimeter clusters to the originally generated particles.

- The analyses are performed using the physics analysis software, **DaVinci**, analogous to the processing of real data (Sec. 2.3.6).

Chapter 3

Alignment of the Tracking Stations

The LHCb detector has been designed to achieve the physics goals discussed in Sec. 1. However, these measurements are only possible if the detector is built with a high precision, or if the actual position and performance of the detector can be known precisely. For example, if the VELO is not aligned correctly, the primary vertices are not reconstructed precisely and the time of flight of long-lived mesons cannot be measured correctly. The tracking stations provide a momentum estimate for the tracks that cross the detector. For low-multiplicity decays, this momentum estimate dominates the mass resolution. The LHCb tracking methods and alignment strategy are presented in this section. First, the track types going through the detector are described, followed by a discussion of the tracking methods used at the LHCb. Next, the global alignment strategy is discussed for each subdetector internal alignment strategy. Finally, the LHCb alignment framework used in particular for the alignment of the tracking stations is detailed.

3.1 Track reconstraction at LHCb

The pp collisions that occur in the LHCb yield a high track density. The average track multiplicity in the VELO is around 100 in a typical $b\bar{b}$ event. In this environment, finding the trajectory of the particles through the various subdetectors is a challenging task. The track reconstruction is divided into the pattern recognition (assignment of clusters to the tracks) and fitting (determining the optimal track parameters). The aim 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 for the determination of the b-hadron lifetime or flight distance.

3.1.1 Track types

Depending on their origin vertex and their momentum, the charged tracks have different trajectories through the detector, leading to the following classification, pictured in Fig. 3.1:

- **VELO tracks:** tracks with large polar angles or which go backwards with respect to the detector. These tracks will only traverse the VELO and are useful for the precise fitting of the primary vertex.
- **Upstream tracks:** low momentum tracks bent out of the LHCb geometrical acceptance before they can reach the tracking stations. These tracks only leave clusters in the VELO and TT and their momentum resolution is of the order 10-20%. These tracks are used for the RICH pattern recognition and for the reconstruction of slow pions from D^* meson decays.
- **Downstream tracks:** tracks originating outside of the VELO and traversing the TT and tracking stations. Such tracks can be decay products of long-lived particles (K^0_S, Λ). These tracks are important for the reconstruction of b-hadron decays, such as $B^0 \to J/\psi K^0_S$ and $B^0 \to \phi K^0_S$.
- **T tracks:** secondary interactions in the material of the upstream detectors lead to particles that only traverse the tracking stations. They are mainly used for the RICH2 pattern recognition and for the tracking station internal alignment.
- **Long tracks:** tracks traversing the entire spectrometer. They are the most useful tracks for the physics studies, which benefit from the precise determination of the track parameters.

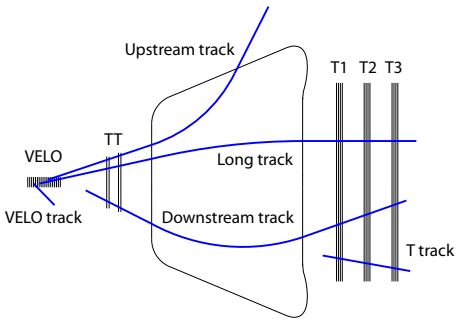


Fig. 3.1 Schematic view of the different track types defined in the LHCb. The view shows the $x - z$ plane. The scheme is not to scale and is taken from [70].

3.1.2 Pattern recognition

The track reconstruction starts with the pattern recognition, which consists of the building of track segments in the VELO and the tracking stations and then matching these, or using seeds in one detector to extend tracks to another. Several algorithms are run in a sequence to find the various track types defined in Sec. 3.1.1:

- **VELO seeds:** VELO clusters are collected along straight lines [71]. This follows from the low magnetic field in this region, meaning that no momentum information can be

assigned to the track segment at this stage. These VELO seeds are used as input to further tracking algorithms.

- **T seeds:** Segments of tracks are reconstructed in the tracking stations using both IT and OT clusters [72]. The algorithm starts by taking one x hit in each of the three tracking stations. A parabola is made out of them and other x hits are collected in a window around it. Next, compatible u/v hits are added to the track candidate (Sec. 2.3.1). Quality cuts are then used to clean the track sample.

- **Forward tracking:** Starting from the parameters of a VELO seed and an individual hit in the tracking stations, tracks are searched for by looking for hits in the other layers of the tracking stations in a window around the predicted track [73]. If the track candidate passes quality criteria, it becomes a long track. TT hits are then added to the track by collecting hits in windows around the predicted positions. About 90% of the long tracks are reconstructed using this algorithm.

- **Track matching:** This algorithm takes as an input both VELO and T seeds and tries to match them by extrapolating both segments to the bending plane in the magnet and comparing their positions, slopes and the number of compatible TT hits [74]. Once a track is built from two seeds, TT hits are added if they are close enough to the track. This algorithm reconstructs an additional 5% of long tracks.

- **Up/Downstream tracking:** Tracks are made of VELO/T seeds respectively, if at least three TT hits can be matched to the seeds.

- **VELO/T tracking:** Any VELO or T seed that has not been used in any of the previous algorithms is kept as a VELO or a T track respectively. The momentum resolution for these tracks is of dp/p ∼10-20%.

- **Clone killing:** Tracks may be reconstructed by more than one reconstruction algorithm. Clones are tracks that share a certain percentage of their hits. The tracks with the smaller number of hits are rejected by the clone-killing algorithm [75]. In case two or more tracks have the same number of hits, the highest quality one (based on its χ^2) is retained.

3.1.3 Track fitting

Once a track has been found by one of the tracking algorithms, it needs to be fitted in order to determine the most precise estimate of the track parameters. These parameters are then used to match the tracks with hits in the particle identification (PID) subsystems, locate the origin and decay vertices and calculate invariant masses in physics analyses.

A track is described by a set of track states, which each consists of a track vector $\vec{x}$ and a 5×5 covariance matrix C. Given the geometry of the detector, the track states are naturally

parametrized as a function of z. The state vectors are calculated, at a given z position, by the x and y positions, the track slopes and the particle momentum:

$$\vec{x} = \begin{pmatrix} x \\ y \\ dx/dz \\ dy/dz \\ q/p \end{pmatrix}, \tag{3.1}$$

where q is the signed charge and p is the particle momentum. This fifth parameter is obtained through the curvature of the track in the magnetic field. For practical reasons, the track states are chosen in the track fit at the measurement planes. These can be used to obtain the optimal track coordinates at other measurement planes either by extrapolation or interpolation.

The tracks are fitted using a Kalman filter, which is described in detail in Ref. [76]. The Kalman filter is a recursive least-squares method that gives optimal estimates of track parameters and allows multiple scattering and energy loss corrections due to ionisation to be properly taken into account. The principle of this technique is to add the measurements one after the other during the fitting procedure, updating the local track state at each step. This means that a state depends on all the previous states on the track. The method is based on a χ^2 minimization of the measurement on the tracks, thus making it equivalent to least-squares fit. For each track, a χ^2 is evaluated from track fit residuals as:

$$\chi^2 = [m - h(x)]^T [m - h(x)], \tag{3.2}$$

where m is the measurement vector, $h(x)$ is the measurement model vector, V is a covariance matrix and x is the vector of track parameters. $r(x) = m - h(x)$ is called the residual. All steps of Kalman filter track fit are discussed in App. B.1.

3.1.4 Standard LHCb Particle Selection

Once the tracks are reconstructed using the information from the tracking system and assigned an energy, momentum and PID information from the dedicated subdetectors, a first preselection is made in order to classify the tracks by type: muons, kaons, pions, electrons or protons. Three preselections are used in the performed analysis, the so-called "standard loose" electron, kaon and pion selections.

The *standard loose electrons* are selected amongst the long tracks, defined in Sec. 3.1.1. The selection is based on the number of hits on track that are found in the ECAL detector (Sec. 2.3.3), depending on the momentum of the particle. The requirement for a likelihood difference defined by the calorimeter is $\Delta\ln\mathcal{L}_{e\pi}^{\mathrm{CALO}} > -2$.

A requirement for the *standard loose kaons* selected amongst the long tracks is set on the log-likelihood difference $\Delta\ln\mathcal{L}_{K\pi}^{\mathrm{RICH}} > -5$ given by the RICH counters (Sec. 2.3.2).

Any long track can be selected as a *standard loose pions*. There are no specific requirements for these particles.

3.2 Alignment Strategy

To fulfil LHCb's physics goals relies on the good space and momentum resolution and precise particle identification. These performances are influenced by the overall alignment of the LHCb detector. Poor spatial alignment will lead to a degradation of the mass measurement or to systematic biases, which would degrade sensitive asymmetry measurements. For this reason, each subdetector must be aligned to a precision negligible with respect to its intrinsic resolution and their relative alignment must be precise enough in order to have no impact on physics parameters.

Since the VELO is the most precise device in the experiment, it will be used to define a global coordinate system for LHCb. However, the first VELO silicon-strip being so close to the *pp* collisions ($\sim$8 mm), the two VELO halves will need to be retracted by $\sim$3 cm during the phase where LHC will establish stable beams, i.e. at the beginning of each fill. Since the tracking stations will be aligned with respect to the VELO, the 10 μm precision of the motion controller is expected to be the largest uncertainty on the absolute position precision of the detector.

Misalignments can occur either due to wrong positioning of detector elements in the experimental area or due to inaccuracies during the construction phase. Although care was taken to mount each part precisely, shifts or rotations of up to a few millimetres or milliradians were observed with respect to the design position. A position survey was performed on each system to provide first corrections. However, the obtained precision is still outside the requirements coming from the physics measurements. For example, the IT position has been measured with a precision of 50 μm at the layer and ladder level, which is equivalent to the single hit resolution of the detector. At the box level, this precision is $\sim$1 mm. Another method is needed in order to ensure a better internal alignment and to provide with alignment corrections to the relative position of different subdetectors.

The fact that the polarity of the magnetic field will be inverted periodically during the running of the experiment will also influence the alignment of the spectrometer. Especially, the subsystems close to the magnet may shift slightly when the magnetic field is turned on or the polarity of the magnets is swapped. For these reasons, the following global LHCb alignment strategy has been defined [77]:

1. Internal VELO alignment;

2. Stand-alone IT-OT alignment;

3. Relative alignment of the tracking stations to the VELO;

4. Alignment of the TT using VELO-to-T tracks.

5. Alignment of the RICH, calorimeters and muons system with respect to the tracking stations.

These steps are briefly described in the following sections.

3.2.1 VELO Alignment

The VELO alignment is divided into three main parts. First, the modules inside each of the VELO-halves are aligned in their own boxes [78]. Next, the two VELO halves are aligned to each other [79]. Finally, the sensors inside each module are aligned [80].

The method used is an implementation of the Millepede algorithm [81]. The trajectories of particles are expressed as a linear combination of the track and the alignment parameters. After a χ^2 minimization (Sec. 3.1.3), which contains both the local (track) and global (alignment) parameters, a single matrix inversion leads to the alignment and track parameters. This matrix is usually very large, but can be reduced to a block diagonal form where one of the blocks has a dimension equal to the number of degrees of freedom that are taken into account in the alignment process. This number is 42 modules$\times$6 degrees of freedom = 252 for the alignment of all the modules inside the VELO halves. The alignment of these modules has been tested with MC simulated data. A precision of 1.3 μm is obtained for the translations along the x and y-axes. The accuracy on the rotation around the z-axis is 0.1 mrad [79].

The alignment of the two VELO halves with respect to one another is performed either by selecting tracks that go through the overlap region between the sensors on each side of the beam or by reconstructing the primary vertices. A precision of 12 μm along the two relevant translational degrees of freedom was obtained. The tilts around the x and y axes are known to 36 μrad, which is still within the system requirements [79]. The method using the primary vertices is less constraining than the overlap one. However, it is the only one usable when the VELO is open, which will happen at the beginning of each LHC fill. The precisions of 28 μm for the translations and 108 μrad for the rotations are obtained for the two VELO halves with respect to one another.

The alignment of the two boxes to fix a global LHCb coordinate system is performed by combining the two methods discussed above. The accuracies of 18 μm for the x and y translations and 103 μrad for x and y tilts are found for the two VELO halves with respect to the beam.

Finally, the relative sensor alignment has been tested to be precise to 1.3 μm for translations along the x- and y-axes [80]. Combining these results with the precision of the module alignment, it follows that the absolute position of one sensor inside either VELO half will be determined to better than 2 μm.

3.2.2 Tracking Station Alignment

The tracker consists of the Inner and Outer trackers. Since the tracking stations have a total of 12 active layers, a stand-alone alignment is possible. The IT and OT can be internally aligned separately using tracks going through either tracker only. Alternatively the two systems can be aligned together, using both tracks with larger angles with respect to the beam, and hence going through the two systems, and tracks going through the small overlap region between the silicon sensors and the straw tubes.

After installation of the detector in the experimental area, the position of the boxes was surveyed with a precision of 500 μm. This was carried out whilst the detector was in the open position and does not include possible systematic shifts during the closing. Taking these into account, an overall precision of 1-2 mm is quoted on the box positions [82].

In Ref. [83], a first alignment of the IT boxes and layers in the most sensitive coordinate (x) is discussed. Using a technique based on histograms of residuals, these studies establish the validity of the survey and provide first alignment constants.

Sec. 3.4 discusses the alignment of the tracking stations with a software method described in Sec. 3.3, using MC simulation and data recorded at LHCb during 2011 and 2012.

3.2.3 VELO to Tracking stations Alignment

Once the tracking stations are internally aligned, they are aligned to the VELO, which is the reference for the global coordinate system. This is done using the long tracks presented in Sec. 3.1.1. The results of this method are discussed in Sec. 3.4.

The Tracker Turicensis detector consisting of only four active layers, a stand-alone alignment is not possible in this case. The long or downstream tracks are used to align the TT with respect to both the VELO and the tracking stations. For this detector, upstream tracks reconstructed in the VELO and the TT can also be used for alignment purposes.

3.2.4 RICH, Calorimeters and Muon System Alignment

The three systems that perform the particle identification in the LHCb, namely the RICH counters, the calorimeters and the muon system, have less stringent alignment requirements (in terms of absolute precision) than the tracking system. However, misalignments are not harmless for these subdetectors. For example, misalignments of mirror segments in one of the RICH counters will lead to inaccurate reconstruction of the Cherenkov angle of the photons emitted by the particles that traverse the detector. In the calorimeters, misalignments in the ECAL will degrade the mass resolution of the reconstructed B meson decays involving prompt photons or π^0 by assigning an incorrect photon momentum. The measurement of the positions are also widely used in the process of electron identification. Finally the muon stations, although having a coarser granularity than the tracking system, is of such

importance in the L0 trigger (through muon identification and fast online p_{T} measurement), that misalignments must be taken into account.

These three subdetectors are aligned with respect to the fully aligned tracking system. For the RICHes, several components are necessary to be aligned using various methods. A stand-alone calibration system is used for the Hybrid Photon Detectors (HPDs) (Fig. 2.11). Then the mirrors are aligned by comparing the reconstructed photon positions with the expectation coming from the charged track parameters. The ECAL and the muon system are aligned using clean electron and muon samples to 0.5 and 1 mm, respectively. For the latter, the method used is the same as for the tracking stations, using high momentum muon tracks going through the OT and the muon stations.

3.3 Alignment within the LHCb Framework

The LHCb alignment framework has been developed inside the LHCb Gaudi software framework [58]. It provides a closed-form χ^2 minimization as in the Millepede algorithm [81]. In addition, it gives access to the detector elements and their conditions inside the LHCb framework and uses the same track model and track fit as in the standard reconstruction process. It also provides an easy-to-configure algorithm able to align for any combination of detector elements and for any degree of freedom (three translations and three rotations per element or group of elements). Since it uses the standard LHCb track model, it is able to properly take the multiple scattering, magnetic field and energy loss corrections into account. This is a gain with respect to the Millepede method, which uses a global track model. In addition, it will benefit from any improvement in the track model or track reconstruction software. Since it is able to access the detector elements inside the LHCb geometry databases, it can easily update the alignment constants, where these constants are consistent with the track model used in the reconstruction.

3.3.1 Alignment Parameters

In Sec. 3.1.3, the χ^2 minimization in the track fitting procedure was discussed. However, this computation only took the track parameters into account. The measurement model needs to be extended to include the alignment parameters α of the detector:

$$h(x) \rightarrow h(x, \alpha). \tag{3.3}$$

Contrary to the track parameters, the alignment parameters are common to all the tracks in the sample. They affect the detector elements themselves. Once the track parameters x have been determined by fitting the tracks using a first estimate of the alignment parameters, these can be obtained by minimizing a total χ^2 of a track sample with respect to both track

and alignment parameters:

$$\frac{d\chi^2}{d\alpha} \equiv 0. \tag{3.4}$$

To account for non-linear measurement model vector h, an iterative solution using the Newton-Raphson method is used (App. B.1). The changes in alignment parameters α from a set of initial alignment values α_0 can be expressed as:

$$\Delta\alpha \equiv \alpha - \alpha_0 = -\left(\frac{d^2\chi^2}{d\alpha^2}\right)^{-1}\Bigg|_{\alpha_0} \frac{d\chi^2}{d\alpha}\Bigg|_{\alpha_0}. \tag{3.5}$$

In case of non-linearities of the residuals as a function of the track and alignment parameters, the procedure needs several iterations until the solution for $\Delta\alpha$ converges, i.e. reaches a plateau. The formalism of the χ^2 minimization for alignment parameters is discussed in detail in App. B.2.

3.3.2 Tracking station Alignment Procedure

The framework in which the alignment procedure is run is centred on a single algorithm, which takes the tracks from the standard LHCb track fit as input. In parallel, it processes the detector elements specified by the user and assigns the hits on tracks to the corresponding elements. The correlations between the hits on tracks are computed and the global χ^2 is minimized in order to provide the alignment constants to be written in the condition database.

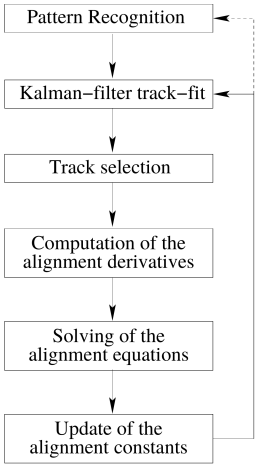


Fig. 3.2 Flow diagram of the software procedure used for the tracking station alignment. The pattern recognition can optionally be rerun at each iteration.

The general software sequence is shown in Fig. 3.2. The alignment algorithm uses a set of dedicated tools. The first one gives access to the detector elements in order to get the initial geometry and update it when the alignment procedure is finished. The track reconstruction is also accessible. Both the pattern recognition and the track fit sequences can be easily

adapted to each use-case. Once the tracks have been reconstructed, a second tool can be used to apply a track quality selection or select subsamples of tracks based on their momentum (in case the magnetic field is turned on), the number of hits in each subdetectors, the number of missed hits on the tracks, etc. The goal of this selection is to obtain a sample of clean tracks while rejecting most of the ghosts. Finally, the solver tools are used to find the solution of the system of M linear equations

$$Ax = b, \tag{3.6}$$

where $A = \frac{\partial^2 \chi^2}{\partial \alpha^2}$, $x = \Delta\alpha$ and $b = \frac{\partial \chi^2}{\partial \alpha}$. M is the number of alignment parameters.

The IT is composed of 3 stations×4 boxes×4 layers×7 = 336 silicon strip IT ladders, each having six degrees of freedom (three translations and three rotations). This adds up to a total of 2016 degrees of freedom (DoFs). On the other hand, the OT is made of 3 stations×4 layers×22 modules = 264 straw-tube OT modules that need to be aligned, adding an extra 1584 DoFs to the problem. The TT consists of 2 layout×15 modules+2 layout×17 modules = 64 silicon strip TT sensors adding 384 DoFs for the alignment procedure. The details of the tracking station design are discussed in Sec. 2.3.1.

Throughout the next section, the following convention is used: the degrees of freedom are called Tx, Ty and Tz (translation of the detector elements along the horizontal measurement direction, the vertical direction and the beam axis respectively) and Rx, Ry and Rz (rotations about the three axes).

3.4 Tracking stations Alignment Studies

The results of an investigation of the TT, IT and OT stations position stability with 2011 and 2012 data and MC samples are discussed in this section. First, general considerations on the event selection and alignment scenario are presented. The verification of the alignment procedure convergence is performed. In addition, the time dependence of alignment constants and dependence on survey constraints are studied. This section has been published as an internal LHCb note [84].

3.4.1 Alignment scenario

In the alignment procedure the detector modules are allowed to move independently. The degrees of freedom assigned to the tracking stations in the alignment study are listed in Table 3.1. The survey positions of the detectors are used as a constraint in the minimization with an uncertainty of the measurement. Several detector elements are taken as a pivot points to inhibit global translations of the spectrometer. Their positions are fixed by artificially increasing survey precision in the constraint. The uncertainties of the survey constraints are

used to fix or allow to move selected elements of the detectors in translations or rotations with respect to their survey positions.

Table 3.1 Degrees of freedom of the tracker stations alignment.

Detector	Alignable	DoFs
TT	Layers	Tz
	Modules	$TxRz$
IT	Boxes	$TxTzRz$
	Layers	None
OT	CFrames	$TxRz$
	CFrameLayers	Tz
	Modules	$TxRz$

The study is performed using a sample of the $D^0 \to K^-\pi^+$ and $J/\psi \to \mu^+\mu^-$ decays. The selection criteria applied to the sample are discussed in App. B.3.1. The number of analysed events is $2{\cdot}10^6$ for 2011 and 2012 data sample of the D^0 meson decay and 10^4 for 2011 data sample of the J/ψ meson decay. In case of the simulation, $1.5{\cdot}10^6$ events are used in the alignment studies. The samples contain an equal amount of events for both magnet polarities except for J/ψ data sample.

3.4.2 Alignment convergence study

As a first study, a verification of the alignment procedure convergence is performed meaning that the procedure assumes that the detector is misaligned and tries to move it back to its ideal position. Each box of the first IT station (IT1) is shifted from official database (DB) position. The x position of every box is misaligned by 100 or 200 μm in different directions as listed in Table 3.2. A different size for the box shift is applied to check how the convergence depends on the initial misalignment. Three iterations of the alignment procedure are applied to minimize the difference between artificial misalignment and official DB values. The 2011 and 2012 data and 2011 simulated samples are used to perform this alignment study.

Table 3.2 Artificial misalignment in IT1Boxes.

IT/Station1/Box	x shift [μm]
Top	-100
Bottom	+100
ASide	-200
CSide	+200

The misalignment in the x position is observed on the residual distributions of IT1Boxes for both polarization. The shift is shown in Fig. 3.3 where black triangles correspond to official values taken from DB, blue crosses are values with artificial misalignment, red circles and green boxes are values after first and second iteration, respectively. After three iterations (green boxes), the distribution is indistinguishable with the one of an official alignment.

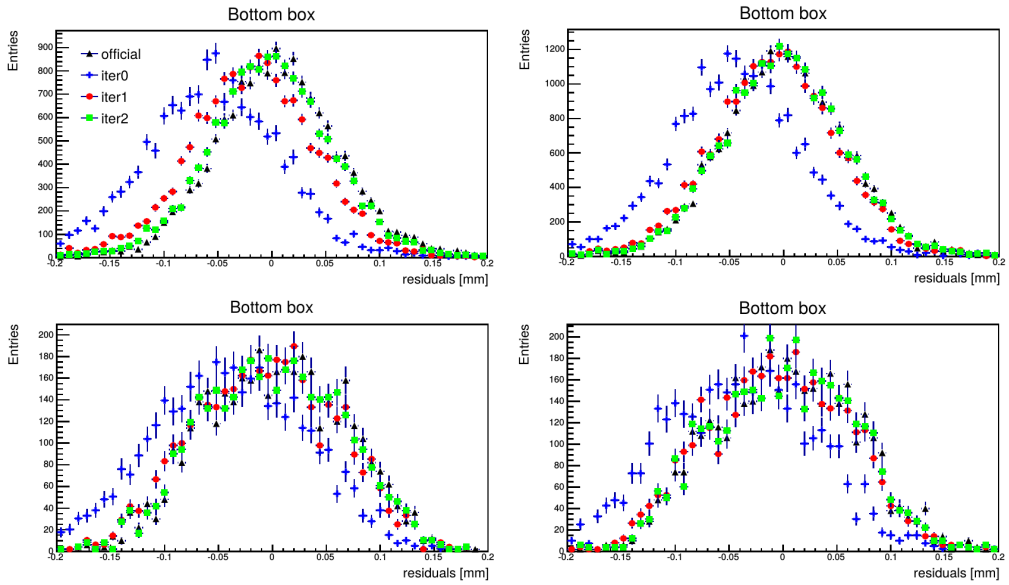


Fig. 3.3 Residual distributions with 2011 (top) data sample and (bottom) simulation for (left) Down and (right) Up polarizations.

The result of the alignment convergence study is as follows. The convergence of the alignment procedure confirmed that for small misalignments performing 3 iterations is sufficient. The observed difference with official values is below 60 μm and comparable with IT station resolution (σ_{det} $\sim$50μm). The discrepancy between opposite magnet polarizations is up to 25 μm. Therefore, there is no strong evidence for dependence of detector positions on magnetic field polarization. The investigation results are comparable between data samples and simulation. The detailed description of the alignment convergence study is summarized in App. B.3.2.

3.4.3 Time dependence of the alignment

The variation of the alignment parameters within one full time period of data taking has been investigated using a test with finer segmentation. At least one data taking period of each polarization for 2011 and 2012 is considered. Each time period is divided into several parts of two-four days. The alignment procedure has been probed using five different configurations of the survey constraints. These constraints have different strength on the position of considered detector stations. In addition, the time dependence of the alignment procedure has been studied using $D^0 \to K^-\pi^+$ and $J/\psi \to \mu^+\mu^-$ data samples. All results presented in this section are obtained with three alignment iterations[1].

[1] All figures in the section are in global frame.

Survey constraints study

The alignment constants variation in a time period is determined using different survey constraints on the elements of the OT station. The configuration descriptions and conditions of the OT survey constraints are defined below and are listed in Table 3.3:

- *Set1*: alignment performed per OT module. Set tight constraint to survey position for OT3 (CFrameLayers) as a pivot point;
- *Set2*: alignment performed for OTCFrameLayers. Set tight constraint to survey position for OT1 and OT3 (CFrameLayers);
- *Set3*: alignment performed per OT module. Set tight constraint to survey position for OT1 and OT3 (CFrames and CFrameLayers). These options are used in the alignment convergence study (Sec. 3.4.2);
- *Set4*: alignment performed for OTCFrameLayers. Set tight constraint to survey position for OT1 and OT3 (CFrames and CFrameLayers);
- *Set5*: alignment performed for OTCFrameLayers with OTModules allowed to rotate around x-axis. Set loose constraint to survey position for OT1 and OT3 (CFrames and CFrameLayers).

Table 3.3 Study conditions with different survey constraints of the OT station.

Alignable	Element	DoFs	Uncertainties, [mm]				
			set1	set2	set3	set4	set5
OTCFrLayers	T1/X1U	Tx	None	10^{-5}	10^{-5}	10^{-5}	0.5
		Tz	None	10^{-5}	10^{-5}	10^{-5}	0.5
	T3/X1U	Tx	0.5	10^{-5}	10^{-5}	10^{-5}	0.5
		Tz	10^{-5}	10^{-5}	10^{-5}	10^{-5}	0.5
OTCFrame	T3X1U(A|C)Side	Tx	None	0.5	10^{-5}	10^{-5}	0.5
OTModules			$TxRz$	None	$TxRz$	None	Rx

The study is performed using the $D^0 \to K^-\pi^+$ decay in 2012 data sample with Up polarization[2]. The fluctuation in Tx for OT1CFrame and IT1Boxes and in Tz for TTLayers and IT1Boxes is shown in Fig. 3.4 where the first point is the official alignment constant, other points correspond five sets of the OT survey constraint configurations, respectively. The error bars represent the parameter uncertainty returned by the alignment procedure. The alignment parameter fluctuations for all investigated detector stations is reported in App. B.3.3.

The difference between configuration sets for the aligned elements of IT and OT station are reported in Table 3.4 where the resolution of each detector is called σ_{det}. The time dependence result for all elements of tracker detectors are discussed below:

[2]The data sample corresponds to 3114 fill.

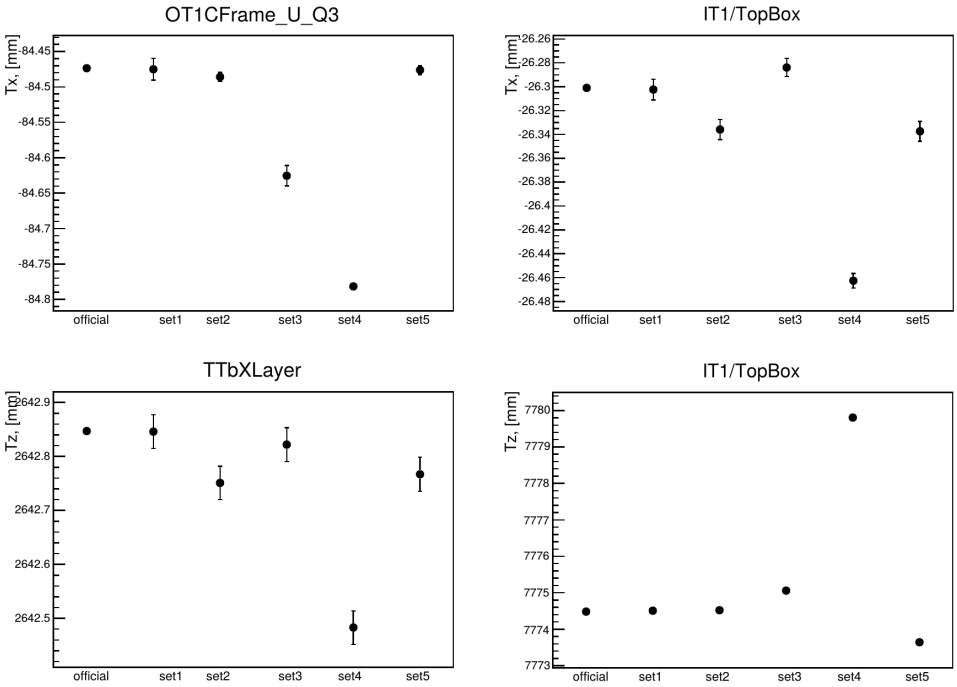


Fig. 3.4 (top) Tx and (bottom) Tz alignment parameter fluctuation of OT1CFrame, IT1Boxes and TTLayers.

Table 3.4 Time dependence result of the alignment using 2012 data sample.

Alignable	Element	DoFs	Difference [μm]			
			set2	set3	set4	set5
$\sigma_{\text{det}} \sim 50$ μm						
ITBoxes		Tx	5-40	5-40	80-200	5-40
		Tz	50-400	300-1200	$\geq$2500	50-1300
$\sigma_{\text{det}} \sim 200$ μm						
OTCFrs	T(1|2)X1U(A|C)Side T3X1U(A|C)Side	Tx	5-40	100-180 10^3	360-500	5-40
OTCFrLrs	T1X1U T2X1U T3X1U	Tz	200 150 $5{\cdot}10^{-6}$	200 10-80 $5{\cdot}10^{-6}$	200 230 10^{-5}	$\geq 10^3$

- *Set1-2.* The set1 and set2 configurations give a similar result. However there are important differences between constraints used for both cases. In set2 the OTModules are fixed with respect to the CFrames while for set1 they are aligned for two degrees of freedom, Tx and Rz. Moreover, only one OT station is fixed with survey constraint in set1. No significant difference between those two sets in IT and OT stations is observed. The shift in Tz DoF of TT station (Fig. 3.4) in set2 compared to the official values most likely originates from the fact that both OT1 and OT3 CFrames are fixed in z position. It will affect the angles of the tracks and could result in scaling effect for the whole spectrometer.

- *Set4.* The OTCFrames and OTCFrameLayers are tightly fixed in the set4 configuration. Since the OT stations are displaced further from the beam line in x direction, it has an impact on track angles and results in a shift of TT stations in Tz as shown in Fig. 3.4. In addition, the offset of IT in Tx and Tz DoFs is observed most likely as an effect of TT movement and mass constraint used in the alignment procedure.

- *Set3.* The difference between set3 and set4 configurations is that the OTModules are allowed to move independently of the OTCFrames. The effects described for set4 are still visible for set3 configuration but are largely diminished by movement of the OTModules in Tx DoF (Fig. 3.4).

- *Set5.* In this configuration the importance of fixed pivot points in the alignment procedure is evaluated. The result shows that the alignment procedure without a fixed reference point leads to global translations. All TT, IT and OT stations are displaced in Tz in one direction with comparable size of the offset in OT and IT detectors.

Both tight constraints and very loose ones lead to significant deviation from the official results. It is interesting to verify which configuration set gives the best description of physical observables. The D^0 invariant mass and its resolution distribution is reconstructed to investigate it. Fig. 3.5 presents the D^0 meson mass and resolution value for different configurations where the official value is a black square while considered configurations from set1 to set5 are indicated by a red circle, blue and green triangle, purple cross and blue star, respectively. The vertical error bars correspond to the uncertainty of the fit to the mass spectrum. The alignment procedure for all configurations gives a consistent D^0 mass within observed resolution, $m(D^0)$=1864.84±0.17 MeV/c^2 [3]. The alignment procedure for the set1 configuration gives the best mass value with respect to nominal D^0 mass [3] used in the official alignment. The largest discrepancy of the mass and resolution distribution to the official value is observed for the set4 configuration where the elements of OT station are tightly fixed.

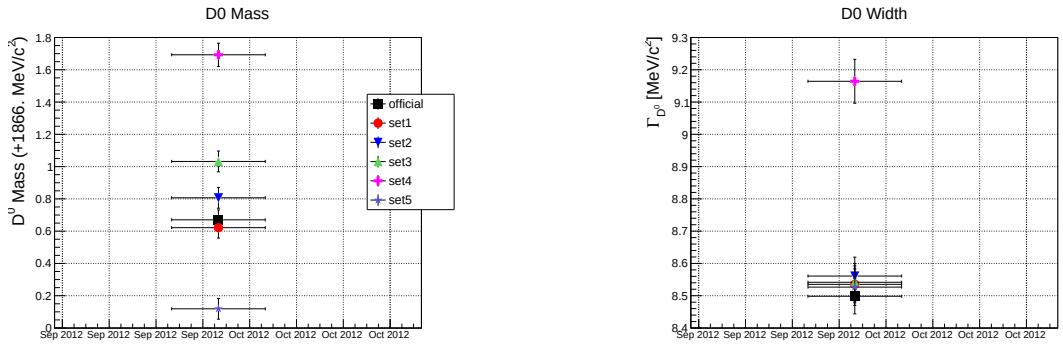


Fig. 3.5 Distribution of the D^0 meson (left) mass and (right) resolution used data range.

Time dependence study

The time dependence of the alignment procedure is verified using 2011 and 2012 data sample of the $D^0 \to K^-\pi^+$ and $J/\psi \to \mu^+\mu^-$ decays. Primary focus is given to effects of polarity change for the LHCb dipole magnet. The behaviour of the alignment parameters in IT and OT stations is investigated.

2011 $D^0 \to K^-\pi^+$ data

Two time periods of 2011 data are selected to perform a time dependence study: 26.04. - 10.06. and 11.06. - 29.06. with Down and Up magnet polarities, respectively (Table 3.5).

Table 3.5 Time periods of 2011 data sample.

Part	Day	Fill number	Part	Day	Fill number
Down polarity			11	07.06	1856
1	26.04	1739	12	09.06	1862
2	28.04	1745	Up polarity		
3	30.04	1749	1	11.06	1863
4	02.05	1755	2	13.06	1867
03.05-15.05 -Technical Stop			3	16.06	1871
5	19.05	1795	4	18.06	1880
6	24.05	1804	5	21.06	1883
7	27.05	1809	6	24.06	1889
8	30.05	1816	7	26.06	1894
9	02.06	1836	8	28.06	1901
10	04.06	1845			

During 03.05.2011 - 15.05.2011 time period, there was a technical stop, hence no data is available. The fluctuation of the alignment parameters are evaluated for TT, IT and OT stations with the set2 configuration that gave the smallest difference to the official parameters. The results of the alignment procedure are discussed for each detector separately below and presented in Figs. 3.6-3.7 where the official parameters are marked by a blue and red lines, the set2 configuration values are indicated by a green and brown lines for Down and Up magnet polarities, respectively. The observed differences with official values for short time period before the technical stop originate from an error in the condition DB - the validity of the alignment parameters value after the technical stop was extended to just before its start. The alignment parameter fluctuations for all investigated detector stations is presented in App. B.3.3.

- *Tracker Turicensis.* The Tx alignment parameter fluctuations within one period of data taking are within the spatial resolution of the TT detector ($\sim$50 μm) (Fig. 3.6). Although the OTCFrameLayers are tightly constrained to survey position in x direction, it gives a very small effect on the TT detector. The observed difference in Tz alignment parameter is up to 150 μm as shown in Fig. 3.6.

- *Inner Tracker.* Observed fluctuation in Tx alignment parameter is of the order 50 μm (Fig. 3.7). The observed difference between both magnet polarisations are consistent

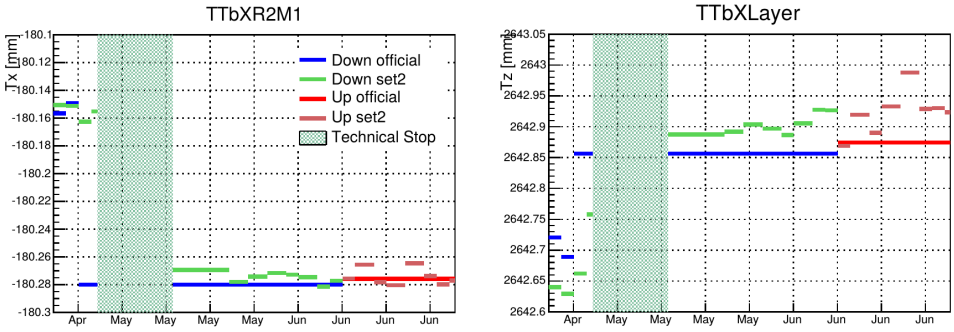


Fig. 3.6 Alignment parameter fluctuation of (left) TTModules in Tx and (right) TTLayers in Tz with 2011 $D^0 \to K^-\pi^+$ data sample.

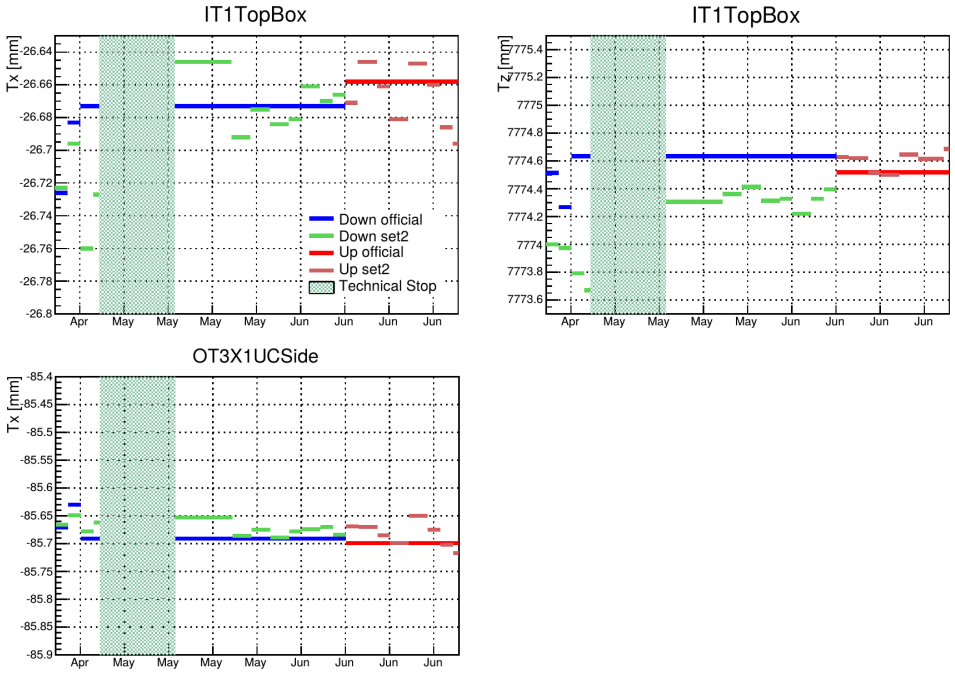


Fig. 3.7 Alignment parameter fluctuation of (top) IT1Boxes in Tx and Tz and (bottom) OTCFrames in Tx with 2011 $D^0 \to K^-\pi^+$ data sample.

with the fluctuations within period. However, the variation in Tz alignment parameter is up to 400 μm, this remains within the fluctuation size observed within analysed period.

- *Outer Tracker*. The variation between both configurations within considered period are consistent to each other which is in agreement with the IT detector results (Fig. 3.7).

The movement in Rz DoF for all detectors is negligible. No discrepancy in x and z translations is observed between Up and Down magnet polarizations for analysed period. The results of the time dependence study are summarized in Table 3.6.

Table 3.6 Results of the time dependence study with 2011 $D^0 \to K^-\pi^+$ data sample.

Detector	Alignable	DoFs	Difference [μm]
			set2
$\sigma_{\text{det}} \sim 50$ μm			
TT	Layers	Tz	5 - 80
	Modules	Tx	5 - 20
$\sigma_{\text{det}} \sim 50$ μm			
IT	Boxes	Tz	10 - 500
		Tx	5 - 50
$\sigma_{\text{det}} \sim 200$ μm			
OT	CFrames	Tx	10 - 50
	CFrLayers	Tz	10 - 100

The distribution of the D^0 meson mass and resolution in 2011 time period is shown in Fig. 3.8 where the official values are indicated by the blue and red squares, the set2 configuration points are marked by the green and brown triangles for Down and Up magnet polarizations, respectively. The vertical error bars correspond to the uncertainty of the fit to the mass spectrum.

The observed fluctuations within time period of 2011 data taking are consistent with results obtained with 2012 data and discussed in detail in App. B.3.3.

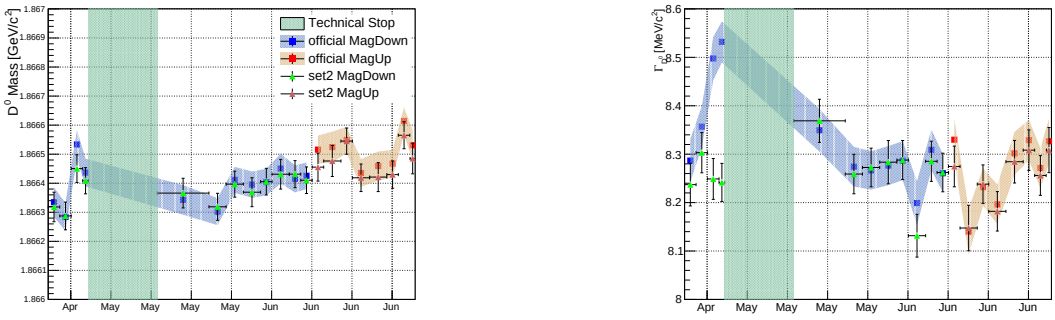


Fig. 3.8 Distribution of the D^0 meson (left) mass and (right) resolution in 2011 data sample.

2011 $J/\psi \to \mu^+\mu^-$ data

The time dependence of the alignment procedure is also verified using the $J/\psi \to \mu^+\mu^-$ decay following the same procedure as for the $D^0 \to K^-\pi^+$ decay. The 2011 data sample consisting of one time period with Down magnet polarization (parts 5 to 9 from Table 3.5) is used. The alignment procedure results for the J/ψ meson decay are presented in Fig. 3.9 (App. B.3.3) where the official parameters are marked by a blue line, the set2 configuration values are shown by a green line and the "set2, no Tz" values are indicated by a pink line. Since a large discrepancy with set2 configuration is observed in Tz DoF of the IT station (Table 3.6), the "set2, no Tz" configuration is introduced in the study with J/ψ meson decay. The "set2, no Tz" is the same as the set2 configuration except for the alignment in z direction.

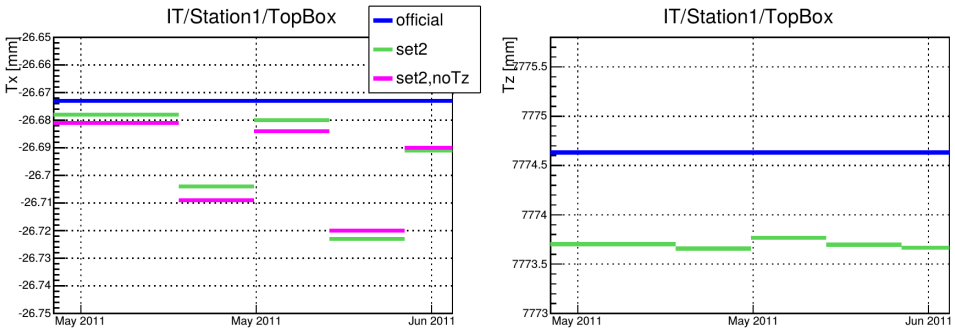


Fig. 3.9 (left) Tx and (right) Tz alignment parameter fluctuation of IT1Boxes with 2011 $J/\psi \to \mu^+\mu^-$ data sample.

Observed fluctuations in TT and OT stations within one period are consistent with alignment results obtained using the D^0 data sample. The difference between three configurations is in agreement with the tracking detector resolution. The best agreement with official parameters is obtained with the "set2, no Tz" configuration. On the other hand, a visible discrepancy of both set2 configurations remains for Tx DoF in Bottom and Top boxes of IT station as shown in Fig. 3.9. The alignment procedure in Tz DoF using set2 configurations leads to a discrepancy with official values of about 2 mm. The conclusion is that the J/ψ data sample is not suitable for the Tz alignment in IT detector. That effect is discussed in detail in App. B.3.3.

As in the case of the D^0 data sample, no significant dependence of the physical observables with different alignment configurations is observed as shown in Fig. 3.10 where the official values of the J/ψ meson mass and resolution distribution are shown by blue squares, the set2 and "set2, no Tz" configuration points are indicated by green and purple triangles, respectively.

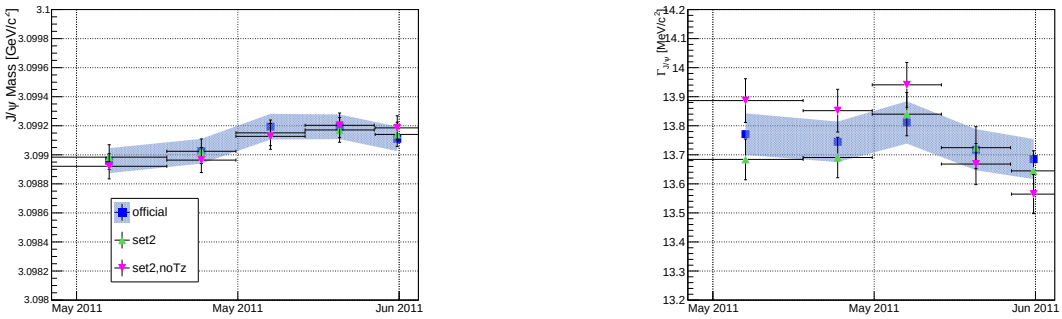


Fig. 3.10 Distribution of the J/ψ meson (left) mass and (right) resolution in 2011 data sample.

Reconstruction with independent data sample

To verify existence of a possible bias on the D^0 (J/ψ) mass value due to its usage as a

constraint in the alignment procedure, a sample of the D^0 (J/ψ) meson decay is reconstructed. The alignment DB obtained after alignment procedure of the J/ψ (D^0) data sample is used. The observed mass and resolution values are compared using two scenarios:

- same data sample is used for the alignment and reconstruction;
- independent data sample is used for the reconstruction: the D^0 (J/ψ) mass distribution is determined using a sample aligned with a J/ψ (D^0) mass constraint.

The 3114 fill of 2012 data taking is divided into consecutive run numbers (Table 3.7) to investigate the variation of the tracker station alignment parameters during very short time, ∼3 hours.

Table 3.7 Run numbers in 3114 fill of 2012 data sample.

Data set	Runs	Length of runs	Averaged Lumi [pb^{-1}]
file1	129616 - 129619	2h 17m	1.58
file2	129620 - 129630	1h 16m	2.78

The results of the study are presented as the D^0 and J/ψ meson mass and resolution values in Figs. 3.11-3.12 where the black squares and blue triangles are file1 and file2 with the same sample, red circles and green triangles are file1 and file2 with the independent sample. The vertical error bars correspond to the uncertainty of the fit to the mass spectrum.

The mass difference between two scenarios of the D^0 and J/ψ data samples are observed to be in range of 0.15-0.5 MeV/c^2. The maximal mass resolution difference is ∼0.2 MeV/c^2 for both samples. Therefore, observed mass and resolution values for two data samples are consistent with the statistical uncertainty.

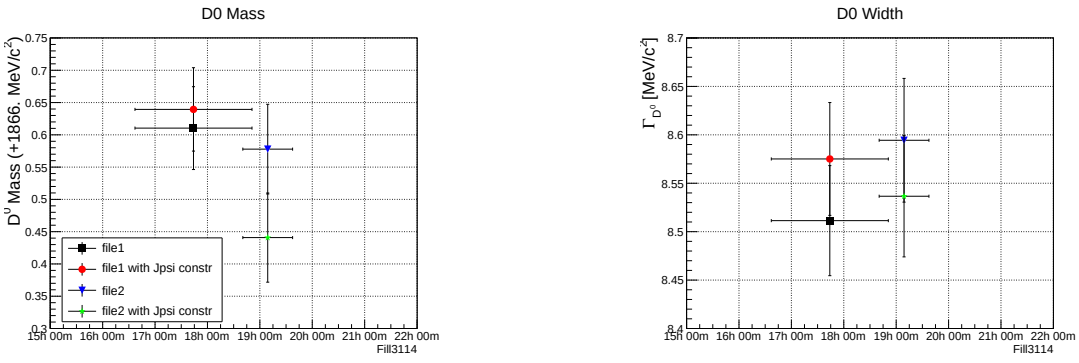


Fig. 3.11 Value of the D^0 meson (left) mass and (right) resolution for different runs in 3114 fill and two alignment scenarios.

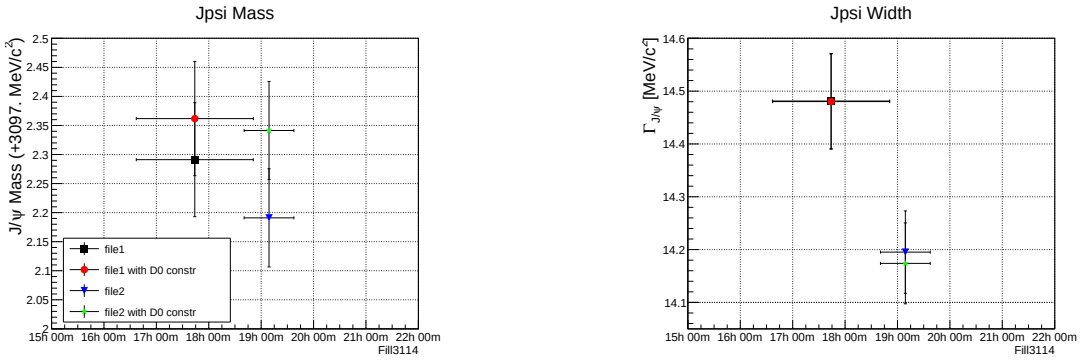


Fig. 3.12 Value of the J/ψ meson (left) mass and (right) resolution for different runs in 3114 fill and two alignment scenarios.

3.4.4 Results of the Alignment Studies

The following conclusions can be drawn from the tracking station alignment studies:

Alignment convergence study. The discrepancy in Tx DoF between realigned and official parameters is up to 60 μm with 2011 and 2012 data samples and is below 20 μm with 2011 simulation. However, these results are within the IT spatial resolution. The results of MC sample are in agreement with the results of data samples. The observed difference in Tx between two magnetic field polarizations is about 25 μm. Therefore, there is no strong evidence of the detector positions dependence on magnetic field polarization.

Time dependence study. A good agreement between different configurations with loose Tx constraints in OTCFrames and fixed OTModules is seen for both magnet polarizations. The observed fluctuations within one period are up to 60 μm and are in agreement with tracker detector resolutions. No discrepancy is observed between Up and Down magnet polarizations for analysed data within uncertainties. The variation of the alignment parameters does not have a significant impact on physical observables like resonance masses or their resolutions.

Several survey constraint configurations of the OT station are investigated. Significant deviations in tracker detector positions are observed for tight or loose constraints. Nevertheless, the resulting D^0 mass values are consistent with nominal mass value within 3σ uncertainty.

The difference between official and realigned IT station parameters for the J/ψ data sample proved to be 2-3 times worse than for the D^0 data sample. This is not the case for the OT and TT stations. In the IT detector, there are far less tracks from the J/ψ meson decays compared to the D^0 meson decays for the same number of the decay vertices. This leads to large fluctuation of IT positions with the J/ψ data sample. For both J/ψ and D^0 data samples, the mass and resolution value variation between alignment procedure using the resonance mass as a constraint to independent one is within 0.15-0.5 MeV/c^2. Hence, no significant bias on physical observables is introduced.

Chapter 4

Analysis of the $B_s^0 \to J/\psi\phi$ decay

The measurement of the time dependent $\mathcal{CP}$ asymmetry for the decay $B_s^0 \to J/\psi\phi$ allows for the determination of the $\mathcal{CP}$-violating phase ϕ_s. This thesis presents a measurement of the ϕ_s by fitting the theoretical $B_s^0 \to J/\psi\phi$ decay rates, introduced in Sec. 1.3, to the measured decay time and angular distributions of the reconstructed $B_s^0 \to J/\psi\phi$ decays with following subsequent decays[1]: $J/\psi \to e^+e^-$ and $\phi \to K^+K^-$. The fit relies on the maximum likelihood method that can be used to estimate physical parameters from a given data sample by maximizing a likelihood function. The principle idea and implementation of the maximum likelihood fit will be discussed in Sec. 4.6.

Several topics have to be addressed before the $\mathcal{CP}$-violating phase ϕ_s can be determined from the fit:

- The data sample of selected $B_s^0 \to J/\psi\phi$ candidates is polluted with physical and combinatorial background with different decay time and angular distributions than the signal. In order to get a correct measurement of the phase ϕ_s, the background components have to be identified and modelled in the fit.

- The measured decay time and angular distributions are distorted by acceptance effects introduced by the geometrical coverage of the detector and by selection and trigger requirements. The fit has to include corrections for these acceptance effects. The corrections of the decay time distribution are determined with help of a control $B^0 \to J/\psi K^*(892)^0$ decay. The corrections to the angular distribution are defined using a signal simulated sample.

- The decay time resolution of the detector influences the sensitivity on ϕ_s and has to be modelled in the maximum likelihood fit performed on the data sample of selected $B_s^0 \to J/\psi\phi$ candidates. The parameters of the time resolution model are extracted simultanously from the fit. An external input from $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ channel is used as a constraint on the parameter values.

[1]Throughout the thesis a $B_s^0 \to J/\psi\phi$ notation indicates subsequent decays of: $J/\psi \to e^+e^-$ and $\phi \to K^+K^-$, otherwise they are specified explicitely.

- The production flavour of the B_s^0 mesons must be determined to distinguish between the decay rates of B_s^0 and $\bar{B}_s^0$ in the fit. This is done with flavour tagging algorithms that exploit either the hadronisation properties of the signal B_s^0 meson or the properties of a second b-hadron that might be produced from the $b\bar{b}$ quark pair.

Each topic is discussed in more detail below.

4.1 Data samples

The events for the analysis are selected from data collected during 2011 at $\sqrt{s} = 7$ TeV and 2012 at $\sqrt{s} = 8$ TeV. They are reconstructed with "version 14" of the LHCb reconstruction software (`Reco14`). The analisys is performed on preselected samples with loose selection criteria aimed at fast reduction of data without signal candidates - so called "stripping"[2]. In both cases, the `StrippingBs2JpsieePhiDetachedLine` selection is applied using `DaVinci v37r2p4` package [57], the pre-selection details are summarized in Table 4.3. Further processing was performed offline in order to apply the selection criteria, single candidate choice for events with more than one, and obtain sWeighted particle mass distribution, the procedure for which is described in Sec. 4.2.

The Monte Carlo simulated signal events used for this thesis contain one $B_s^0 \to J/\psi(e^+e^-)$ $\phi(K^+K^-)$ signal decay in every event. To achieve this, repeated hadronization of the generated $b\bar{b}$ pair is performed until a B_s^0 is found. The B_s^0 meson is then forced to decay into the signal channel. For the generation of the signal decay, `EvtGen` [65] and `Photos++` [66] programs use the time dependent decay amplitudes given in Sec. 1.3. To save on computational costs, the event is only kept if the final state particles of the signal decay lie in the acceptance of the LHCb detector (10 mrad $< \theta <$ 400 mrad). The parameters used in the MC generation are summarized in Table 4.1.

Table 4.1 Parameters used in the generation of the MC simulated signal events. The parameter values are based on those reported in Ref. [28].

Physics parameter	Value
ϕ_s	0.07 rad
Γ_s	0.6653 ps^{-1}
$\Delta\Gamma_s$	0.0917 ps^{-1}
$\lvert A_0(0)\rvert^2$	0.722
$\lvert A_\parallel(0)\rvert^2$	0.480
$\lvert A_\perp(0)\rvert^2$	0.499
$\delta_\parallel - \delta_0$	3.30 rad
$\delta_\perp - \delta_0$	3.07 rad

Besides the obvious use of the simulated signal events to test the extraction of the B_s^0 meson mixing phase and to optimize the selection criteria, the simulated events are also used

[2]The `21r1` (`Stripping21r1-Merging-DV-v36r1`) and `21` (`Stripping21-Merging-DV-v36r1`) stripping versions are used for 2011 and 2012 data, respectively.

Table 4.2 Number of MC events, quoted is a sum of `Pythia6` (MagUp + MagDown) + `Pythia8` (MagUp + MagDown) for 2011 and 2012.

Decay Channel	Number of events
$B_s^0 \to J/\psi(ee)\phi$	20M
$B^0 \to J/\psi(ee)K^*(892)^0$	12M
$J/\psi \to ee$ (incl)	20M
$B_s^0 \to J/\psi(ee)$X	11M
$B^0 \to J/\psi(ee)$X	11.6M
$\Lambda_b \to J/\psi(ee)pK$	10M
$\Lambda_b \to J/\psi(ee)p\pi$	10M

for the determination of possible acceptance effects due to detector geometry, reconstruction and selection.

The MC samples used in the analysis are produced with `Sim08` configuration [85] and summarized in Table 4.2. For each decay channel, half of the sample is produced with `Pythia6` [63], half with `Pythia8` [64], each of them being split equally into Up and Down magnet polarizations[3]. A track level smearing has been applied to the simulated events in order to match momentum resolution to the one observed in data.

4.1.1 Trigger selection

In order to select the signal candidates while rejecting background events, the trigger (Sec. 2.3.5) exploits quantities characteristic for the signal decay. The trigger lines that are suitable to select the $B_s^0 \to J/\psi(e^+e^-)\phi(K^+K^-)$ decays are based on the reconstruction of the electrons and hadrons which might form a J/ψ and ϕ mesons.

There are two L0 lines of interest in this analysis, the electron and hadron lines. The L0 electron line, `L0Electron`, searches for a single electron track with moderately high transverse energy $E_{\rm T}$>2.5 GeV whereas the hadron line, `L0Hadron`, searches for hadron track with a $E_{\rm T}$>3.5 GeV. Only candidates where the electrons from the reconstructed signal B_s^0 meson decay triggered the event are used. These events are called Triggered On Signal events (`TOS`). There are no dedicated HLT (Sec. 2.3.5) trigger lines for electrons in 2011 and 2012 data. Thus, few HLT1 and HLT2 topological lines are used to reduce the background events. The single track line, `Hlt1TrackAllL0`, requires for presence of a single detached high momentum and $p_{\rm T}$ track to identify decays coming from the B_s^0 meson decay. Two types of multi track lines, `Hlt2Topo(2,3,4)BodyBBDT` and `Hlt2TopoE(2,3,4)BodyBBDT` that require for existance of at least 2, 3 or 4 tracks, are used. In the first case, the filtered tracks are taken as an input. In the second case, the filtered tracks with an extra PIDe requirement (Sec. 2.3.3) for at least one of the tracks are taken into account. Finally for the inclusive ϕ

[3]For 2011 MC sample the trigger is simulated with `Moore v12r8g3` package and `TCK 0x40760037` trigger configuration [69] while `Moore v14r8p1` and `TCK 0x409f0045` is used to simulate the trigger of 2012 MC sample (Sec. 2.3.7).

line, `Hlt2IncPhi`, the $\phi \to K^+K^-$ candidates are built in two steps: first, geometric cuts on the track are applied, and afterwards RICH PID (Sec. 2.3.2) is calculated to provide further filtering.

A full list of the criteria used in the L0, HLT1 and HLT2 trigger lines is given in App. C.1. In total the L0 lines contain ∼72% of the $B_s^0 \to J/\psi\phi$ candidates used in this analysis, the Hlt1 and Hlt2 lines include ∼28%.

4.1.2 Offline selection

After the trigger selection, an additional offline selection is performed in order to further reduce the background contribution. The specific selection criteria are optimized to maximize the sensitivity on the phase ϕ_s [86]. First, a stripping selection is applied and then a tighter final selection is used. The detailed selection requirements are listed in Table 4.3 and documented in the following.

Table 4.3 Stripping and offline selection criteria of the $B_s^0 \to J/\psi\phi$ candidates.

Decay mode	Cut parameter	Stripping	Offline selection
$J/\psi \to e^+e^-$	$\Delta\ln\mathcal{L}_{e\pi}$	>0	-
	$\chi^2_{\text{track}}/\text{ndf}(e)$	<5	<3
	$\chi^2_{\text{IP}}(e)$	-	>0
	$p_{\text{T}}(e)$	>500 MeV/c	-
	$\chi^2_{\text{vtx}}/\text{ndf}(J/\psi)$	<15	-
	$p_{\text{T}}(J/\psi)$	-	>400 MeV/c
	$m(J/\psi)$	∈[2500, 3300] MeV/c^2	-
$\phi \to K^+K^-$	$\Delta\ln\mathcal{L}_{K\pi}$	-	>0
	$\chi^2_{\text{track}}/\text{ndf}(K)$	-	<3
	$p_{\text{T}}(K)$	-	>200 MeV/c
	$p(K)$	-	>2000 MeV/c
	$\text{GhostProb}_{\text{track}}(K)$	-	<0.5
	$p_{\text{T}}(\phi)$	>1000 MeV/c	-
	$\chi^2_{\text{vtx}}/\text{ndf}(\phi)$	<15	<9
	$m(K^+K^-)$	∈[990, 1050] MeV/c^2	-
$B_s^0 \to J/\psi\phi$	$m(B_s^0)$	∈[4500, 6000] MeV/c^2	∈[4600, 6000] MeV/c^2
	$\chi^2_{\text{vtx}}/\text{ndf}(B_s^0)$	<10	-
	$t_{B_s^0}$	>0.3 ps	-

Track selection

The tracks that are used in the reconstruction of the $B_s^0 \to J/\psi\phi$ decays have to fulfill basic selection criteria. In order to reject bad quality tracks, a requirement on the χ^2 of the track fit divided by the number of degrees of freedom, $\chi^2_{\text{track}}/\text{ndf}$, is made.

J/ψ meson reconstruction

A J/ψ meson candidate is reconstructed using two electron candidates made of tracks with a positive electron probability $\Delta\ln\mathcal{L}_{e\pi}$, the difference between the logarithm of the electron and pion likelihoods. In addition, a minimum electron transverse momentum p_{T} is required to suppress electrons from the pp primary interaction vertex (PV). The χ^2_{IP} of the electron defines the difference between the χ^2 of the PV reconstructed with and without electron track. The particles produced in the pp interactions have, in general, a smaller p_{T} than daughters from the B meson decays due to the large boost. The quality of the J/ψ vertex combination is ensured by cutting on the $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$ of the J/ψ vertex fit. The invariant mass of the dielectron system $m(e^+e^-)$ should be within a certain range of the true J/ψ mass, world average $m(J/\psi) = 3096.9\pm0.006$ MeV/c^2 is used [3].

ϕ meson reconstruction

A ϕ meson candidate is reconstructed by combining two kaon candidates with a positive kaon probability $\Delta\ln\mathcal{L}_{K\pi}$ (Sec. 2.3.2). The ghost probability for kaons refers to the probability that the reconstructed track is made up from false hits in the detector and is not left by real particles traversing the tracking detector. The fitted ϕ vertex must have a good $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}(\phi)$ and the invariant mass of the combined kaons $m(K^+K^-)$ has to be in an interval around the true ϕ mass, world average $m(\phi) = 1019.461\pm0.019$ MeV/c^2 is used [3]. In addition, a cut on the transverse momentum $p_{\mathrm{T}}(\phi)$ is placed to suppress ϕ mesons produced directly in the pp interactions.

B_s^0 meson reconstruction

Finally, a B_s^0 meson candidate is reconstructed from the J/ψ and the ϕ candidates, again requiring a good vertex fit $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}(B_s^0)$ of the formed B_s^0 vertex. The allowed interval of $m(B_s^0)$ is relatively large in order to have enough events in the mass sidebands to perform background studies. The B_s^0 candidates with a short reconstructed decay time are rejected to reduce a huge amount of the background arising through wrongly combined J/ψ and ϕ particles originating from the PV.

Moreover, besides the simple vertex fit, a kinematic fit of the full decay tree (decay tree fit, DTF), is performed [87]. The parameters of the fit are the vertex positions and momenta of the particles in the decay chain. They are constrained by the momentum conservation at each vertex and an additional constraint is applied in DTF, fixing the mass of the dielectron system to the true J/ψ mass. The B_s^0 candidates are required to have a good χ^2/ndf of the global kinematic fit, $\chi^2_{\mathrm{DTF}}/\mathrm{ndf}(B_s^0)$.

In the end, in events with more than one selected B_s^0, the candidate with the best decay tree fit quality is chosen. This avoids the selection of $B_s^0 \to J/\psi\phi$ candidates with one or more tracks in common.

A total number of the B_s^0 candidates remaining after the stripping and offline selection process is about $6.9{\cdot}10^6$. Their invariant mass distribution with and without constraint on the J/ψ mass is given in Fig. 4.1. The peak around B_s^0 mass is not visible. The number of background events is still very high. To further suppress background contribution two multivariate strategies are compared as discussed in Sec. 4.1.4.

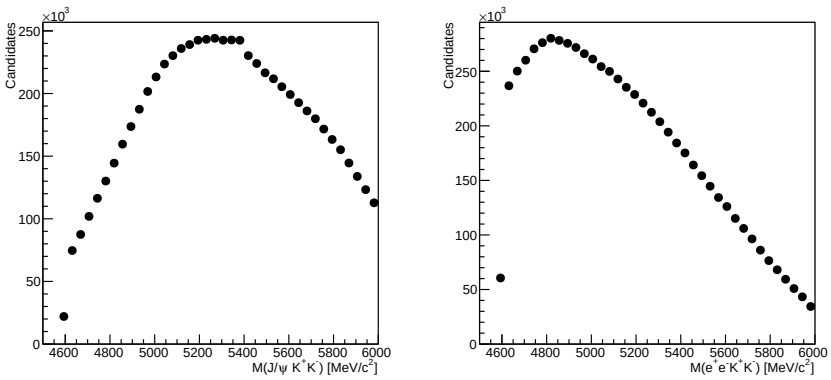


Fig. 4.1 Reconstructed B_s^0 mass distribution of the selected $B_s^0 \to J/\psi\phi$ candidates with and without J/ψ mass constraint.

4.1.3 Monte Carlo reweighting

The PID information is not well represented in simulated events (Fig. 4.2). Therefore, instead of directly applying the selection on the PID variables calculated in the MC, the events are reweighted using PID resampling. This correction is a function of track kinematics (momentum and pseudorapidity) and event multiplicity (number of tracks). The correction is done with an unbinned approach by utilisation of probability density functions (PDFs) in four dimensions (PID, p_{T}, η, $\mathrm{N}_{\mathrm{tracks}}$) that are described by a kernel density estimation procedure using the Meerkat library [88]. The `Urania` package[4] [89] is used to perform the PID resampling for all final state particle types, $K^{\pm}$ and $e^{\pm}$. In case of the charged kaons, the log-likelihood of the kaon hypothesis (ProbNNK) is resampled while the difference of the log-likelihood between electron and pion hypothesis ($\Delta\ln\mathcal{L}_{e\pi}$ is also called PIDe) is reweighted for the electron pair. The corrected PID efficiencies are used to weight the event. After applying PID correction, the ProbNNK and PIDe values for MC are in agreement with data variables (Fig. 4.2). The corrected PIDe($e^{\pm}$) and ProbNNK($K^{\pm}$) variables are used in the further selection.

4.1.4 Multivariate selection requirements

In order to reduce combinatorial background using the signal decay kinematics, a Boosted Decision Tree (BDT) [90] and Neural Network (NN) [91] are used as a multivariate discriminant.

[4]The `Urania v7r0` package is used.

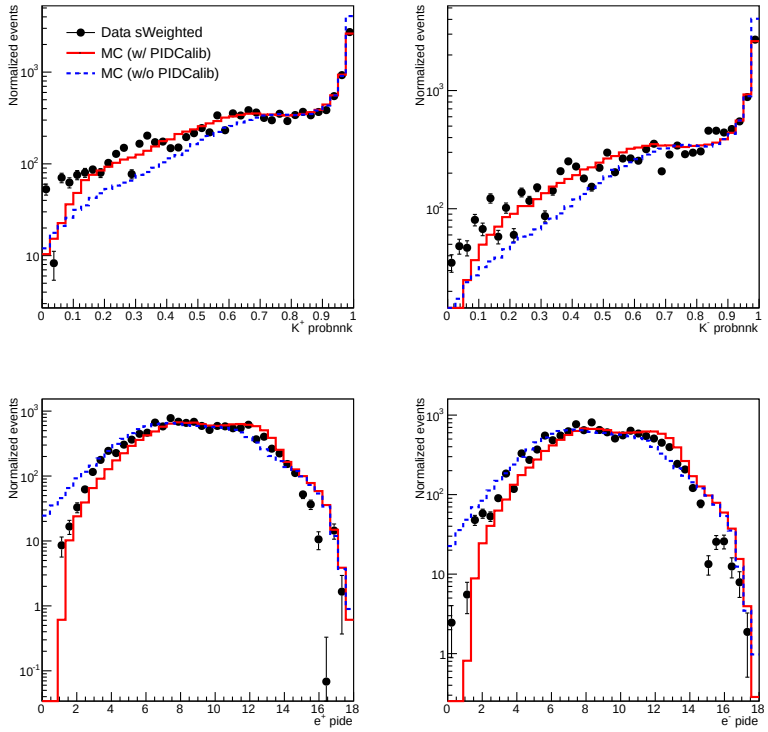


Fig. 4.2 ProbNNK($K^{\pm}$) and PIDe($e^{\pm}$) distributions of the selected $B_s^0 \to J/\psi(e^+e^-)\phi$ candidates for sWeighted data (black points), MC (blue dashed line) and resampled MC (red line). The distributions are in a logarithmic scale.

The multivariate analysis (MVA) is based on the statistical principle of multivariate statistics, which involves observation and analysis of more than one statistical outcome variable at a time. This technique is used to perform trade studies across multiple dimensions while taking into account the effects of all variables on the responses of interest [92].

Boosted Decision Tree

A decision tree is a binary tree structured classifier similar to the one sketched on Fig. 4.3. Repeated left/right (yes/no) decisions are taken on one single variable at a time until a stop criterion is fulfilled. The phase space is split this way into many regions that are eventually classified as a signal or background, depending on the majority of training events that end up in the final leaf node. The boosting of a decision tree is a procedure that combines many classifiers (decision trees) to form a powerful discriminator. The trees are sequentially derived from the same training ensemble by reweighting events based on their misclassification[5], and are finally combined into a single classifier which is given by a (weighted) average of the individual decision trees. Boosting stabilizes the response of the decision trees with respect to fluctuations in the training sample and is able to considerably enhance the performance with respect to a single tree. The BDT response is the combined vote of many individual

[5]Events that are misclassified are given a higher weight in subsequent iterations.

decision trees, derived from the same training sample by boosting events. A more detailed overview of the BDT is given in Ref. [92].

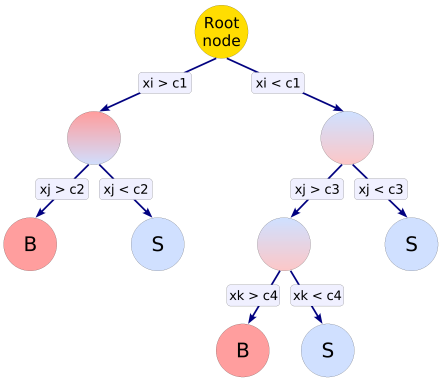


Fig. 4.3 Diagram showing the principle of a decision tree. For a given variable, x_i, the value giving the best separation, c_i, is found. This is repeated for all variables until a given signal purity is reached, or until a minimum number of event in the node is reached. A category is then assigned to the leaf depending on whether signal or background is the dominant contribution.

In the thesis the BDT classifier is considered to discriminate the signal and background candidates based on quantities describing kinematics and quality of the candidates. In order to have a good agreement between data and simulation samples, a simple set of kinematic variables is chosen for the BDT discriminator. In particular, 8 kinematic variables are taken:

- $p_{\mathrm{T}}(J/\psi)$ and $p_{\mathrm{T}}(\phi)$ - transverse momentum of the J/ψ and ϕ mesons;
- $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$ of the reconstructed secondary vertex;
- PIDe - difference of log-likelihood between electron and pion hypothesis for the electron pair;
- ProbNNK - log-likelihood of the kaon hypothesis for charged kaons;
- χ^2 of the constrained kinematical vertex fit of B^0_s meson, χ^2_{DTF}.

The training of the BDT classifier is done using the following samples:

- Signal: $B^0_s \to J/\psi\phi$ simulated sample is used. The sample is required to pass exactly the same stripping and offline selection criteria described above and, furthermore, the MC truth information is used to require that the reconstructed candidate matches with the generated decay.
- Background: a wrong sign $B^0_s \to J/\psi\phi$ data sample is used. In particular, the $B^0_s \to J/\psi(e^+e^+)\phi(K^+K^+)$, $B^0_s \to J/\psi(e^+e^+)\phi(K^-K^-)$, $B^0_s \to J/\psi(e^-e^-)\phi(K^+K^+)$, $B^0_s \to J/\psi(e^-e^-)\phi(K^-K^-)$, $B^0_s \to J/\psi(e^+e^-)\phi\ (K^+K^+)$, $B^0_s \to J/\psi(e^+e^-)\phi(K^-K^-)$, $B^0_s \to J/\psi(e^+e^+)\phi(K^+K^-)$ and $B^0_s \to J/\psi(e^-e^-)\phi(K^+K^-)$ events make the background sample. These events are also required to pass all selection steps.

Table 4.4 Ranking of variable importance used in the BDT training.

Rank	Variable	Importance
1	PIDe$(e^+)_{\text{corr}}$	1.804e-01
2	PIDe$(e^-)_{\text{corr}}$	1.679e-01
3	log(ProbNNK)$(K^+)_{\text{corr}}$	1.424e-01
4	log(ProbNNK)$(K^-)_{\text{corr}}$	1.314e-01
5	$p_{\text{T}}(J/\psi)$	1.051e-01
6	$p_{\text{T}}(\phi)$	1.007e-01
7	$\log(\chi^2_{\text{DTF}})(B_s^0)$	9.268e-02
8	$\chi^2_{\text{vtx}}(B_s^0)$	7.951e-02

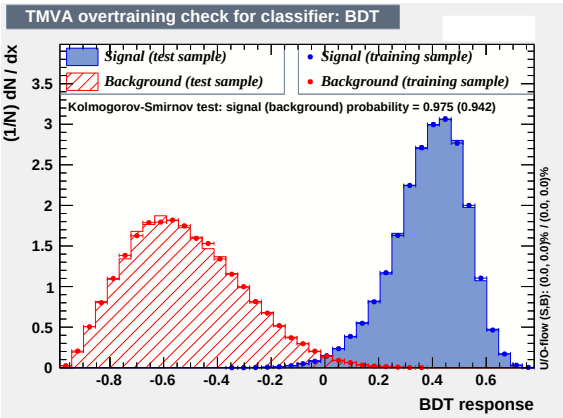


Fig. 4.4 BDT classifier distribution for the signal and background $B_s^0 \to J/\psi\phi$ samples.

The ranking of variable importance used to train the BDT is shown in Table 4.4 where the PIDe$(e^\pm)_{\text{corr}}$ and ProbNNK$(K^\pm)_{\text{corr}}$ variables are reweighted using `Meerkat` tool described in Sec. 4.1.3. The comparison of the signal and background distributions of the BDT input variables is shown in App. C.2. The BDT response distribution (Fig. 4.4) can be used as a univariate discriminant to distinguish signal candidates from the background events. A figure of merit (FoM) value [93] is applied to choose the BDT response criterium. The FoM value is defined as an effective size of the signal from sPlot technique with the signal sWeight w_i [94]:

$$\text{FoM} = \frac{(\sum_{i=1}^{N} w_i)^2}{\sum_{i=1}^{N} w_i^2} \tag{4.1}$$

The dependence of the B_s^0 FoM value on the BDT response and the obtained dependence between efficiency and purity of the BDT selection are shown in Fig. 4.5. The optimal BDT response criterium which reduces the maximal number of background events and has a reasonable efficiency/purity percentage is larger than 0.2.

Neural Network

Another multivariate discriminant tested in the analysis is the Neural Network reviewed in detail in Ref. [95]. The NN is a Multi Layer Perceptron (MLP) feed-forward network with

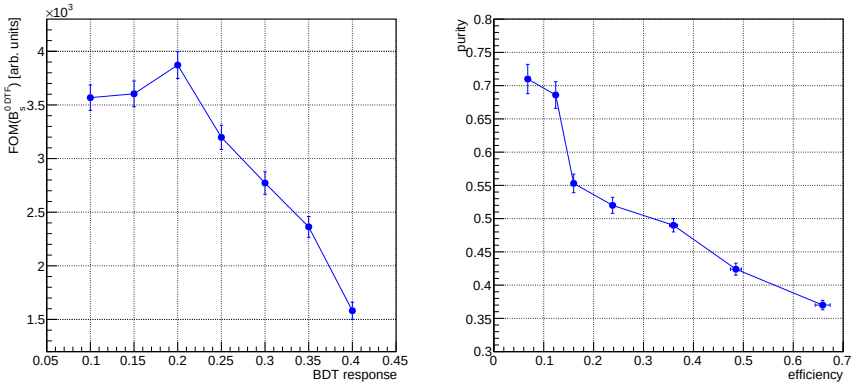


Fig. 4.5 (left) Distribution of the FoM value as a function of the BDT response. (right) Distribution of the efficiency and purity of the signal selection depending on the BDT criteria.

following structure (Fig. 4.6). The input layer contains as many neurons as input variables used in the multivariate analysis. The output layer contains a single neuron for the signal weight. Between the input and output layers there is a variable number of n hidden layers with an arbitrary numbers of neurons.

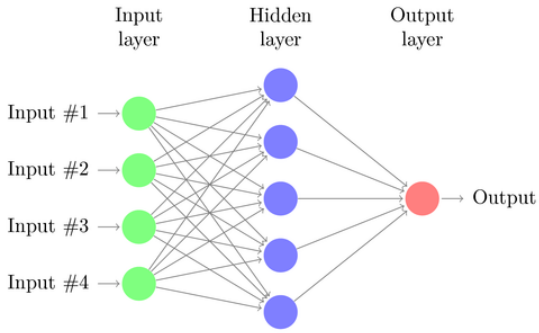


Fig. 4.6 Diagram showing the principle of a Neural Network.

As indicated in Fig. 4.6, all neuron inputs to a layer are linear combinations of the neuron output of the previous layer. The transfer from input to output within a neuron is performed by the activation function $f_{act}(\Sigma)$, where $\Sigma = \vec{x}_{in}^{T} \cdot \vec{w}$ is the dot product of the neuron input vector and neuron weight vector. In general, the activation function of a neuron can be zero (deactivated), one (linear), or non-linear.

A freeware package for building neural networks `NetMaker` [96] is used in the analysis. In order to have a good agreement between data and simulation samples, a set of 7 kinematic variables is selected with corresponding neurons of the input layer. The NN discriminator was chosen to have a single hidden layer composed of 10 neurons with a non-linear activation function, sigmoid $f_{act}(\Sigma) = 1/(1+e^{-\Sigma})$ [97]. In particular, the selected input variables consist of:

- IP(B_s^0) - impact parameter[6] of the reconstructed B_s^0 meson with respect to the PV;

[6]The Impact Parameter (IP) is the distance of the closest approach of a track to the PV. It is used in the

- χ^2/ndf of the reconstructed B_s^0 vertex;
- $p_{\mathrm{T}}(J/\psi)$ and $p_{\mathrm{T}}(\phi)$ - transverse momentum of the J/ψ and ϕ mesons;
- PIDe - difference of log-likelihood between electron and pion hypothesis for e^+ particle;
- ProbNNK - log-likelihood of the kaon hypothesis for positive kaon;
- $\chi^2_{\mathrm{IP}}(e^+)$ of the impact parameter for the positron.

The training of the NN discriminator is done using the same samples as for the BDT training:

- Signal: $B_s^0 \to J/\psi\phi$ simulated sample is used. This sample is required to pass the stripping and offline criteria described in 4.1.2. Furthermore, the MC truth information is used to ensure that the reconstructed candidate matches with the generated decay.
- Background: a wrong sign $B_s^0 \to J/\psi\phi$ data sample is used. In particular, $J/\psi \to e^\pm e^\pm$ and $\phi \to K^\pm K^\pm$ events make the background sample. Also, these events are required to pass all selection steps.

The NN discriminator is trained in a supervised process to obtain the desired network answers for events in the training data set. For the signal events a target vector of 1 is chosen, whereas for the background events a target vector of 0 is used. The target vectors are also used to calculate the error when the network is tested.

The obtained NN response distribution and efficiency-purity curves for training and testing samples are presented in Fig. 4.7a-b. The criterium of the NN response is defined using the same FoM definition as used to determine the BDT response criterium. The Fig. 4.7c shows the B_s^0 FoM value as a function of the NN response. The optimal NN criterium which reduces the maximal number of background events is larger than 0.997.

The fit to the B_s^0 mass distribution with J/ψ mass constraint (Fig. 4.8) assists to determine which of the multivariate analysis provides better background reduction. In both discriminants, the fit model of the B_s^0 mass distribution with J/ψ mass constraint is defined as:

$$S_M(x; f, \mu, \sigma) = \sum_{i=i}^{2} f_i \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-\mu)^2}{2\sigma_i^2}} + f_3 \frac{1}{(x-\mu)^2 + \frac{1}{4}\sigma_3^2}, \tag{4.2}$$

where the first term is a sum of two Gaussian functions [98] with fractions f_i, a mass mean μ and the widths σ_i. The second term represents a Breit-Wigner function [99] with fraction $f_3 = 1 - f_1 - f_2$ and width σ_3. An exponential (Eq. 4.6) and Gaussian functions describe combinatorial and partially reconstructed background events determined in detail in Sec. 4.2.1, respectively.

selections and the LHCb trigger to identify long-lived particles and reject short-lived background.

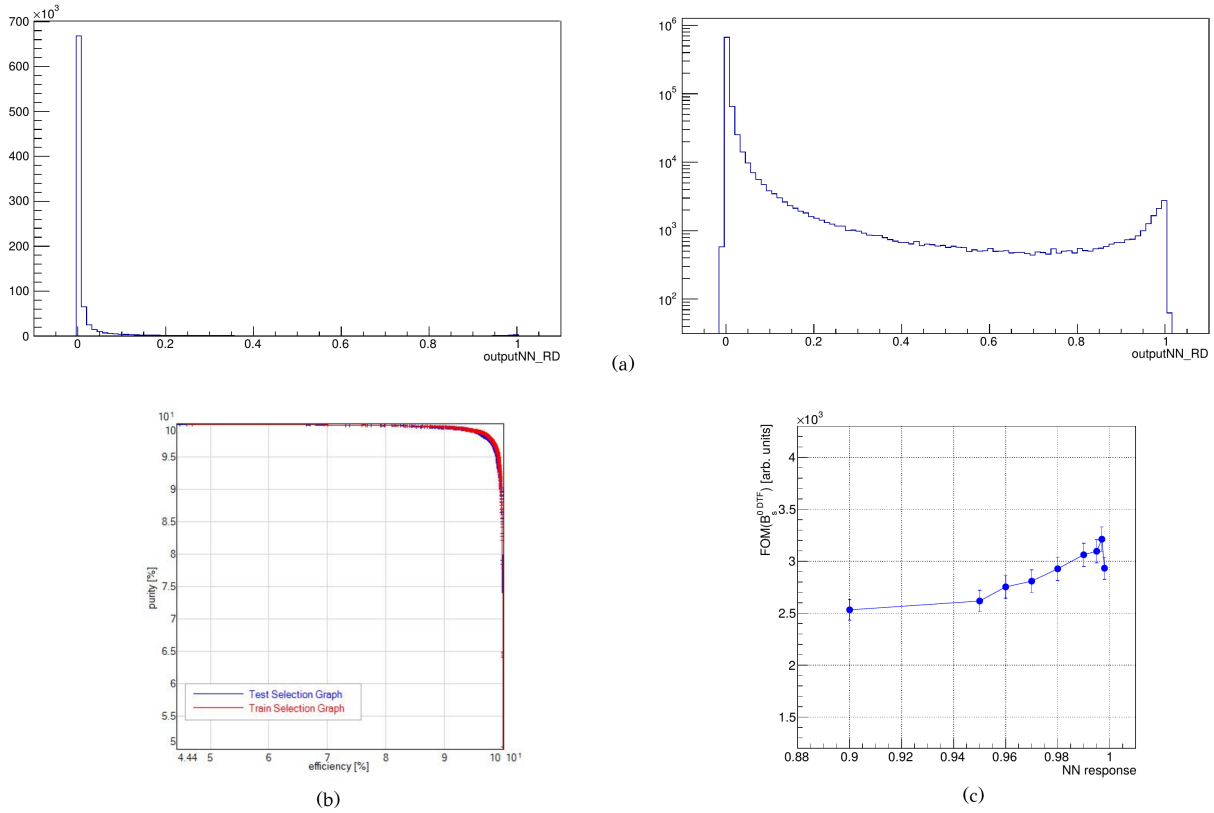


Fig. 4.7 (a) NN classifier distribution for $B_s^0 \to J/\psi\phi$ training sample where the right figure is in a logarithmic scale. (b) Efficiency distribution as a function of purity distribution for training and testing samples. Testing result is shown as a blue line while the training result is a red line. (c) NN response distribution as a function of the FoM value distribution.

The performance of the two multivariate analyses is verified on 2011 data set corresponding to a 1.0 fb^{-1} integrated luminosity. The B_s^0 mass fit and their parameters for both discriminators are shown in Fig. 4.8 and Table 4.5. A similar number of background events remain for both selections while the BDT selection retains more signal candidates. The larger FoM values are observed for the BDT selection than after NN selection as presented in Figs. 4.5 and 4.7c. Thus, the BDT selection is used for further analysis.

4.1.5 Peaking background

Two possible sources of the background that the peak under the signal B_s^0 peak are the $\Lambda_b \to J/\psi pK$ decays where the p has been misidentified as a K^+, and $B^0 \to J/\psi K^*(892)^0$ decays where a π^- from the $K^*(892)^0 \to K^+\pi^-$ decay[7] has been misidentified as a K^-.

In order to better understand the structure of these backgrounds in the reconstructed B_s^0 mass distribution, the simulated events of the $\Lambda_b \to J/\psi pK$ and $B^0 \to J/\psi K^*$ decays are reconstructed with $B^0 \to J/\psi(e^+e^-)\phi(K^+K^-)$ hypothesis and the full selection procedure

[7] In the following, the $K^*(892)^0$ is indicated as the K^*, otherwise it is indicated.

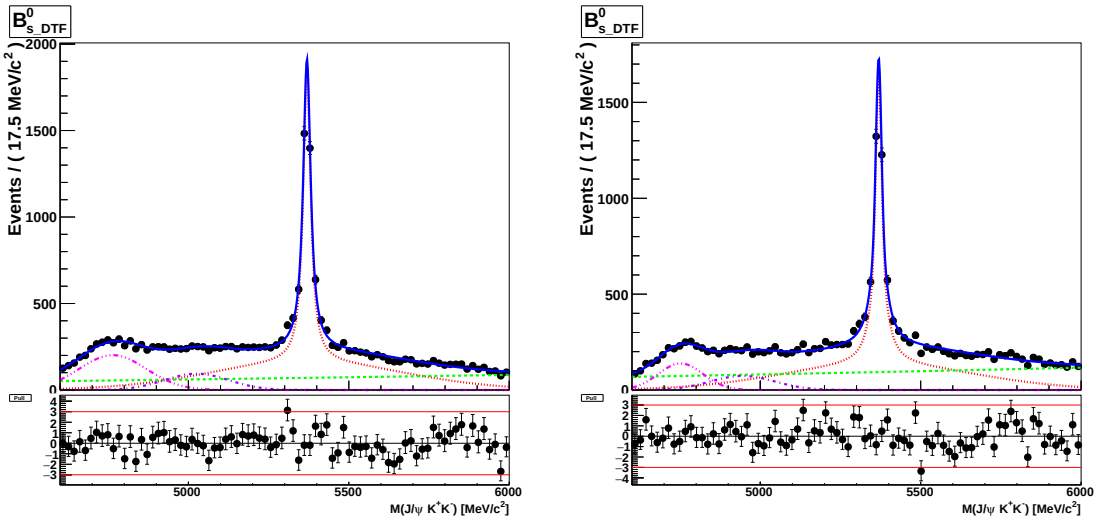


Fig. 4.8 Fit of the B_s^0 mass distribution with J/ψ mass constraint (left) after BDT selection and (right) after NN selection. The blue line shows the total fit which includes the signal (red line) and combinatorial background (green line) contributions. A partially reconstructed background from the $B_s^0 \to \psi(2S)\phi$ and $B_s^0 \to \chi_{c1}(1P)\phi$ decays is indicated by pink and purple lines, respectively (Sec. 4.2.1).

Table 4.5 Result of the fit to the B_s^0 mass distribution with J/ψ mass constraint after BDT selection and after NN selection.

Parameter	BDT	NN
σ_1[MeV/c^2]	272±155	298±4.8
σ_2[MeV/c^2]	79±29	3.9±6.5
σ_3[MeV/c^2]	25±1.6	25±3.0
f_1	0.33±0.031	0.71±0.019
f_2	0.12±0.030	0.003±0.010
μ[MeV/c^2]	5369±0.30	5369±0.52
$\sigma_{B_s^0\to\psi(2S)\phi}$[MeV/c^2]	88±4.3	120±86
$\mu_{B_s^0\to\psi(2S)\phi}$[MeV/c^2]	4759±3.9	4807±8.8
$\sigma_{B_s^0\to\chi_{c1}(1P)\phi}$[MeV/c^2]	100±2.1	99±2.4
$\mu_{B_s^0\to\chi_{c1}(1P)\phi}$[MeV/c^2]	5029±15	5075±25
α_{slope}	2.6±1.5	0.14±0.18
$N_{B_s^0\to\psi(2S)\phi}$	2842±75	1356±59
$N_{B_s^0\to\chi_{c1}(1P)\phi}$	1351±77	1054±72
N_{cbkg}	5501±177	7331±178
N_{sig}	10989±165	9362±156

is applied (Sec. 4.1.2, 4.1.4). The mass distribution of background contributions under a signal decay hypothesis is shown in Fig. 4.9 with and without J/ψ mass constrain. The mass distribution of the B^0 meson decay is flat within the interesting mass range, and no clear peaking background contribution is observed. In case of the small contribution from $\Lambda_b \to J/\psi pK$ decay, its branching ratio is taken into account from Ref. [100] to estimate its exact contribution within the signal region. It is found to be 1.03% of the events under the B_s^0 meson peak.

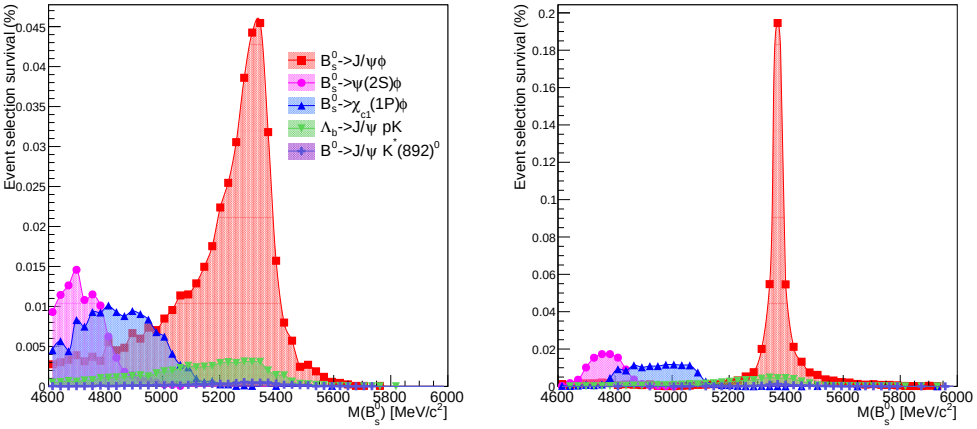


Fig. 4.9 Mass distribution of the reconstructed B_s^0 candidates (left) without and (right) with J/ψ mass constraint. The red area corresponds to simulated signal events. The simulated $\Lambda_b \to J/\psi pK$ (green area) and $B^0 \to J/\psi K^*$ decays (violet area) that are mis-reconstructed as the $B_s^0 \to J/\psi\phi$ due to pK and $K\pi$ misidentification or due to partially random combinations with other tracks. The pink and blue areas show the partially reconstructed background contribution described in Sec. 4.2.1.

To reduce the contribution from the Λ_b meson decay, two additional selections are considered: BDT discriminator or a simple cut on ProbNNp variable. For the first selection, a new BDT training is performed using the $B_s^0 \to J/\psi\phi$ simulated data as a signal sample while the $\Lambda_b \to J/\psi pK$ MC data reconstructed as a signal channel is used as a background sample. The same set of the BDT input variables is used as in the main BDT discriminator (Sec. 4.1.4) except for the ProbNNK(K^+) variable. This kinematic variable is exchanged with the ProbNNp(K^+) to maximize the sensitivity for events with protons in the final state. Fig. 4.10 and Table 4.6 present the BDT classifier distribution and the ranking of training variable importance, respectively. The comparison between signal and background distributions of the BDT input variables is shown in App. C.3.

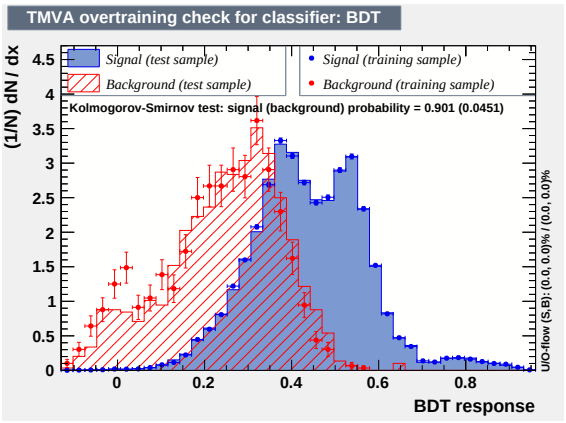


Fig. 4.10 BDT classifier distribution for the signal $B_s^0 \to J/\psi\phi$ and $\Lambda_b \to J/\psi pK$ background samples trained to distinguish the peaking background events.

Table 4.6 Ranking of variable importance used in the BDT training to distinguish the peaking background events.

Rank	Variable	Importance
1	$\log(\text{ProbNNK})(K^-)_{\text{corr}}$	2.727e-01
2	$\log(\text{ProbNNK})(p)_{\text{corr}}$	2.669e-01
3	$p_{\text{T}}(\phi)$	9.546e-02
4	$p_{\text{T}}(J/\psi)$	8.873e-02
5	$\text{PIDe}(e^-)_{\text{corr}}$	8.855e-02
6	$\text{PIDe}(e^+)_{\text{corr}}$	7.963e-02
7	$\log(\chi^2_{\text{DTF}})(B_s^0)$	6.402e-02
8	$\chi^2_{\text{vtx}}(B_s^0)$	4.399e-02

In case of the second selection, a direct cut on the log-likelihood of the proton hypothesis for a positive kaon, ProbNNp(K^+), is applied. The ProbNNp in the signal simulated sample is resampled as described in Sec. 4.1.3 and shown in Fig. 4.11. The signal $B_s^0 \to J/\psi\phi$ data and simulation have a similar ProbNNp distribution while for Λ_b MC sample a completely different structure is observed.

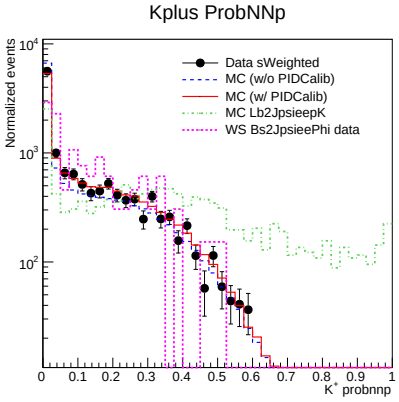


Fig. 4.11 ProbNNp(K^+) distribution for the signal $B_s^0 \to J/\psi\phi$ data sample (black), $B_s^0 \to J/\psi\phi$ simulation sample before (blue) and after (red) MC reweighting, wrong-sign $B_s^0 \to J/\psi\phi$ data (violet) and the $\Lambda_b \to J/\psi pK$ MC sample (green).

The ratio between the signal and background events determines a FoM of peaking background criteria:

$$\text{FoM} = \frac{\text{N}_{\text{sig}}}{\sqrt{\text{N}_{\text{sig}} + \text{N}_{\text{bkg}}}}, \tag{4.3}$$

where N_{sig} is the number of the $B_s^0 \to J/\psi\phi$ simulated events that pass the appropriate BDT or ProbNNp selection. The N_{bkg} is the number of the $\Lambda_b \to J/\psi pK$ MC events which pass the same selection. The resulting FoM distributions as a function of the BDT response and ProbNNp(K^+) variable are presented in Fig. 4.12. The signal-to-background ratio starts to decrease for values of the BDT response larger than 0.15, whereas for the ProbNNp selection a decrease is observed for cut values below 0.6. The dependence of efficiency on purity shows that there is no significant difference between the two selections within the uncertainty. Thus,

the selection of ProbNNp(K^+)<0.6 is chosen as the optimal criterion for further analysis. The background rejection of this selection, with respect to the number of left over $\Lambda_b \to J/\psi pK$ events after the main BDT selection (Sec. 4.1.4), is 19.8%, evaluated on the Λ_b component of the inclusive MC sample. The corresponding signal efficiency evaluated on the signal simulated sample is 99.3%.

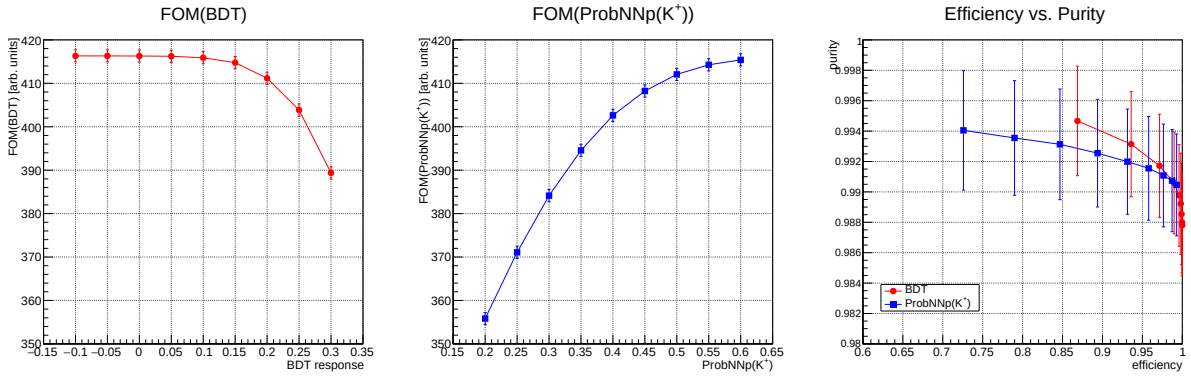


Fig. 4.12 Distribution of the FoM value as a function of the (left) BDT response and (middle) ProbNNp(K^+) variable selection. (right) Distribution of the efficiency and purity depending on the BDT (red) and ProbNNp(K^+) (blue) selection.

4.2 Mass fitting procedure

The physics parameters of interest are extracted via a log-likelihood fit of the signal PDF to the unbinned decay time and angular distributions as defined in Eq. 1.32. For the first step, the events are weighted to statistically subtract the combinatorial and partially reconstructed background using the sPlot method [94] with the $e^+e^-K^+K^-$ mass as the discriminating variable.[8]

In order to generate the sWeights for use in the time dependent angular analysis, a fit is performed to the reconstructed $m(e^+e^-K^+K^-)$ distribution. The mass model used for the signal candidates is a double Crystal Ball function [101]:

$$S_M(x;\alpha,n,\mu,\sigma) = N \begin{cases} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, & \text{for } \frac{x-\mu}{\sigma} > -\alpha \\ A(B-\frac{x-\mu}{\sigma})^{-n} & \text{for } \frac{x-\mu}{\sigma} \leq -\alpha \end{cases} \tag{4.4}$$

where

$$\begin{aligned} A = \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{|\alpha|^2}{2}}, \quad B = \frac{n}{|\alpha|} - |\alpha|, \quad N = \frac{1}{\sigma(C+D)}, \\ C = \frac{n}{|\alpha|}\frac{1}{n-1}e^{-\frac{|\alpha|^2}{2}}, \quad D = \sqrt{\frac{\pi}{2}}\left(1+erf\left(\frac{|\alpha|^2}{2}\right)\right). \end{aligned} \tag{4.5}$$

[8] This is referred to as sWeighting and the weights are called sWeights.

The n is a normalization factor and erf is the error function [102]. This function describes a line shape composed of a central Gaussian (mean μ, standard deviation σ) with a decreasing power law tail smoothly connected to the Gaussian at $(\mu-\alpha\sigma)$. The combinatorial background is modelled as an exponential function (Eq. 4.6). The partially reconstructed background from the $B_s^0 \to \psi(2S)\phi$ and $B_s^0 \to \chi_{c1}(1P)\phi$ decays described in Sec. 4.2.1 are characterized using double and single Gaussian function (first term of Eq. 4.2), respectively. To better describe the left tail of the B_s^0 mass distribution arising due to Bremsstrahlung radiation [103], the $m(e^+e^-K^+K^-)$ distribution is split into three Bremsstrahlung categories depending on the number of recovered photon clusters: no Bremsstrahlung photon found (0γ), one photon found (1γ) or more (2γ). Fig. 4.13 shows the fit to the $m(e^+e^-K^+K^-)$ distribution in the simulated and data samples of the $B_s^0 \to J/\psi\phi$ events. The B_s^0 mass region (4700,5600) MeV/c^2 is applied to limit the data size for the analysis. The model parameters are fixed to the MC fit to better describe the shape of the mass distribution in regions with large background. To only fix the shape and take into account different sizes and resolution of mass determination between signal and simulated samples, the parameters of the mass PDF are taken from the MC except for the mean mass value μ and width σ of Gaussian distributions. The values of fitted parameters are given in Table 4.7, along with the signal and background yields for each Bremsstrahlung category. Since these categories are independent, the final sWeights are computed as a sum of Bremsstrahlung category fits and are used for further analysis. The number of signal $B_s^0 \to J/\psi(e^+e^-)\phi$ candidates obtained on 2011 and 2012 data sets is 12195±497 which corresponds to 12.7% of the leading muon mode of the $B_s^0 \to J/\psi\phi$ statistics [39]. Although, the masses of the daughter particles are not used in the fit, a simultaneous fit to the $J/\psi K^+K^-$, ϕ and J/ψ mass distributions is performed for illustration and is included in App. C.4.

Table 4.7 Results of the fit to the $m(e^+e^-K^+K^-)$ distribution in data sample divided into three Bremsstrahlung categories. The scale parameter is a difference between sigma of the first and second Crystal Ball function and $n_1 = n_2 = n$. The shape of the partially reconstructed background is fixed to MC fit reported in Table 4.8.

Parameter	0γ	1γ	2γ	
α_1	0.134±0.005	0.199±0.006	0.37±0.01	fixed to MC fit
α_2	-1.28±0.3	-0.742±0.02	-0.546±0.02	fixed to MC fit
σ[MeV/c^2]	29.0±3.7	44.9±1.4	69.7±8.3	float
scale	1.15±0.1	1.5±0.05	0.96±0.03	fixed to MC fit
μ[MeV/c^2]	5339.8±3.9	5337.1±2.6	5352.9±4.9	float
f_{CB1}	0.987±0.008	0.7±0.006	0.59±0.01	fixed to MC fit
n	5.3±0.5	30.0±6.0	21.0±5.0	fixed to MC fit
α_{slope}	-0.00037±0.0001	-0.00053±0.00005	-0.0005±0.00008	float
$\mathrm{N}_{\mathrm{cbkg}}$	8142±344	23365±365	17113±410	
$f_{B_s^0\to\psi(2S)\phi}$	0.387±0.014	0.432±0.008	0.453±0.010	float
$\mathrm{N}_{B_s^0\to\psi(2S)\phi}$	127±38	298±79	203±85	
$\mathrm{N}_{B_s^0\to\chi_{c1}(1P)\phi}$	200±105	391±198	245±206	
$\mathrm{N}_{\mathrm{sig}}$	3146±268	6159±246	2890±339	

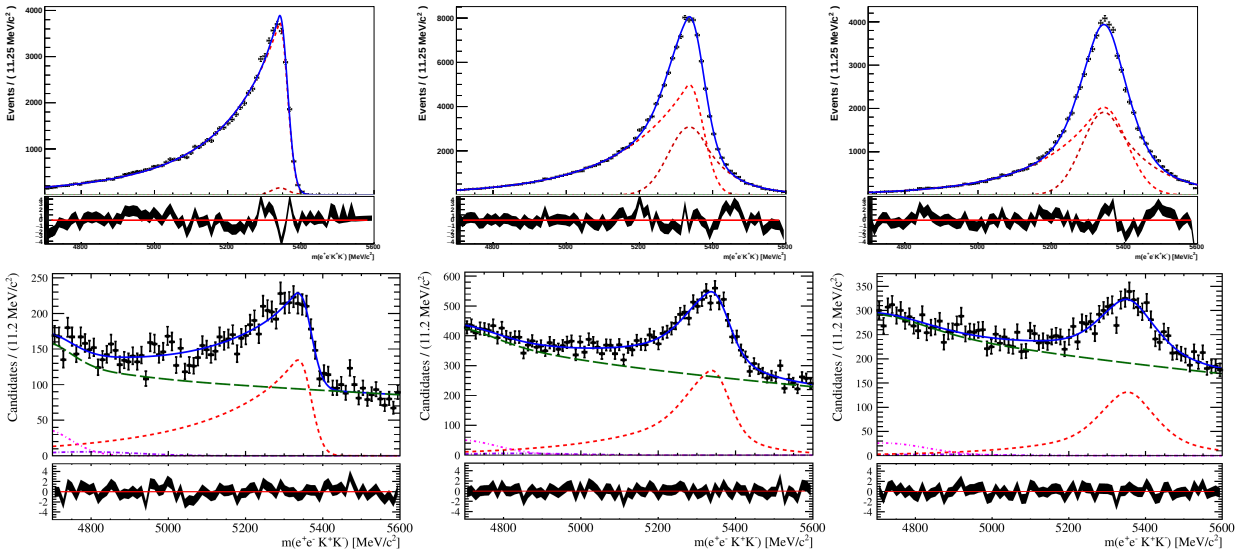


Fig. 4.13 Distribution of the $m(e^+e^-K^+K^-)$ in the (top) simulated and (bottom) data samples of $B_s^0 \to J/\psi\phi$ events divided into three Bremsstrahlung categories: (left) without, (middle) one and (right) more photons originating from the electron radiation. The blue line shows the total fit which includes the signal (red line) and combinatorial background (green line) contributions. The contribution of partially reconstructed background from the $B_s^0 \to \psi(2S)\phi$ and $B_s^0 \to \chi_{c1}(1P)\phi$ decays is indicated by pink and purple lines, respectively.

4.2.1 Background studies

Combinatorial background

The random combinatorial background arises when within the four final tracks, $e^{\pm}$ and $K^{\pm}$, not all originate from the B_s^0 meson decay. This is modelled by a normalized exponential function:

$$B_M(x;\alpha_{\text{slope}}) = \frac{e^{-\alpha_{\text{slope}}\cdot x}}{\int_{min}^{max} e^{-\alpha_{\text{slope}}\cdot x}\mathrm{d}x}, \tag{4.6}$$

where the slope parameter α_{slope} is estimated in the fit. The integral in the normalization is calculated from the lower (min) to the upper (max) edge of the fitted mass region.

Partially reconstructed background

The partially reconstructed background arises from the true B meson decays but with one or more tracks missing from the reconstruction. In case of the $B_s^0 \to J/\psi(e^+e^-)\phi$, there are two sources of the partially reconstructed events: those from the hadronic part, such as events with $\phi(1020)$ resonances (called partially reconstructed hadronic background in the following), and those from the J/ψ part (called partially reconstructed J/ψ background in the following), such as events from $\psi(2S)$ and $\chi_{c1}(1P)$ meson decays. In order to study this background, $8{\cdot}10^6$ inclusive $B_s^0 \to J/\psi(e^+e^-)X$ simulated events are produced. In all 10 977 events passed the selection criteria (Sec. 4.1.2) as well as a veto on true $B_s^0 \to J/\psi(e^+e^-)\phi$ events. Most of the events originate from the J/ψ part (black points in Fig. 4.14) where

41.30% is from $\psi(2S)$ and 55.56% from $\chi_{c1}(1P)$ decaying into the J/ψ meson and neutral particles. The remaining part of the events is produced from the $\chi_{c0}(1P)$ (0.49%), $\chi_{c2}(1P)$ (1.59%) and $h_c(1P)$ (1.03%) meson decays. Their contributions to the B_s^0 mass distribution are presented in Fig. 4.15. Only 4% of the partially reconstructed background is due to the hadronic part (blue points in Fig. 4.14) which is the $\eta^{'}$, B_s^{0*} and $b\bar{b}$ decays.

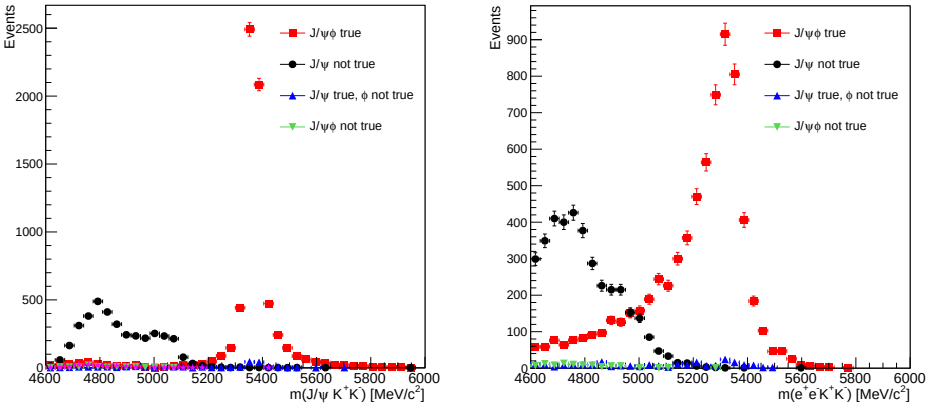


Fig. 4.14 Background from the partially reconstructed events due to missing particles from the J/ψ and/or hadronic part: B_s^0 mass distribution (left) with and (right) without J/ψ mass constraint. The true signal events are shown by red squares.

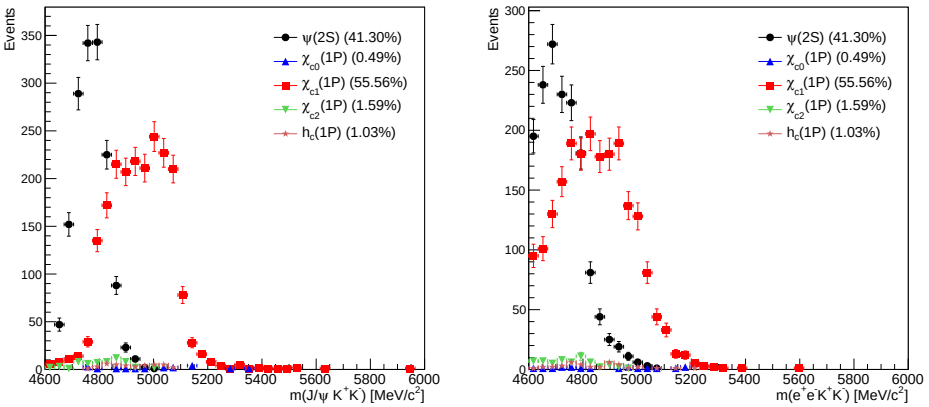


Fig. 4.15 J/ψ partially reconstructed background from the excited charmonium resonances for B_s^0 mass distribution (left) with and (right) without J/ψ mass constraint.

The $B_s^0 \to J/\psi X$ simulated data sample is used to distinguish the different types of partially reconstructed background in the B_s^0 mass distribution of the data sample. The mass fit is applied to the simulated events that give the largest contribution in background studies: $B_s^0 \to \psi(2S)\phi$ and $B_s^0 \to \chi_{c1}(1P)\phi$ decay modes. Their simulated data samples are also divided into three Bremsstrahlung categories and the mass fit is shown in Fig. 4.16. The B_s^0 mass range (4600,5600) MeV/c^2 is applied to better describe the distribution shape. The obtained fit parameters are reported in Table 4.8. The mean μ, width σ and fraction f are

used in the mass fit performed on data sample (Table 4.7). In order to better describe the behaviour of the partially reconstructed background contribution in data, the fraction of the $B_s^0 \to \psi(2S)\phi$ events to the number of partially reconstructed background[9] is calculated from the simulation and used as a Gaussian constraint for the mass fit on the data sample, $f_{B_s^0 \to \psi(2S)\phi}$.

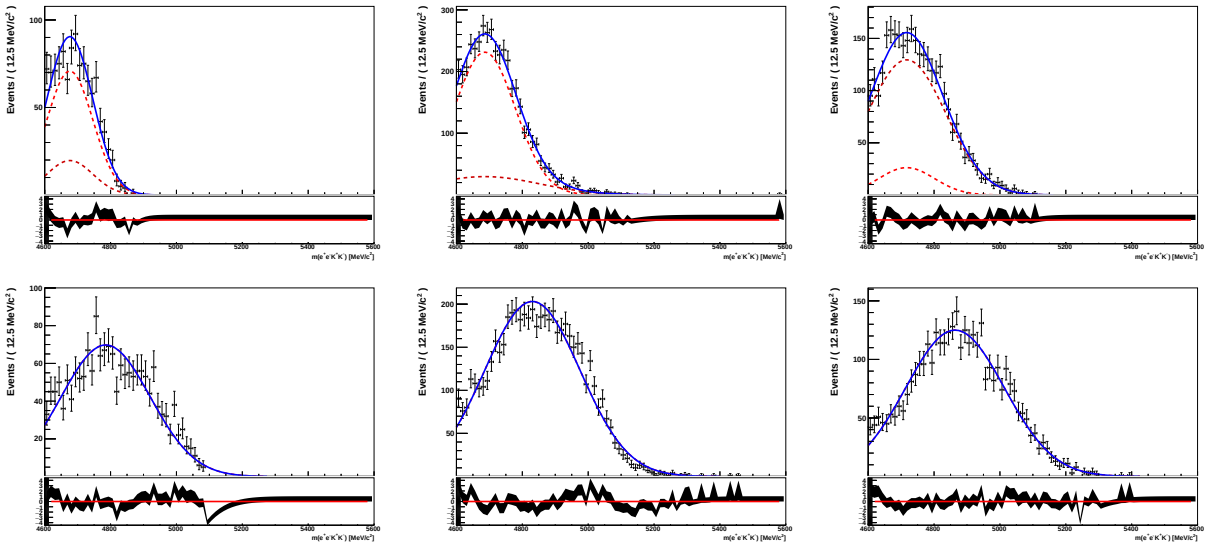


Fig. 4.16 Distribution of the $m(e^+e^-K^+K^-)$ for the (top) $B_s^0 \to \psi(2S)\phi$ and (bottom) $B_s^0 \to \chi_{c1}(1P)\phi$ simulated data sample divided into three Bremsstrahlung categories: (left) without, (middle) one and (right) more photons coming from the electron radiation. The blue line shows the total fit, in case of the $B_s^0 \to \psi(2S)\phi$ decay, consisting of two Gaussian functions (red line).

Table 4.8 Results of the fit to the $m(e^+e^-K^+K^-)$ distribution in the $B_s^0 \to \psi(2S)\phi$ and $B_s^0 \to \chi_{c1}(1P)\phi$ simulated data sample divided into three Bremsstrahlung categories. The shape is modelled by a double or single Gaussian function. The last line shows a fraction of the $B_s^0 \to \psi(2S)\phi$ events with respect to the number of partially reconstructed background.

Parameter	0γ	1γ	2γ
	$B_s^0 \to \psi(2S)\phi$		
μ[MeV/c^2]	4676±4	4686±3	4717±4
σ_1[MeV/c^2]	70±3	93±5	84±31
σ_2[MeV/c^2]	70±7	166±17	119±6
f_1	0.8±0.7	0.84±0.08	0.1±0.2
N_{evt}	1093±33	4238±65	2983±55
	$B_s^0 \to \chi_{c1}(1P)\phi$		
μ[MeV/c^2]	4785±5	4831±2	4864±3
σ[MeV/c^2]	135±3	145±2	150±2
N_{evt}	1729±42	5570±75	3605±60
$f_{B_s^0 \to \psi(2S)\phi}$	0.387±0.014	0.432±0.008	0.453±0.010

[9]The number of partially reconstructed background is a sum of the $B_s^0 \to \psi(2S)\phi$ and $B_s^0 \to \chi_{c1}(1P)\phi$ events.

4.3 Decay time acceptance and resolution

The sensitivity on the phase ϕ_s is greatly improved through the precise measurement of the B_s^0 meson decay time. Different acceptance effects have to be taken into account to determine the decay time and resolution.

4.3.1 Decay time acceptance

The detector reconstruction, trigger and selection requirements (Sec. 4.1) introduce acceptance effects on the decay time of the B_s^0 meson. The acceptance correction is calculated using a data driven method [42] relying on the $B^0 \to J/\psi(e^+e^-)K^*(892)^0$ decay mode[10] and the known value of the B^0 meson lifetime [20].

Selection of the $B^0 \to J/\psi(e^+e^-)K^*$ sample

The $B^0 \to J/\psi K^*$ decay is selected in a similar way to the signal $B_s^0 \to J/\psi\phi$ decay mode (Sec. 4.1.2). The `Bhadron` stream and `StrippingBd2eeKstarBDTLine` are used to perform the stripping selection of the decay candidates. Since the final state of the control channel has $K^* \to K^+\pi^-$ instead $\phi \to K^+K^-$, it has to be taken into account in the offline selection (Table 4.9)[11]. The selections between the two channels differ in the following criteria:

- The final state of the K^* meson decay includes K^+ and π^- while a charged kaon pair is a product of ϕ decay. The pion candidate is required to have a small probability of being a kaon, $\Delta\ln\mathcal{L}_{K\pi} < 5$. The $K^+\pi^-$ invariant mass has to be in an interval around the true K^* meson mass, world average $m(K^*) = 891.66\pm0.26$ MeV/c^2 is used [3].

- The B^0 meson candidates are formed from combinations of the J/ψ and K^* candidates with an invariant mass in the range $|m(B^0) - 5280|$ <1000 MeV/c^2 where $m(B^0) = 5279.62\pm0.15$ MeV/c^2 is the world average [3].

As for the signal decay mode a MC reweighting (Sec. 4.1.3, Fig. C.8) and trigger selection using the same trigger lines except for `Hlt2IncPhiDecision` (Sec. 4.1.1) are applied to the selected $B^0 \to J/\psi K^*$ candidates. Due to lower B^0 meson mass, $m(B_s^0) - m(B^0)$=87.35±0.20 MeV/c^2 [3], one more difference in the B^0 meson decay selection is required. In order to better determine the $m(e^+e^-K^+\pi^-)$ distribution, a wider range of the J/ψ meson mass is considered: $m(J/\psi) \in$[2200, 4200] MeV/c^2.

For the BDT selection, a set of 8 kinematic variables is chosen where 7 variables are the same as used in the BDT discriminator of the signal decay mode (Sec. 4.1.4). In case of the control decay, the log-likelihood kaon hypothesis for K^- candidate is replaced with a

[10] In the following, the $B^0 \to J/\psi(e^+e^-)K^*(892)^0$ decay used to determine the decay time acceptance is called a control mode.

[11] DIRA (direction angle) is a cosines of the angle between the momentum of the particle and the flight direction from the best PV to the decay vertex.

Table 4.9 Stripping and offline selection criteria of the $B^0 \to J/\psi K^*$ candidates.

Decay mode	Cut parameter	Stripping	Offline selection
$J/\psi \to e^+e^-$	$\Delta\ln\mathcal{L}_{e\pi}(e)$	>-2	-
	$\chi^2_{\rm IP}(e)$	>1	-
	$\chi^2_{\rm track}/{\rm ndf}(e)$	<5	<3
	$p_{\rm T}(e)$	>200 MeV/c	>500 MeV/c
	${\rm GhostProb}_{\rm track}(e)$	<0.5	-
	$\chi^2_{\rm vtx}/{\rm ndf}(J/\psi)$	<16	-
	$\chi^2_{\rm DOCA}(J/\psi)$	<30	-
	$p_{\rm T}(J/\psi)$	-	>400 MeV/c
	$m(J/\psi)$	$\in$[2200, 4200] MeV/c^2	-
$K^* \to K^+\pi^-$	$\Delta\ln\mathcal{L}_{K\pi}(K,\pi)$	>-5,<10	>0,<5
	$\chi^2_{\rm IP}(K,\pi)$	>4	-
	$\chi^2_{\rm track}/{\rm ndf}(K,\pi)$	<5	<3
	$p_{\rm T}(K,\pi)$	>400(250) MeV/c	>200 MeV/c
	$p(K,\pi)$	>3000(2000) MeV/c	>2000 MeV/c
	${\rm GhostProb}_{\rm track}(K,\pi)$	<0.35	<0.5
	$p_{\rm T}(K^*)$	-	>1000 MeV/c
	$\chi^2_{\rm vtx}/{\rm ndf}(K^*)$	<16	<9
	$\chi^2_{\rm vtx}(K^*)$	<25	-
	$\chi^2_{\rm DOCA}(K^*)$	<30	-
	$\lvert m(K^+\pi^-) - 892\rvert$	<150 MeV/c^2	<100 MeV/c^2
$B^0 \to J/\psi K^*$	$\lvert m(B^0) - 5280\rvert$	<1000 MeV/c^2	<1000 MeV/c^2
	$\chi^2_{\rm vtx}/{\rm ndf}(B^0)$	<16	-
	${\rm DIRA}(B^0)$	>0.999	-
	t	-	>0.3 ps

log-likelihood pion hypothesis for π^-. The BDT training is performed using the following samples:

- Signal: $B^0 \to J/\psi K^*$ simulated sample is used. The sample is required to pass the stripping and offline selection criteria described above and, furthermore, the MC truth information is used to require that the reconstructed candidate matches with the generated decay.
- Background: $B^0 \to J/\psi K^*$ data sample from upper sideband $m(B^0) > 5600$ MeV/c^2 is used. Also these events are required to pass all selection steps.

The ranking of the input variable importance used to train the BDT is shown in Table 4.10. The input variable distributions and correlation matrices for the signal and background samples are given in App. C.5.1. A good separation of the signal from background events is observed in the BDT classifier distribution for the control decay mode (Fig. 4.17). The value of the BDT criterion as applied for the B^0_s meson decay (Sec. 4.1.4) is used to select the B^0 signal candidates: BDT>0.2. Possible impact of different BDT criterion is discussed in systematics studies in Sec. 4.7.6. The comparison of the BDT input variable distributions between the signal and control decay data is reported in App. C.5.3.

The resulting $e^+e^-K^+\pi^-$ invariant mass distribution for simulated and data samples is shown in Fig. 4.18 and corresponds to (4500,5600) MeV/c^2 mass range. The mass distribution

Table 4.10 Ranking of the variable importance used in the $B^0 \to J/\psi K^*$ BDT training.

Rank	Variable	Importance
1	PIDe$(e^+)_{\text{corr}}$	1.937e-01
2	$\chi^2_{\text{vtx}}(B^0)$	1.800e-01
3	PIDe$(e^-)_{\text{corr}}$	1.626e-01
4	$p_{\text{T}}(J/\psi)$	1.600e-01
5	log(ProbNNK)(K^+)	1.205e-01
6	log$(\chi^2_{\text{DTF}})(B^0)$	6.701e-02
7	$p_{\text{T}}(K^*)$	6.692e-02
8	log(ProbNNK)(π^-)	4.914e-02

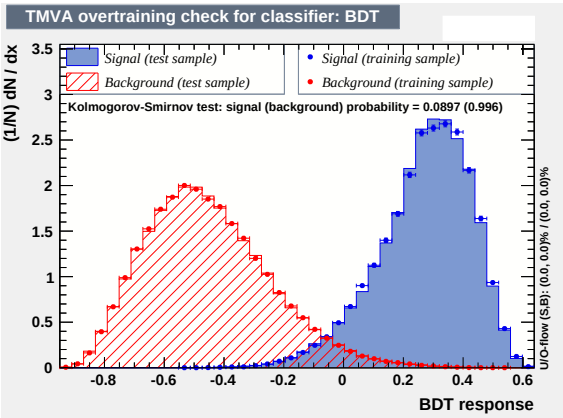


Fig. 4.17 BDT classifier distribution for the signal and background $B^0 \to J/\psi K^*$ samples.

for signal B^0 meson candidates is modelled using a double Crystal Ball function (Eq. 4.4). The combinatorial background events are determined using an exponential function (Eq. 4.6). The partially reconstructed background from $B^0 \to \psi(2S)K^*$ and $B^0 \to \chi_{c1}(1P)K^*$ decays is modelled using double and single Gaussian function, respectively (App. C.5.4). As in the signal decay mode described in Sec. 4.2, the shapes of the mass distributions are fixed to MC fit. The control channel is also divided into three Bremsstrahlung categories to better describe the recovered photons coming from the electron radiation. The parameters describing the shape of the signal candidates mass distribution are fixed to the simulation value as shown in Table 4.11. The fitted signal yield of the $B^0 \to J/\psi(e^+e^-)K^*$ decay mode obtained from a sum of Bremsstrahlung category fits is found to be 55464±543 candidates. The final sWeights computed from the mass fits are used to define the decay time acceptance of the $B_s^0 \to J/\psi\phi$ decay mode.

Calculation of the decay time acceptance

The decay time efficiency for the $B_s^0 \to J/\psi\phi$ signal candidates is computed as the ratio of the sWeighted decay time distribution of the $B^0 \to J/\psi K^*$ control decay with its true decay time distribution, created by generating 10^6 toy events from an exponential distribution

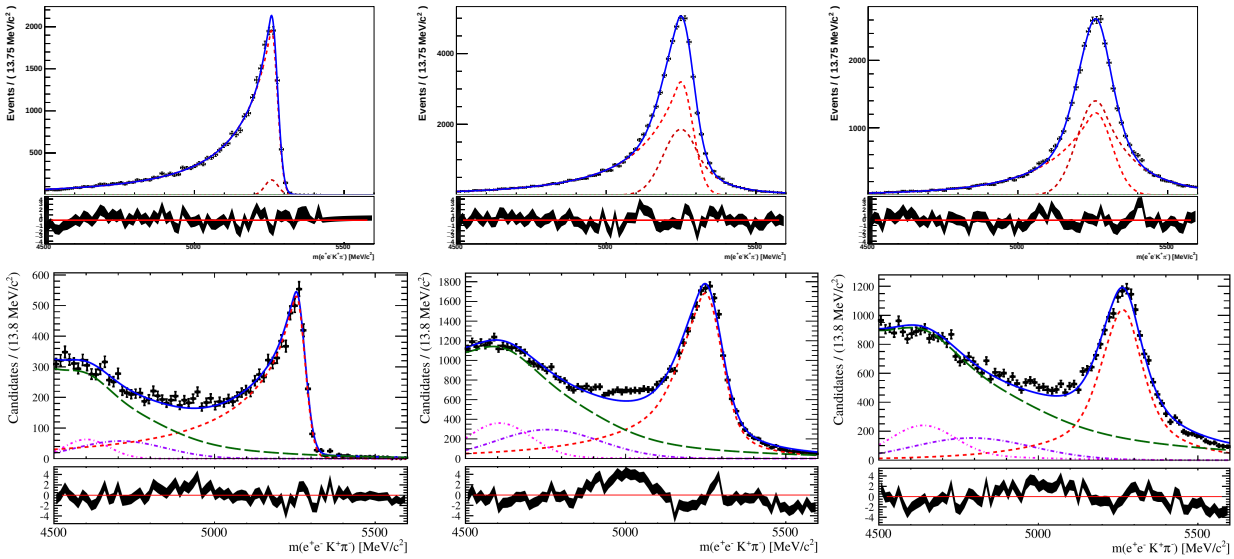


Fig. 4.18 Distribution of the $m(e^+e^-K^+\pi^-)$ in the (top) simulated and (bottom) data sample of $B^0 \to J/\psi K^*$ events divided into three Bremsstrahlung categories: (left) without, (middle) one and (right) more than one photons coming from the electron radiation. The blue line shows the total fit which is included the signal (red line) and combinatorial background (green line) contributions. If there is, a partially reconstructed background from the $B^0 \to \psi(2S)K^*$ and $B^0 \to \chi_{c1}(1P)K^*$ decays is indicated by pink and purple lines, respectively.

Table 4.11 Results of the fit to the $m(e^+e^-K^+\pi^-)$ distribution in data sample divided into three Bremsstrahlung categories. The scale is a difference between sigma of the first and second Crystal Ball function and $n_1 = n_2 = n$. The shapes of the partially reconstructed background are fixed to MC fit reported in Table C.5

Parameter	0γ	1γ	2γ	
α_1	0.19±0.01	0.31±0.01	0.36±0.03	fixed to MC fit
α_2	-1.9±0.2	-0.913±0.03	-0.586±0.02	fixed to MC fit
σ[MeV/c^2]	26.5±0.8	48.4±0.8	57.7±1.3	float
scale	1.12±0.09	1.47±0.06	1.13±0.06	fixed to MC fit
μ[MeV/c^2]	5254.3±1.1	5248.5±0.8	5264.7±1.0	float
$f_{\rm CB1}$	0.97±0.01	0.69±0.01	0.5±0.02	fixed to MC fit
n	2.3±0.1	3.6±0.2	7.1±1.0	fixed to MC fit
$\alpha_{\rm slope}$	-0.00377±0.00011	-0.00294±0.00008	-0.00221±0.00006	float
$N_{\rm cbkg}$	4856±258	19140±815	22172±845	
$f_{B^0\to\chi_{c1}(1P)K^*}$	0.438±0.029	0.457±0.012	0.368±0.014	float
$N_{B^0\to\chi_{c1}(1P)K^*}$	926±113	6424±319	3649±258	
$N_{B^0\to\psi(2S)K^*}$	1189±244	7633±672	6268±645	
$N_{\rm sig}$	8390±143	29376±377	17698±364	

with lifetime fixed to the known B^0 lifetime[12]. The $m(e^+e^-K^+\pi^-)$ distribution is used as a discriminating variable for the sPlot technique [94]. The decay time efficiency of the control decay mode is shown in Fig. 4.19 which is called $\varepsilon^{B^0}_{\rm data}$ in the following. The distribution is divided into 40 bins of variable width.

[12]A world average of the B^0 meson lifetime is $\tau(B^0)$=1.520±0.004 ps [20].

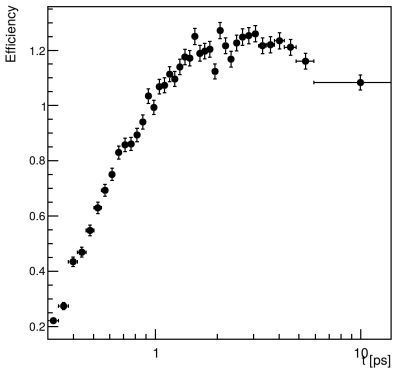


Fig. 4.19 Relative decay time efficiency determined from the sWeighted B^0 decay time distribution and generator level events. The x-axis is in a logarithmic scale.

To account for any differences in the decay time acceptance shape between the signal $B_s^0 \to J/\psi\phi$ decay and control $B^0 \to J/\psi K^*$ decay, the $\varepsilon_{\text{data}}^{B_s^0}$ is corrected by a factor r:

$$\varepsilon_{\text{data}}^{B_s^0} = \varepsilon_{\text{data}}^{B^0} \times r = \varepsilon_{\text{data}}^{B^0} \times \frac{\varepsilon_{\text{MC}}^{B_s^0}}{\varepsilon_{\text{MC}}^{B^0}}, \tag{4.7}$$

where $\varepsilon_{\text{MC}}^{X}$ is the efficiency for mode X determined from the simulated sample. For $\varepsilon_{\text{MC}}^{B_s^0}$, the ratio of the decay time distribution in truth matched events is taken with the decay time distribution of generator level events. In case of $\varepsilon_{\text{MC}}^{B^0}$, the same process as described for signal channel is implemented using 10^6 toy events. Fig. 4.20 shows the resulting form of the factor r. The bias for low decay time is introduced by a small difference in the stripping selection of the $B_s^0 \to J/\psi\phi$ and $B^0 \to J/\psi K^*$ decays (Tables 4.3, 4.9). The size of the B^0 control sample dominates the statistical uncertainty on the final efficiency. The final decay time efficiency $\varepsilon_{\text{data}}^{B_s^0}$ is presented in Fig. 4.21 which is used in the unbinned maximum likelihood fit to determine the $\mathcal{CP}$-violating parameters in the analyzed decay.

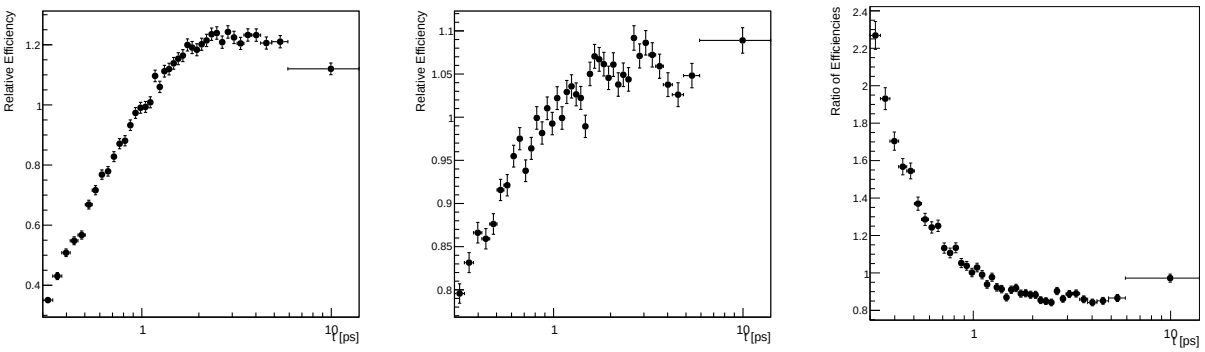


Fig. 4.20 Decay time acceptance from simulated (left) $B^0 \to J/\psi K^*$ and (middle) $B_s^0 \to J/\psi\phi$ events. (right) Ratio r of the acceptances of the simulated B_s^0 decay time over the B^0 decay time. The x-axis is in a logarithmic scale.

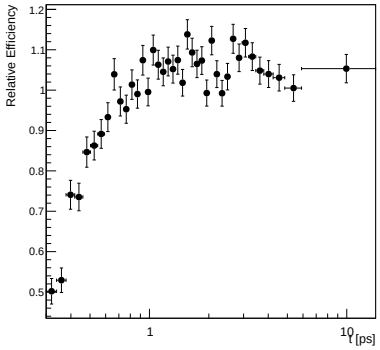


Fig. 4.21 Final decay time efficiency for the signal $B_s^0 \to J/\psi\phi$ events defined in Eq. 4.7. The x-axis is in a logarithmic scale.

4.3.2 Decay time resolution

The precision of the resulting measurement of the phase ϕ_s is dependent on the decay time resolution. It is crucial to have a good understanding of this quantity behaviour for use in the maximum likelihood fit.

To optimally use the information extracted from the $B_s^0 \to J/\psi\phi$ signal sample, a resolution model dependent on the per-event estimated decay time uncertainty, σ_t is used. The decay time resolution is modelled as a sum of two Gaussian distributions with a common mean t_0 and two widths $\delta_i(\sigma_t)$ (one narrow and one wide) each of which depend upon σ_t. The model is:

$$\mathcal{R}(t|\sigma_t) = \sum_{i=1}^{2} f_i \frac{1}{\sqrt{2\pi}\delta_i(\sigma_t)} e^{-\frac{1}{2}\left(\frac{t-t_0}{\delta_i(\sigma_t)}\right)^2}, \tag{4.8}$$

where f_i is the fraction of the i^{th} Gaussian function. Since there are only 2 Gaussian functions single parameter $f = 1 - f_1$ and $f = f_2$ is defined. In order to help with the convergence of the fit to the decay time distribution, a parameter transformation is used to reduce the correlations between fit parameters [39],

$$\begin{aligned} \delta_1 &= -\sqrt{\frac{f}{1-f}}\delta''(\sigma_t) + \delta'(\sigma_t), \\ \delta_2 &= \sqrt{\frac{1-f}{f}}\delta''(\sigma_t) + \delta'(\sigma_t). \end{aligned} \tag{4.9}$$

A quadratic relation is expressed in the equivalent form

$$\begin{aligned} \delta'(\sigma_t) &= \sigma_t(c' + b'(\sigma_t - \bar{\sigma}_t)), \\ \delta''(\sigma_t) &= \sigma_t(c'' + b''(\sigma_t - \bar{\sigma}_t)), \end{aligned} \tag{4.10}$$

where the parameter $\bar{\sigma}_t$ is the mean estimated decay time uncertainty for the data sample and has been introduced to further reduce correlations between c and b parameters. Their values for each year of data taking are given in Table 4.12.

Table 4.12 Average estimated decay time resolution.

year	$\bar{\sigma}_t$ [ps]
2011	0.0307
2012	0.0298

Using this parametrization, a decay time fit can be performed to a data sample of prompt $J/\psi \to e^+e^-$ events selected using the following decay time unbiased stripping and trigger lines:

- The stripping selection `StrippingBs2JpsieePhiLine` which is identical to the selection `StrippingBs2JpsieePhiDetachedLine` except for the B_s^0 decay time selection, $t(B_s^0)$ >0.3 ps (Sec. 4.1);
- `L0Electron` or `L0Hadron` trigger lines is considered (Sec. 4.1.1).

Such procedure was successfully used to determine the decay time resolution for the $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ decay analysis [39]. However, no $J/\psi \to e^+e^-$ mass peak is observed in the data (Fig. 4.22), even after testing various stricter selection criteria to try and isolate the signal candidates. The J/ψ mass peak arises using a constraint on the B_s^0 decay time $t > 0.3$ ps.

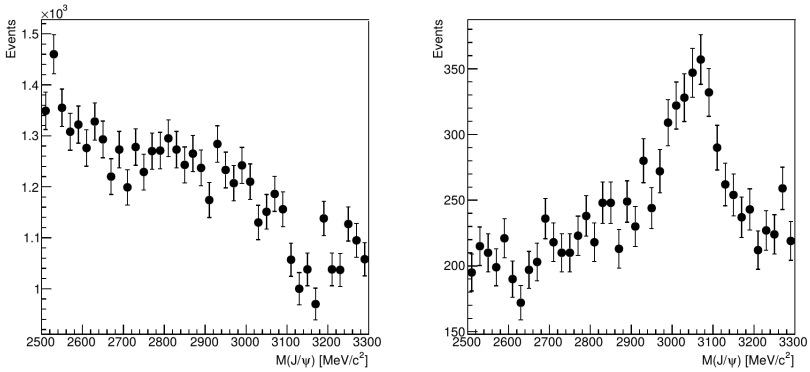


Fig. 4.22 Distribution of the $m(e^+e^-)$ for events selected using the `StrippingBs2JpsieePhiLine` and unbiased `L0` triggers (left) before and (right) after applying $t(B_s^0)$ >0.3 ps.

As an alternative, the simulated $B_s^0 \to J/\psi(e^+e^-)\phi$ and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ events are used to search for significant difference in the resolution between the two modes. The PDF that is used to fit the decay time distribution of the simulated sample is given by

$$\mathrm{PDF}(t|\sigma_t) = \mathcal{R}(t|\sigma_t) + N_{\mathrm{wpv}}\mathcal{G}(t), \tag{4.11}$$

where the first term is the decay time resolution model (Eq. 4.8) and the second term is a component associated with the wrong PV. These candidates have an incorrect decay time and a wide Gaussian function has been added to the PDF to take this component into account. A

fit is performed to the difference between the reconstructed and true decay time distributions for each year. The projections of the fit to both simulated decay modes and the values of their parameters are shown in Table 4.13 and Figs. 4.23, 4.24. In case of 2012 simulated sample, the decay time resolution model corresponds to an effective dilution of 0.782±0.0002 for the $B_s^0 \to J/\psi(\mu\mu)\phi$ decay and 0.742±0.001 for the $B_s^0 \to J/\psi(ee)\phi$ events calculated using Eq. 4.12. The effective resolution is 39.5±0.001 fs and 43.6±0.008 fs, respectively. A similar value of the effective resolution and dilution is obtained for the 2011 simulated sample.

The corresponding dilutions are defined as

$$D = \sum_{i=1}^{2} f_i e^{\left(-\frac{\Delta m_s^2 \sigma_{i,true}^2}{2}\right)}, \qquad D_{\text{effective}} = \sqrt{\frac{1}{S_W}\sum_i w_i D_i^2}, \tag{4.12}$$

where w_i is a weight of the event, S_W is a sum of the weights and Δm_s^2 is the B_s^0 oscillation frequency taken as 17.711 ps^{-1} [3].

Table 4.13 Resolution model parameters obtained from a fit to the difference between the reconstructed and true decay time using 2011 and 2012 simulated $B_s^0 \to J/\psi(\mu\mu)\phi$ and $B_s^0 \to J/\psi(ee)\phi$ events.

Parameters	2011		2012	
	$B_s^0 \to J/\psi(\mu\mu)\phi$	$B_s^0 \to J/\psi(ee)\phi$	$B_s^0 \to J/\psi(\mu\mu)\phi$	$B_s^0 \to J/\psi(ee)\phi$
$b^{'}$	1.993±0.095	-7.689±0.289	1.492±0.099	-7.314±0.229
$b^{''}$	4.304±0.197	-6.899±0.524	4.200±0.215	-6.478±0.436
$c^{'}$	1.2668±0.0011	1.4414±0.0059	1.3054±0.0011	1.4763±0.0048
$c^{''}$	0.3201±0.0026	0.6531±0.0133	0.3227±0.0028	0.6423±0.01
f	0.0773±0.0027	0.1316±0.0047	0.0724±0.0026	0.156±0.0046
μ_G [ps]	-0.0005±0.00003	-0.0023±0.00012	-0.0004±0.00004	-0.002±0.0001
Calculate on $B_s^0 \to J/\psi(ee)\phi$ data				
eff res [fs]	40.7±0.004	42.8±0.012	39.5±0.001	43.6±0.008
eff dilution	0.77±0.001	0.75±0.002	0.782±0.0002	0.742±0.001

Since the kinematics of the electron and muon modes of the $B_s^0 \to J/\psi\phi$ decay are similar (Sec. C.8), it is expected that the description of the real resolution by simulated sample is comparable for the two decay modes. Thus, the values that parametrize the decay time resolution model in the $B_s^0 \to J/\psi(e^+e^-)\phi$ data sample can be determined from the muon sample. To take into account differences between the two decay modes the parameters are free in the final unbinned maximum likelihood fit with a Gaussian constraint as listed in Table 4.14. The constraint value is taken as the fit parameters value of the decay time resolution obtained on the $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ data sample [39]. The width of the constraint is defined as a double difference between the c, b and f parameter values from the fit to the $B_s^0 \to J/\psi(e^+e^-)\phi$ and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ simulated samples as reported in Table 4.13.

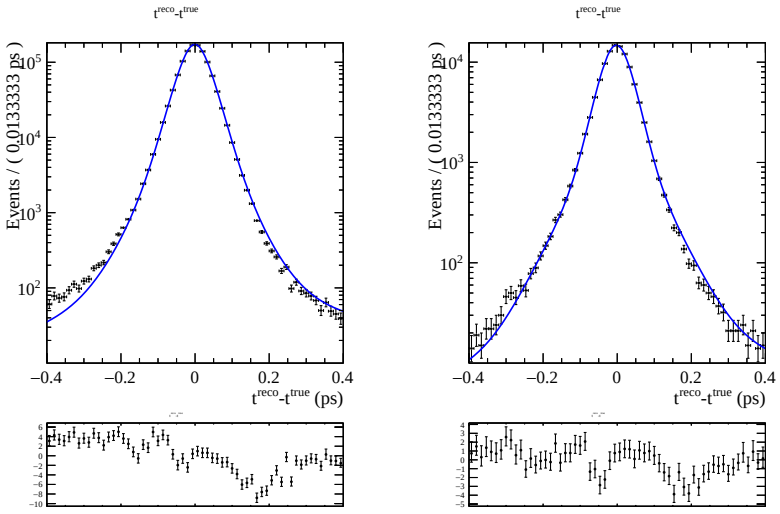


Fig. 4.23 Distribution of the $t = t^{\text{reco}} - t^{\text{true}}$ of 2011 simulated (left) $B_s^0 \to J/\psi(\mu\mu)\phi$ and (right) $B_s^0 \to J/\psi(ee)\phi$ samples. The solid blue line shows the fit result of Eq. 4.11. The black points with error bars represent data.

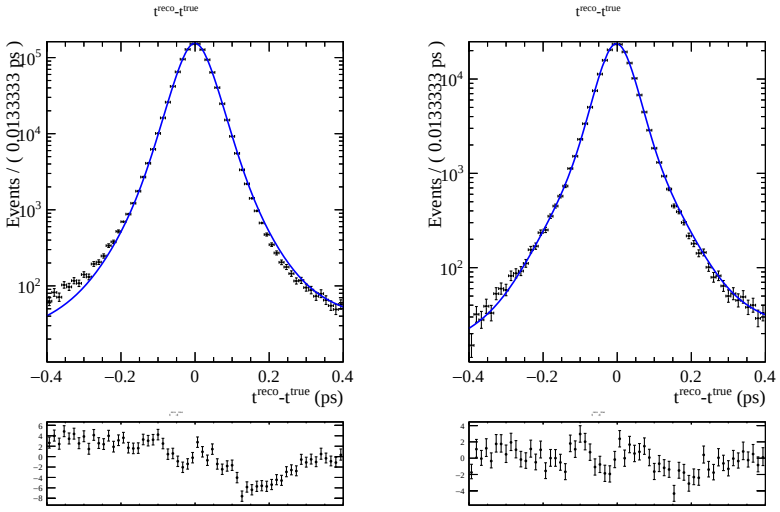


Fig. 4.24 Distribution of the $t = t^{\text{reco}} - t^{\text{true}}$ of 2012 simulated (left) $B_s^0 \to J/\psi(\mu\mu)\phi$ and (right) $B_s^0 \to J/\psi(ee)\phi$ samples. The solid blue line shows the fit result of Eq. 4.11. The black points with error bars represent data.

The resolution model parameters extracted from an unbinned maximum likelihood fit to 2011 and 2012 data sets of the $B_s^0 \to J/\psi(e^+e^-)\phi$ decay are reported in Table 4.15. As described in Sec. 4.6.1 a constraint on $\Delta\Gamma_s$ and Γ_s to the final fit is applied: $\Delta\Gamma_s$=0.0805±0.0091 ps^{-1} and Γ_s=0.6603±0.0027 ps^{-1} [39]. The corresponding dilutions of the 2011 and 2012 data samples are 0.743±0.077 and 0.737±0.055, respectively. The effective overall average resolutions are 43.5±0.423 fs and 44.1±0.304 fs, respectively.

Table 4.14 Constraint on the free fit parameters of the decay time resolution model for 2011 and 2012 data samples.

Parameters	2011	2012
$b^{'}$	-2.85±19.46	-4.96±17.824
$b^{''}$	-1.55±22.407	-2.88±21.832
$c^{'}$	1.4207±0.381	1.4811±0.3636
$c^{''}$	0.3778±0.6968	0.4057±0.6598
f	0.244±0.1073	0.239±0.1666

Table 4.15 Resolution model parameters obtained from an unbinned maximum likelihood fit to 2011 and 2012 background subtracted data of the $B_s^0 \to J/\psi(e^+e^-)\phi$ decay.

Parameters	2011	2012
$b^{'}$	-2.678±16.181	-3.469±8.012
$b^{''}$	0.652±15.124	-3.289±9.888
$c^{'}$	1.4464±0.2007	1.4754±0.178
$c^{''}$	0.3886±0.3638	0.3999±0.3236
f	0.2445±0.0533	0.2389±0.0827
eff res [fs]	43.5±0.428	44.1±0.304
eff dilution	0.743±0.077	0.737±0.055

4.4 Angular acceptance and resolution

The final state of the $B_s^0 \to J/\psi\phi$ decay is an admixture of $\mathcal{CP}$-even and $\mathcal{CP}$-odd eigenstates, which need to be separated in order to extract the $\mathcal{CP}$-violating parameters. This separation can be performed on a statistical basis using the distinct angular distributions of the different $\mathcal{CP}$ eigenstates as shown in Sec. 1.3.1. Experimentally these angular distributions are modified by imperfect detection and reconstruction of the decay chain. A particularly important effect is a non-uniform angular acceptance, i.e. the event reconstruction efficiency varies as a function of the angular variables. A good understanding and reliable correction of the angular acceptance and resolution is not only necessary for controlling the systematics of the experimental measurements, but it is also vital for checking the validity of the model used to describe the data.

4.4.1 Angular acceptance

The reconstruction and selection processes (Sec. 4.1) introduce a distortion on the three angles ($\Omega = \{\cos\theta_K, \cos\theta_e, \phi_h\}$) defined in Fig. 1.7. In the PDF that describes the decay as a function of decay time and angles, the angular acceptance effects are taken into account by multiplying the expression for the differential decay rate (Eqs. 1.35, 1.36) by $\varepsilon_a(\Omega)$. In case only the signal $B_s^0 \to J/\psi\phi$ dataset is described, all terms in the PDF are proportional to $\varepsilon_a(\Omega)$. If the acceptance function contains no free parameters it appears as a constant term in the log-likelihood function, which can be ignored.

However, the angular acceptance must be included in the normalization calculated for the log-likelihood in the form of normalization weights, $\xi_i = \int \mathrm{d}\Omega \varepsilon_a(\Omega) f_i(\Omega)$ [104]. Note that for a flat acceptance equal to one ($\varepsilon_a(\Omega)$=1), the weights are given by the integrals of the angular functions $f_i(\Omega)$. In that case, they would vanish for the terms that describe the interference between final state polarizations: $\xi_{4,5,6,8,9,10}$=0. The procedure of the normalization weights determination is briefly summarized in App. C.6, while a more detailed description can be found in Ref. [104].

The angular acceptance normalization weights are calculated using 2011 and 2012 simulated data samples as shown in Fig. 4.25. The 10 weights corresponding to the 10 terms of the PDF (Eq. 1.32) are shown in Table 4.16, split by running year. No reweighting of the final state kinematics has been applied to determine these numbers, since the momentum distributions in the simulation are observed to well reproduce those in data as shown in Fig. C.11. As no significant difference is observed between the two years (Fig. 4.25), the weights are determined and normalized separately for each running year and then combined to the final set of the weights as reported in Table 4.17. The obtained normalization weights for simulated signal events are applied in the unbinned maximum likelihood fit (Sec. 4.6). The correlations between the errors of the weights are shown in Table C.6.

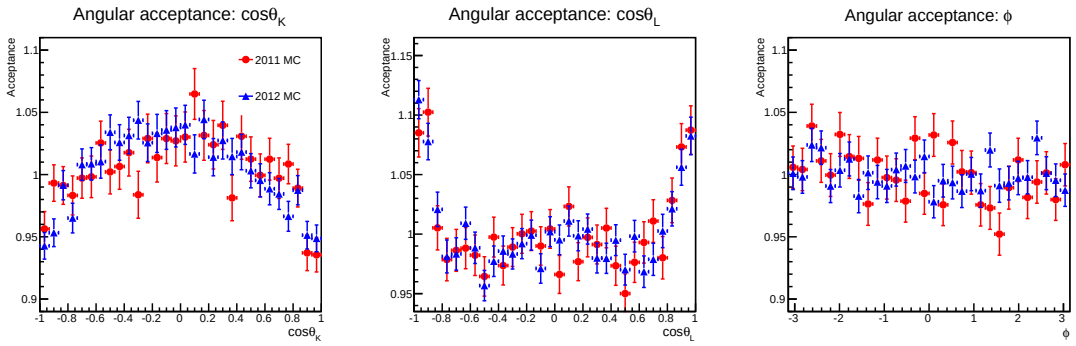


Fig. 4.25 Angular acceptance projections calculated from the 2011 and 2012 simulated sample for the three helicity angles.

4.4.2 Angular resolution

The definition of the angular observables $(\cos\theta_K, \cos\theta_e, \phi_h)$ in the rest frame of the mesons are shown in Fig. 1.7. The differences between the true and reconstructed values of the observables are caused by imperfections in the tracking[13]. Since the polarization amplitudes A_i are determined from the angular distribution (Sec. 1.3.1), non-zero angular resolutions lead to a decrease of the sensitivity to these physics parameters.

The one-dimensional angular resolutions (angle$^{\mathrm{true}}$ - angle$^{\mathrm{reco}}$) determined from the signal $B_s^0 \to J/\psi\phi$ simulated events are shown in Fig. 4.26. For angles θ_K and θ_e, the fit model is

[13] Due to boosts and rotations, it is not intuitive to translate tracking resolutions in the lab frame to angular resolutions in the rest frame of the decay particles.

Table 4.16 Angular acceptance weights for the 2011 and 2012 simulated samples. The last column shows the difference in σ between two samples.

		2011	2012	
	k	ξ_k/ξ_1	ξ_k/ξ_1	Diff(σ)
1	(00)	0.9817±0.0022	0.9791±0.0017	+0.9
2	($\parallel\parallel$)	1.0181±0.0027	1.0212±0.0021	-0.9
3	($\perp\perp$)	1.0185±0.0026	1.0224±0.0020	-1.2
4	($\parallel\perp$)	0.0005±0.0029	0.0001±0.0023	+0.1
5	(0$\parallel$)	0.0001±0.0020	0.0013±0.0015	-0.5
6	(0$\perp$)	0.0014±0.0019	0.0016±0.0015	-0.1
7	(SS)	0.9957±0.0017	0.9998±0.0013	-1.9
8	(S$\parallel$)	-0.0017±0.0026	0.0017±0.0020	-1.0
9	(S$\perp$)	-0.0006±0.0026	0.0022±0.0020	-0.9
10	(S0)	-0.0058±0.0058	-0.0073±0.0045	+0.2

Table 4.17 Angular acceptance weights for full simulated sample.

	k	ξ_k/ξ_1
1	(00)	0.9801±0.0014
2	($\parallel\parallel$)	1.0200±0.0017
3	($\perp\perp$)	1.0209±0.0016
4	($\parallel\perp$)	0.0003±0.0018
5	(0$\parallel$)	0.0008±0.0012
6	(0$\perp$)	0.0015±0.0012
7	(SS)	0.9983±0.0011
8	(S$\parallel$)	0.0004±0.0016
9	(S$\perp$)	0.0012±0.0016
10	(S0)	-0.0067±0.0036

composed of a sum of three Gaussian functions while for ϕ angle a sum of double Gaussian and Voigtian [105] component that is required in order to better model the large non-Gaussian tails. The parameters determined from these fits are reported in Table 4.18. Based upon studies performed in Ref. [28], the effect from the angular resolution is negligible. It is considered as a source of systematic uncertainty detailed in Sec. 4.7.10.

Table 4.18 Fit results to the angular resolution distributions for each of the helicity angles taken from the $B_s^0 \to J/\psi\phi$ simulated sample. The last row shows the effective resolution computed for θ_K and θ_L angles.

	3 Gaussian	3 Gaussian	2 Gaussian+Voigtian
Parameter	θ_L	θ_K	ϕ
σ_1[mrad]	20.60±0.33	12.45±0.13	17.42±0.26
σ_2[mrad]	7.190±0.14	20.71±0.19	36.52±0.91
σ_3[mrad]	61.42±0.72	43.70±1.20	38.90±1.80
σ_4[mrad]	-	-	84.60±2.90
f_1	0.4151±0.0051	0.3740±0.0130	0.5140±0.0180
f_2	0.1968±0.0063	0.5750±0.0110	0.3920±0.0140
f_3	0.3881±0.0081	0.0510±0.0170	0.0940±0.0228
σ_{eff}[mrad]	40.63±0.57	20.05±0.84	-

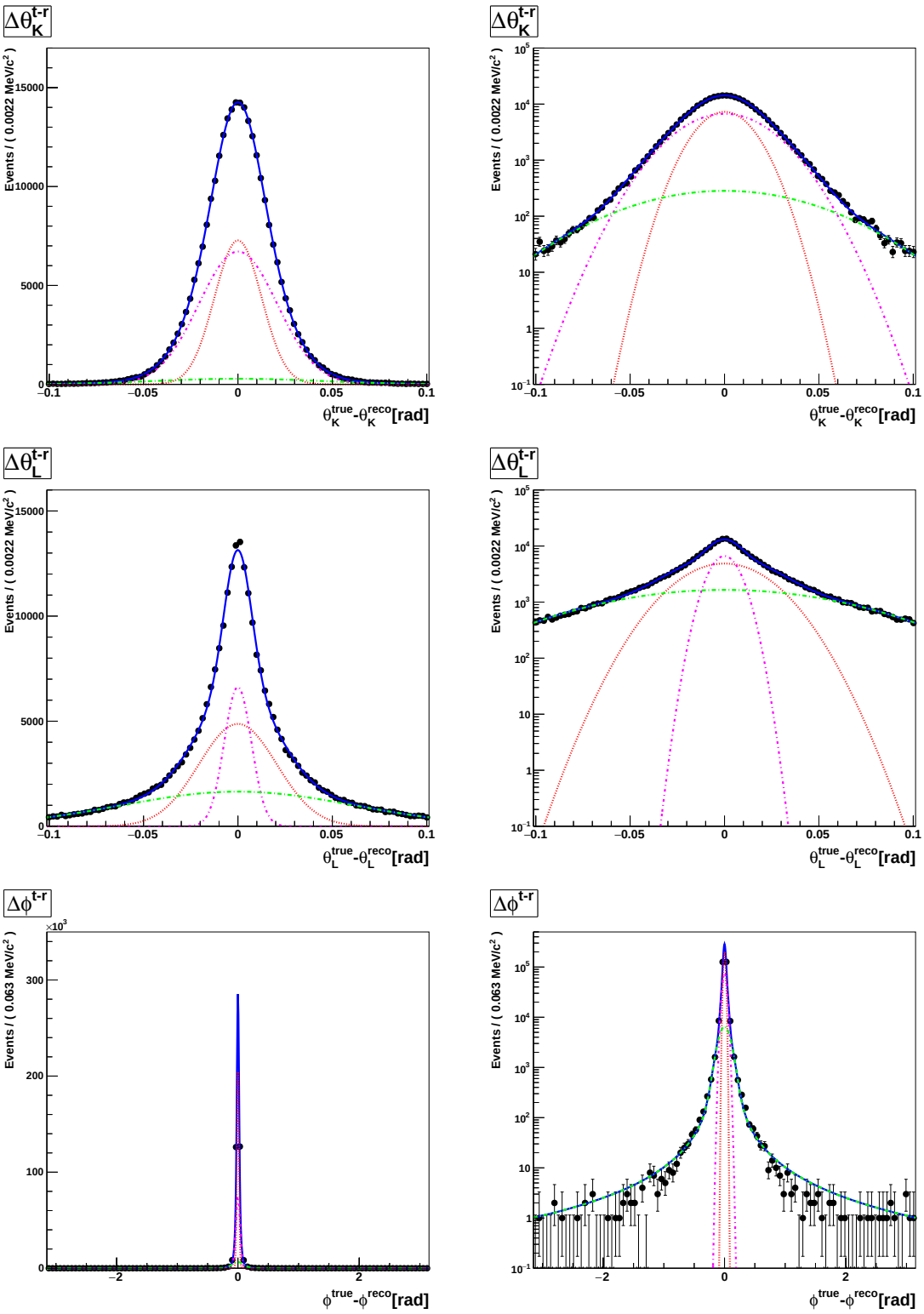


Fig. 4.26 Distributions of the difference between true and reconstructed angle from the $B_s^0 \to J/\psi\phi$ simulated sample. The right plots are in a logarithmic scale. The units are radians in all cases.

4.5 Flavour tagging

The measurement of the $\mathcal{CP}$ asymmetry in the interference of mixing and decay requires the knowledge of the production flavour of the B mesons in order to distinguish the differential decay rates for B_s^0 and $\bar{B}_s^0$ mesons as defined in Sec. 1.3.2. The procedure to determine the production flavour is called flavour tagging. Several algorithms are used in the LHCb experiment to extract the production flavour from the B meson decay kinematics and event properties. The tagging algorithms provide for each event a tagging decision q with $q = +1$ for a produced B_s^0, q = -1 for a produced $\bar{B}_s^0$ and $q = 0$ if no decision could be made. The tagging efficiency is defined as

$$\varepsilon_{tag} = \frac{N_{\mathrm{R}} + N_{\mathrm{W}}}{N_{\mathrm{R}} + N_{\mathrm{W}} + N_{\mathrm{U}}}, \tag{4.13}$$

where N_{R} is the number of events with a correct tagging decision, N_{W} is the number of incorrectly tagged events and N_{U} is the number of untagged events. The quality of the tagging is described by the mistag probability ω giving the fraction of wrong tagging decisions:

$$\omega = \frac{N_{\mathrm{W}}}{N_{\mathrm{R}} + N_{\mathrm{W}}}. \tag{4.14}$$

The influence of the tagging quality on the measurement of the phase ϕ_s can be seen by looking at the differential decay rates (Eqs. 1.32, 1.33). The terms providing the most information on ϕ_s are those proportional to $\sin\phi_s \cdot \sin(\Delta m_s t)$. These terms appear with different signs in the decay rates of B_s^0 and $\bar{B}_s^0$ meson. A finite mistag probability for the determination of the production flavour leads to an additional damping factor in the $\mathcal{CP}$ asymmetry (Eq. 1.29):

$$\mathcal{A}_{\mathcal{CP}}(t) \propto \mathcal{D} \sin\phi_s \cdot \sin(\Delta m_s t), \tag{4.15}$$

where $\mathcal{D} = (1 - 2\omega)$ is the tagging dilution [106]. The sensitivity on the phase ϕ_s as the measurement of the amplitude of the $\mathcal{CP}$ asymmetry depends therefore directly on the mistag probability of the tagging. The effective statistical reduction of the data sample due to imperfect tagging with respect to an event sample with perfect tagging is given by the effective tagging power ε_{eff},

$$\varepsilon_{eff} = \varepsilon_{tag}\mathcal{D}^2 = \varepsilon_{tag}(1 - 2\omega)^2. \tag{4.16}$$

In order to determine the production flavour of neutral B mesons, two classes of algorithms are available at the LHCb [107]:

- Opposite-side tagging algorithms (OS) exploit the fact that b-hadrons are always produced in pairs with opposite flavour. They use the properties of the second b-hadron in the event to determine the production flavour of the signal B_s^0 meson candidate;
- Same-side tagging algorithms (SS) exploit the properties of the signal B_s^0 meson decay

to get an estimation of the production flavour.

Fig. 4.27 shows a schematic overview of the different tagging algorithms. Both opposite-side and same-side tagging algorithms are used in the analysis.

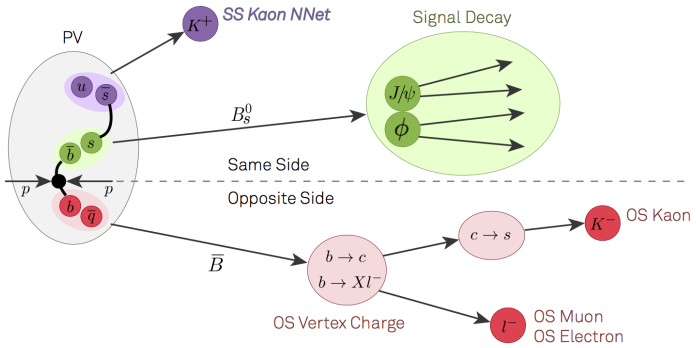


Fig. 4.27 Schematic overview of tagging algorithms at LHCb. Two main types are visible: opposite-side and same-side taggers. The figure is taken from [108].

4.5.1 Opposite-side tagging

The opposite-side taggers use the information of the second b-hadron decay in the event to determine the flavour of the signal B_s^0 meson candidate [109]. There are four algorithms available:

- *Single particle taggers.* The single particle taggers use the charge of leptons from semileptonic b-hadron decays (electron and muon tagger) or the charge of the kaon from the $b \to c \to s$ decay chain, for example, in the $B^0 \to D^{*-}\mu^+\nu_\mu$ decay, to determine the production flavour of the signal B meson. They require the muon, electron and kaon track candidates to have a large impact parameter significance, $\mathrm{IP}/\sigma_{\mathrm{IP}}$, and large transverse momentum p_{T} in order to select only candidates from long-living b-hadrons. In addition, the information from the particle identification detectors is used by applying $\Delta\ln\mathcal{L}$ requirements (Sec. 2.3.2, 2.3.4). In case when a single particle tagger selects more than one track candidate per event, the one with the highest transverse momentum is chosen to define the tagging decision.

- *Vertex charge tagger.* The vertex charge tagger determines the overall charge associated to the decay vertex of the second b-hadron in the event. The vertex candidates are reconstructed by combining two tracks with high impact parameter significance and transverse momentum to a common vertex. For each vertex candidate, the probability to originate from a b-hadron decay is estimated, taking into account the quality of the vertex fit as well as geometric and kinematic properties of the vertex candidate. The vertex with the highest probability is chosen as decay vertex of the second b-hadron. Afterwards, all tracks in the event that are compatible with originating from this

vertex candidate but not from PV are assigned to the tagging vertex. After applying additional selection requirements on the sum of the associated tracks, the overall vertex charge is calculated by summing the charges of the tracks weighted by their transverse momentum. The tagging decision is made depending on the value of the overall vertex charge [109].

4.5.2 Same-side tagging

In the hadronisation of a b quark to a B_s^0 meson, the $s\bar{s}$ quark pair is produced from the vacuum. The s quark forms together with the $\bar{b}$ quark the signal meson and the $\bar{s}$ quark might hadronise to a charged kaon (Fig. 4.27). In this case, the charge of the kaon indirectly determines the flavour of the B_s^0 meson. To select the kaon tracks from the hadronisation, the same-side kaon tagging algorithm is used. It considers tracks that are close to the signal B_s^0 phase space by requiring a maximum polar angle difference between a track and the B_s^0 meson candidate. A maximum $\mathrm{IP}/\sigma_{\mathrm{IP}}$ with respect to the PV is required, to ensure that the kaon candidate originates from the B_s^0 hadronisation. To reject low momentum background tracks from the PV, a high p_{T} is required for the kaon candidate. The $\Delta\ln\mathcal{L}$ variable is used to distinguish kaons from pions and protons. In case more than one kaon track candidate is selected per event, the one with the highest transverse momentum is chosen to determine the tagging decision [110].

4.5.3 Calibration of taggers

Besides the tagging decision q, each of the single tagging algorithms provides an estimate for the mistag probability, η, for every event. The estimated η is determined by a neural network depending on the B_s^0 kinematic and event properties. The measurement of the phase ϕ_s profits from the additional information provided by the estimated mistag probability because the B_s^0 meson candidates with a more reliable tagging decision get a larger weight in the $\mathcal{CP}$ asymmetry. In order to use the estimated mistag probability, it has to be calibrated to make sure it reflects the true mistag probability.

The performance of the tagging algorithms might in addition depend on the B_s^0 production flavour, resulting in different true mistag probabilities ω^{alg} and $\bar{\omega}^{\mathrm{alg}}$ (alg = OS, SS) for produced B_s^0 and $\bar{B}_s^0$ mesons, respectively. Ignoring this difference would introduce an artificial asymmetry and lead to a wrong measurement of the $\mathcal{CP}$ violation. Therefore, the estimated mistag probability η is calibrated separately for produced B_s^0 and $\bar{B}_s^0$ mesons. For OS tagger, the calibration is done using the $B^{\pm} \to J/\psi(ee)K^{\pm}$ decays. Since the charged $B^{\pm}$ meson does not oscillate, their production flavour is determined by a charge of the kaon in the decay. In the calibration procedure, the estimated mistag probability η is compared to the true mistag probabilities ω^{alg} and $\bar{\omega}^{\mathrm{alg}}$ measured with $B^+ \to J/\psi(ee)K^+$ and $B^- \to J/\psi(ee)K^-$ decays, respectively [109].

The calibration of SS tagger can only be done with B_s^0 meson decays, as a kaon needs to be produced in the hadronisation process of the B_s^0 meson (Sec. 4.5.2). The $B_s^0 \to D_s^- \pi^+$ decay is used for the calibration. As the B_s^0 meson oscillates, a time dependent analysis is necessary to determine the production flavour [110] and to measure the true mistag probabilities ω^{alg} and $\bar{\omega}^{\text{alg}}$. The details of the SS flavour tagging calibration are given in App. C.9.

In the calibration procedure a linear dependence between estimated and true mistag probabilities is assumed to derive a calibrated mistag probability η^{alg}. In order to correctly propagate the statistical and systematic uncertainties of the calibration, the true mistag probabilities ω^{alg} and $\bar{\omega}^{\text{alg}}$ are expressed as a function of the calibrated mistag probability η^{alg},

$$\begin{aligned}
\omega^{\text{alg}} &= \left(p_0^{\text{alg}} + \frac{\Delta p_0^{\text{alg}}}{2}\right) + \left(p_1^{\text{alg}} + \frac{\Delta p_1^{\text{alg}}}{2}\right)(\eta^{\text{alg}} - \langle\eta^{\text{alg}}\rangle), \\
\bar{\omega}^{\text{alg}} &= \left(p_0^{\text{alg}} - \frac{\Delta p_0^{\text{alg}}}{2}\right) + \left(p_1^{\text{alg}} - \frac{\Delta p_1^{\text{alg}}}{2}\right)(\eta^{\text{alg}} - \langle\eta^{\text{alg}}\rangle),
\end{aligned} \tag{4.17}$$

where $\langle\eta^{\text{alg}}\rangle$ is the average of the calibrated mistag probabilities for all events and p_i^{alg} (i=0, 1) are the calibration parameters; Δp_i^{alg}, called mistag asymmetries, are the difference between the calibration parameters as measured for B_s^0 and $\bar{B}_s^0$ mesons. For a perfectly calibrated tagger, the calibration parameters are $p_0 = \langle\eta^{\text{alg}}\rangle$ and $p_1 = 1$ and the true mistag probabilities are identical to the calibrated mistag probability. For $\Delta p_i^{\text{alg}} = 0$, the true mistag probability does not depend on the production flavour.

After the calibration of the individual OS taggers, the tagging decisions and calibrated mistag probabilities are combined to one single OS decision q^{OS} and mistag probability η^{OS}. To account for correlations between the individual OS taggers, the calibration with $B^\pm \to J/\psi(ee)K^\pm$ decays has to be repeated, resulting in a single calibrated mistag probability η^{OS} for the OS taggers.

4.5.4 Tagging strategy for fit

After the calibration and combination of the single OS taggers each B_s^0 candidate has one tagging decision and calibrated mistag probability from the OS taggers q^{OS}, η^{OS}, and one from the SS tagger, q^{SS}, η^{SS}. The calibrated mistag probabilities are related to the true mistag probabilities ω^{OS} and ω^{SS} by the parameterization given in Eq. 4.17.

For the measurement of the phase ϕ_s, the $B_s^0 \to J/\psi\phi$ candidates are classified in four tagging categories, depending on the opposite and same side tagging decisions. A candidate is called "OS-only" tagged when $q^{\text{SS}} = 0$ and $q^{\text{OS}} \neq 0$ and "SS-only" tagged when $q^{\text{OS}} = 0$ and $q^{\text{SS}} \neq 0$. The candidates that have non-zero tagging decision from both kind of taggers (overlapping decisions) are called "OS&SS" tagged with combined tagging decision $q^{\text{OS\&SS}}$ and calibrated mistag probability $\eta^{\text{OS\&SS}}$. The last category is the untagged events $q^{\text{OS}} = q^{\text{SS}} = 0$ which are about 22% and 26% of the data and simulation samples, respectively.

From the decisions and mistag probabilities of OS and SS, a combined probability $\mathcal{P}(b)$ that the B_s^0 meson contains a b quark is calculated. The probability that the B_s^0 meson contains a $\bar{b}$ quark is given by $\mathcal{P}(\bar{b}) = 1 - \mathcal{P}(b)$. Depending on the value of $\mathcal{P}(b)$, $q^{\text{OS\&SS}}$ and $\eta^{\text{OS\&SS}}$ are determined as [109]:

$$q^{\text{OS\&SS}} = -1,\ \eta^{\text{OS\&SS}} = 1 - \mathcal{P}(b) \text{ if } \mathcal{P}(b) > \mathcal{P}(\bar{b}),$$

$$q^{\text{OS\&SS}} = +1,\ \eta^{\text{OS\&SS}} = 1 - \mathcal{P}(\bar{b}) \text{ if } \mathcal{P}(b) < \mathcal{P}(\bar{b}). \tag{4.18}$$

The calibration parameters, p_i and Δp_i $(i = 0, 1)$, relating the calibrated mistag probability η to the true mistag probabilities ω and $\bar{\omega}$ (Eq. 4.17) are given in Table 4.19 for the three different tagging categories. The stated uncertainties of the calibration parameters include both statistical and systematic uncertainties from the calibration procedure.

Table 4.19 Calibration parameters for two tagging categories "OS-only" and "SS-only" relating the calibrated mistag probabilities to the true mistag probabilities for each B_s^0 candidate [111, 112].

Parameter	OS-only	SS-only
$\langle\eta\rangle$	0.3768	0.4349
$p_0 - \langle\eta\rangle$	0.0086±0.0021	0.0018±0.007
$\Delta p_0/2$	0.0083±0.0021	-0.0047±0.0033
p_1	0.931±0.021	0.886±0.1036
$\Delta p_1/2$	0.0115±0.0205	0.0744±0.0518

The quality of the tagging for the selected $B_s^0 \to J/\psi\phi$ event sample is characterized by the tagging efficiency ε_{tag}, the average mistag probabilities ω and tagging power ε_{eff} of the three tagging categories. The resulting average mistag ω and dilution $\mathcal{D}$ of the $\mathcal{CP}$ asymmetry can be calculated following Eq. 4.14 as

$$\omega = \frac{1}{N}\sum_N \omega^{\text{alg}}, \qquad \mathcal{D} = \frac{1}{N}\sum_N \mathcal{D}^{\text{alg}} = \frac{1}{N}\sum_N (1 - 2\omega^{\text{alg}}), \tag{4.19}$$

where the sums are calculated over N selected signal $B_s^0 \to J/\psi\phi$ candidates. The distributions of the averaged mistag ω^{alg} in two tagging categories are shown in Fig. 4.28 for signal candidates. They are unfolded from the background events with the sPlot technique [94] by using $m(e^+e^-K^+K^-)$ distribution as a discriminating variable. In the "SS-only" tagged category, the calibrated signal candidates have in average a higher tagging mistag compared with the "OS-only" category.

Finally, the tagging efficiency, the average dilution and the resulting tagging power

$$\varepsilon_{eff} = \varepsilon_{tag}\frac{1}{N}\sum_N \mathcal{D}^2 \tag{4.20}$$

of the selected $B_s^0 \to J/\psi\phi$ signal candidates for data and simulation samples are given in

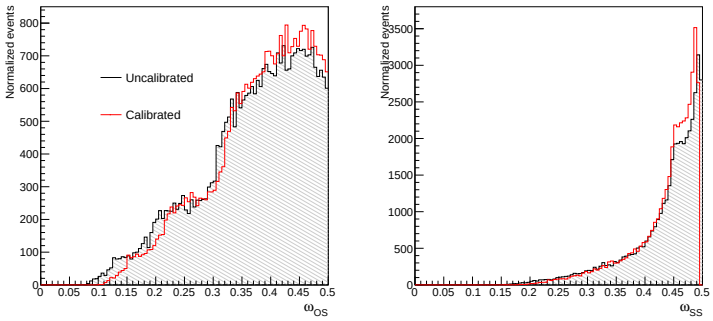


Fig. 4.28 Normalized distributions of the averaged mistag ω for $B_s^0 \to J/\psi\phi$ signal events from the (left) "OS-only" and (right) "SS-only" tagging categories before (black) and after (red) the calibration.

Table 4.20 for the three tagging categories. The overall tagging power is given by a sum of the tagging power in the three categories as $\varepsilon_{eff} = (4.93\pm0.16)\%$ and $(5.59\pm0.02)\%$ for data and MC events, respectively.

Table 4.20 Tagging efficiency ε_{tag}, square average tagging dilution $\mathcal{D}^2$ and tagging power ε_{eff} of the selected B_s^0 signal candidates for data and simulation samples in the three tagging categories. The column "Fraction" reports the fraction of events in each category out of the all tagged events.

Category	Fraction(%)	$\varepsilon_{tag}(\%)$	$\mathcal{D}^2$	$\varepsilon_{eff}(\%)$
		Data		
OS-only	15.6	11.75±0.75	0.1089±0.0103	1.28±0.09
SS-only	52.5	39.50±1.14	0.0272±0.0013	1.07±0.04
OS&SS	31.9	23.97±1.00	0.1077±0.0067	2.58±0.12
Total	100	75.22±1.69	0.0655±0.0026	4.93±0.16
		Simulation		
OS-only	13.8	10.60±0.06	0.1047±0.0011	1.11±0.01
SS-only	56.2	43.07±0.09	0.0360±0.0002	1.55±0.01
OS&SS	30.0	22.91±0.08	0.1279±0.0010	2.93±0.02
Total	100	76.58±0.13	0.0730±0.0003	5.59±0.02

The time dependent angular decay rate without taking into account any resolution and acceptance effect, $R(t,\Omega|q^{\rm alg},\eta^{\rm alg})$, can be written as [113]:

$$R(t,\Omega|q^{\rm alg},\eta^{\rm alg}) = (1+q^{\rm OS}(1-2\omega^{\rm OS}))(1+q^{\rm SS}(1-2\omega^{\rm SS}))R(t,\Omega|B_s^0)$$
$$+ (1-q^{\rm OS}(1-2\bar{\omega}^{\rm OS}))(1-q^{\rm SS}(1-2\bar{\omega}^{\rm SS}))R(t,\Omega|\bar{B}_s^0), \quad (4.21)$$

where $R(t,\Omega|B_s^0)$ and $R(t,\Omega|\bar{B}_s^0)$ are the time dependent angular decay rates for an initial B_s^0 and $\bar{B}_s^0$ defined in Eqs. 1.32 and 1.33, respectively; ω and $\bar{\omega}$ are derived from Eq. 4.17. By considering $q^{\rm OS}$=0 and $q^{\rm SS}$=0 as a special tagging decision, Eq. 4.21 provides a unified form of the time dependent angular decay rate for any of the four independent categories considered, i.e. no matter a tagging decision is (or both decisions are) zero. By considering

the efficiency of the OS&SS category, and the OS and SS efficiencies, a negligible correlation between OS and SS tag decisions is estimated, $|1 - \varepsilon_{\text{OS\&SS}}/(\varepsilon_{\text{OS}}\varepsilon_{\text{SS}})| \simeq 4\%$. The correlation between OS and SS mistag probabilities is also negligible (1.59% from the two-dimensional distribution in Fig. 4.29).

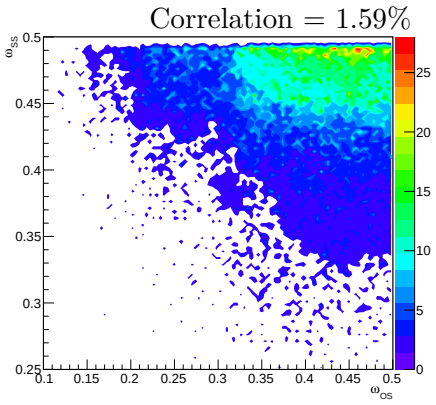


Fig. 4.29 OS mistag vs. SS mistag probability for $B_s^0 \to J/\psi\phi$ signal events of the OS&SS category.

4.6 Time dependent angular fit

The physics parameters are extracted from a four-dimensional maximum likelihood fit to the measured decay time and angular distributions. This requires the modelling of the distributions with theoretical functions that need to be corrected for acceptance and resolution effects. The basic concepts of a maximum likelihood estimation is discussed in detail in Refs. [114, 115]. The specific implementation for this analysis is explained in the following chapter.

4.6.1 PDF for the $B_s^0 \to J/\psi\phi$ decay analysis

The aim of the maximum likelihood fit is the determination of the physics observables in the differential decay rate of the B_s^0 meson decays. The parameters $\vec{\lambda}$ occurring in the PDF for the $B_s^0 \to J/\psi\phi$ can be separated in $\vec{\lambda} = \{\vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuis}}\}$, where $\vec{\lambda}_{\text{phys}}$ denotes the parameters that describe the physics observables and $\vec{\lambda}_{\text{nuis}}$ represents all additional parameters (nuisance parameters) that arise, for example, in the description of resolution effects or background contributions.

The dataset used to measure $\vec{\lambda}_{\text{phys}}$ is divided into several subsamples. They are distinguished by their tagging category (Sec. 4.5.4). Most of the parameters extracted in the maximum likelihood fit, in particular $\vec{\lambda}_{\text{phys}}$, are the same for all subsamples and can be fitted simultaneously exploiting the full statistical power of the dataset.

A dataset composed of signal and background events is generally described by a PDF consisting of a signal part $S(\vec{x};\vec{\lambda})$ and a background part $B(\vec{x};\vec{\lambda})$:

$$\mathcal{PDF}(\vec{x}|\vec{\lambda}) = f_{sig} \cdot S(\vec{x}|\vec{\lambda}) + (1 - f_{sig}) \cdot B(\vec{x}|\vec{\lambda}), \tag{4.22}$$

where f_{sig} is the relative fraction between signal and background. Both signal and background component depend on five measured observables $\vec{x}$:

- decay time t,
- three decay angles $\Omega = \{\cos\theta_K, \cos\theta_e, \phi_h\}$,
- reconstructed mass of the B_s^0 candidate, m, which is used to separate the signal and background components.

In addition, they depend conditionally on the estimated uncertainty of the decay time σ_t, the tagging decision q and calibrated mistag probability η of each event.

According to the tagging strategy discussed in Sec. 4.5.4, the events are divided into three tagging categories that are fitted simultaneously: q and η are either taken from OS and SS tagger or are a combination of both, OS&SS. For untagged events q is zero, for all other events q and η are defined as follows

$$\begin{array}{ccc} q = q^{\mathrm{OS}}, & \eta = \eta^{\mathrm{OS}} & \text{for } q^{\mathrm{OS}} = \pm 1 \text{ and } q^{\mathrm{SS}} = 0, \\ q = q^{\mathrm{SS}}, & \eta = \eta^{\mathrm{SS}} & \text{for } q^{\mathrm{SS}} = \pm 1 \text{ and } q^{\mathrm{OS}} = 0, \\ q = q^{\mathrm{OS\&SS}}, & \eta = \eta^{\mathrm{OS\&SS}} & \text{for } q^{\mathrm{OS}} = \pm 1 \text{ and } q^{\mathrm{SS}} = \pm 1, \\ q = 0, & \eta = 0 & \text{for } q^{\mathrm{OS}} = 0 \text{ and } q^{\mathrm{SS}} = 0. \end{array} \tag{4.23}$$

In the following sections the implementation of the signal and background components are discussed in detail.

Description of the signal decay

The reconstructed B_s^0 mass is uncorrelated to the decay time and angular distribution. Therefore, the signal PDF used to determine the physics parameters can be factorized into a term describing the reconstructed mass distribution S_M and a term describing the time and angular dependent decay rate S_t. As $S_t = S_t(t, \sigma_t, \Omega, \eta, q|\vec{\lambda})$ depends conditionally on the quantities σ_t, η and q, the signal PDF needs to contain as well the probability density functions of the estimated decay time uncertainty S_{σ_t}, calibrated mistag probability S_T and the tagging efficiency S_q:

$$S(\vec{x}|\vec{\lambda}) = S_M(m|\vec{\lambda}_{\mathrm{nuis}}) \cdot S_t(t, \sigma_t, \Omega, \eta, q|\vec{\lambda}_{\mathrm{phys}}, \vec{\lambda}_{\mathrm{nuis}}) \cdot S_{\sigma_t}(\sigma_t) \cdot S_T(\eta) \cdot S_q. \tag{4.24}$$

S_{σ_t} and S_T are represented in the fit by histograms shown in Figs. 4.30 and 4.31. $S_q = S_q^{\mathrm{OS}} \cdot S_q^{\mathrm{SS}} \cdot S_q^{\mathrm{OS\&SS}}$ is defined by the tagging efficiencies of the three categories with

$$S_q^{\mathrm{OS}} = \begin{cases} \frac{\varepsilon_{sig}^{\mathrm{OS}}}{2} & \text{for } q^{\mathrm{OS}} = \pm 1 \\ (1 - \varepsilon_{sig}^{\mathrm{OS}}) & \text{for } q^{\mathrm{OS}} = 0 \end{cases}, \tag{4.25}$$

and similar expressions for S_q^{SS} and $S_q^{\mathrm{OS\&SS}}$. The three tagging efficiencies for the signal decays can be determined directly in the minimization and are free parameters in the fit.

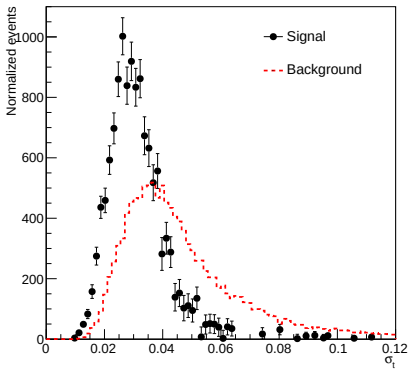


Fig. 4.30 Estimated decay time uncertainty of the selected $B_s^0 \to J/\psi\phi$ signal and background candidates.

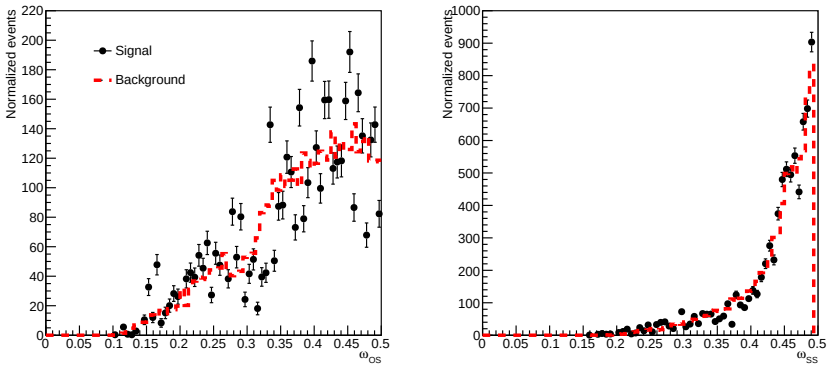


Fig. 4.31 Normalized distributions of the averaged mistag ω of the selected $B_s^0 \to J/\psi\phi$ signal and background candidates from the (left) "OS-only" and (right) "SS-only" tagging categories.

Mass distribution

The signal mass shape of the B_s^0 candidate is detaily described in Sec. 4.2 and is presented by a double Crystal Ball function given in Eq. 4.4.

Decay time and angular distributions

The signal distribution of the decay time and the decay angles $s(t, \Omega, \eta, q|\vec{\lambda})$ is the main part of the PDF containing all physics observables of interest. It is built from the differential decay rates for produced B_s^0 and $\bar{B}_s^0$ mesons as given in Eqs. 1.35 and 1.36, with $P_{B_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}) = \frac{\mathrm{d}\Gamma(B_s^0 \to J/\psi\phi)}{\mathrm{d}t\mathrm{d}\Omega}$ and $P_{\bar{B}_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}) = \frac{\mathrm{d}\Gamma(\bar{B}_s^0 \to J/\psi\phi)}{\mathrm{d}t\mathrm{d}\Omega}$:

$$s(t, \Omega, \eta, q|\vec{\lambda}) = \frac{1 + q\mathcal{D}_{tag}}{2} P_{B_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}) + \frac{1 - q\bar{\mathcal{D}}_{tag}}{2} P_{\bar{B}_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}), \tag{4.26}$$

where q is the tagging decision; $\mathcal{D}_{tag} = (1 - 2\omega_{tag})$ and $\bar{\mathcal{D}}_{tag} = (1 - 2\bar{\omega}_{tag})$ are the tagging dilutions with the mistag probability ω_{tag} and $\bar{\omega}_{tag}$ of produced B_s^0 and $\bar{B}_s^0$ mesons, respectively. The physics parameters are $\vec{\lambda}_{\text{phys}} = \{\phi_s, \Delta\Gamma_s, \Gamma_s, |\lambda|, \Delta m_s, |A_0|^2, |A_\perp|^2, F_S, \delta_\perp, \delta_\parallel, \delta_S\}$. The differential decay rate is normalized in such a way that the P-wave amplitudes satisfy the condition: $|A_\perp|^2 + |A_0|^2 + |A_\parallel|^2 = 1$. Therefore, only two of them are free parameters in the fit. As only phase differences are measurable quantities, the value of one strong phase can be fixed. The convention $\delta_0 = 0$ is chosen. The S-wave fraction is parametrized such that $F_S = 1/(|A_\perp|^2 + |A_0|^2 + |A_\parallel|^2 + |A_S|^2)$. The parametrization of δ_S is chosen as $\delta_S - \delta_\perp$ to decrease correlations between δ_S and $\delta_\perp$ in the fit. All physics parameters $\vec{\lambda}_{\text{phys}}$ are varied in the fit, except $\Delta\Gamma_s$, Γ_s and Δm_s which are constrained to the value $\Delta\Gamma_s = 0.0805 \pm 0.0091$ ps^{-1}, $\Gamma_s = 0.6603 \pm 0.0027$ ps^{-1} and $\Delta m_s = 17.711 \pm 0.056$ ps^{-1}, measured in a muon mode of the $B_s^0 \to J/\psi\phi$ decay [39].

Flavour tagging

For the three possible tagging decisions $q = \pm 1, 0$, Eq. 4.26 can be written as

$$\begin{aligned}
s(t, \Omega, \eta, q = -1|\vec{\lambda}) &= \omega_{tag} \cdot P_{B_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}) + (1 - \bar{\omega}_{tag}) \cdot P_{\bar{B}_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}), \\
s(t, \Omega, \eta, q = +1|\vec{\lambda}) &= (1 - \omega_{tag}) \cdot P_{B_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}) + \bar{\omega}_{tag} \cdot P_{\bar{B}_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}), \\
s(t, \Omega, \eta, q = 0|\vec{\lambda}) &= \frac{1}{2} \cdot P_{B_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}) + \frac{1}{2} \cdot P_{\bar{B}_s^0}(t, \Omega|\vec{\lambda}_{\text{phys}}).
\end{aligned} \tag{4.27}$$

This shows clearly that an imperfect knowledge of the B_s^0 production flavour (ω_{tag} and $\bar{\omega}_{tag}$ are not equal 0) leads to a mixture of the B_s^0 and $\bar{B}_s^0$ differential decay rates in the PDF.

As described in Sec. 4.5.4, ω_{tag} and $\bar{\omega}_{tag}$ depend on calibration parameters and the calibrated mistag probability η. Depending on the tagging category, different calibration parameters appear in the PDF: $p_0^{\text{alg}}, p_1^{\text{alg}}, \eta^{\text{alg}}, \Delta p_0^{\text{alg}}$ and Δp_1^{alg} with alg = OS, SS and OS&SS. All these calibration parameters are varied in the fit inside their uncertainties using Gaussian constraints. Their measured values and uncertainties are given in Table 4.19.

Acceptance and resolution effects

Eq. 4.26 describes the decay time and angular distributions ignoring any detector or resolution effects. Taking these into account and normalizing $s(t,\Omega,\eta,q|\vec{\lambda})$ according to $\int_D \mathcal{PDF}(\vec{x}|\vec{\lambda})\mathrm{d}\vec{x} = 1$ where D denotes the domain of $\vec{x}$, the decay time and angular dependent decay rate leads to

$$S_t(t,\sigma_t,\Omega,\eta,q|\vec{\lambda}) = \frac{\varepsilon(t,\Omega)\cdot s(t,\Omega,\eta,q|\vec{\lambda})\otimes\mathcal{R}(t|\sigma_t)}{\int \varepsilon(t,\Omega)\cdot s(t,\Omega,\eta,q|\vec{\lambda})\otimes\mathcal{R}(t|\sigma_t)\mathrm{d}t\mathrm{d}\Omega}, \tag{4.28}$$

where $\varepsilon(t,\Omega)$ is the detector acceptance including trigger, reconstruction and selection inefficiencies. The $\varepsilon(t,\Omega) = \varepsilon(t)\cdot\varepsilon(\Omega)$ has to be described in four dimensions, as the distributions of the decay time and the three angles can be correlated. The effect of neglecting this correlation is much smaller than the statistical precision of this analysis. The $s(t,\Omega,\eta,q|\vec{\lambda})$ is convoluted for each event with a Gaussian resolution function $\mathcal{R}(t-t_0|\sigma_t)$ with a zero mean and width σ_t to account for the limited decay time resolution of the detector as discussed in Sec. 4.3.2. A possible angular resolution effects are neglected in the analysis as earlier studies proved them to be negligible (Sec. 4.4.2, Ref. [28]).

The decay time acceptance $\varepsilon(t)$ is determined in Sec. 4.3.1. This acceptance effect is included in the PDF as one-dimensional histograms given in Fig. 4.21. The angular acceptance $\varepsilon(\Omega)$ is calculated using simulated events and given by a three-dimensional histogram, described in Sec. 4.4.1. The one-dimensional projections of this histogram are shown in Fig. 4.25, corresponding to 40 bins in each of the decay angles.

The denominator of $S_t(t,\sigma_t,\Omega,\eta,q|\vec{\lambda})$ also depends on the decay time and angular acceptances. Using Eq. 1.35, one can transform the normalization such that the integral factorizes into angular and time dependent components:

$$\int \varepsilon(t)\cdot\varepsilon(\Omega)\cdot P_{B_s^0(\bar{B}_s^0)}(t,\Omega|\vec{\lambda}_{\mathrm{phys}})\mathrm{d}t\mathrm{d}\Omega \to \sum_{i=1}^{10}\int\varepsilon(\Omega)f_i(\Omega)\mathrm{d}\Omega\cdot\int\varepsilon(t)h_i(t)(\bar{h}_i(t))\mathrm{d}t. \tag{4.29}$$

The 10 normalization weights $\xi_i = \int\varepsilon(\Omega)f_i(\Omega)\mathrm{d}\Omega$ are independent of any parameter in the minimization process and can therefore be determined once before the fit (Table 4.17). This is not possible for the time integral which must be calculated numerically for each event. The absolute value of the normalization weight shows that the necessary acceptance corrections are rather small.

Description of the background

Similar to the signal description, the background PDF $B(\vec{x}|\vec{\lambda}_{\mathrm{nuis}})$ is a product of a mass term B_M, a component describing the decay time and angular distribution B_t and the distributions

of the conditional per-event variables for background events, B_{σ_t}, B_T and B_q:

$$B(\vec{x}|\vec{\lambda}_{\text{nuis}}) = B_M(m|\vec{\lambda}_{\text{nuis}}) \cdot B_t(t, \sigma_t, \Omega|\vec{\lambda}_{\text{nuis}}) \cdot B_{\sigma_t}(\sigma_t) \cdot B_T(\eta) \cdot B_q. \tag{4.30}$$

The background in this analysis shows no correlation between decay time and decay angles, therefore one can factorize $B_t(t, \sigma_t, \Omega|\vec{\lambda}_{\text{nuis}}) = B_t(t, \sigma_t|\vec{\lambda}_{\text{nuis}}) \cdot B_\Omega(\Omega)$. Similar to the signal PDF, B_{σ_t} and B_T are represented by histograms determined from combinatorial and partially reconstructed background candidates and shown in Figs. 4.30, 4.31. $B_q = B_q^{\text{OS}} \cdot B_q^{\text{SS}} \cdot B_q^{\text{OS\&SS}}$ takes into account the conditional dependence of B on the tagging decision q and is defined analogously to the signal PDF (Eq. 4.25) as

$$B_q^{\text{OS}} = \begin{cases} \frac{\varepsilon_{bkg}^{\text{OS}}}{2} & \text{for } q^{\text{OS}} = \pm 1 \\ (1 - \varepsilon_{bkg}^{\text{OS}}) & \text{for } q^{\text{OS}} = 0 \end{cases}, \tag{4.31}$$

with similar expressions for the SS and OS&SS categories. All three background tagging parameters are varied in the fit.

Mass distribution

The PDF describing the mass shape of the combinatorial background and partially reconstructed events is discussed in Sec. 4.2.1 and is given by a normalized exponential and Gaussian functions, respectively.

4.6.2 Results of the maximum likelihood fit

Table 4.21 shows the results of the maximum likelihood fit to 3 fb^{-1} of data including decay time and angular acceptance effects and decay time resolution. The tagging calibration parameters, $\Delta\Gamma_s$, Γ_s and Δm_s are Gaussian constrained in the fit. The parameter $|\lambda|$ is consistent with unity, implying no evidence for $\mathcal{CP}$ violation in the $B_s^0 \to J/\psi(e^+e^-)\phi$ decays. The correlation matrix of the fit parameters is shown in Table 4.22.

The MINOS program [116] has been used to determine the asymmetric uncertainties on the parameters. The corresponding one-dimensional likelihood profiles can be found in App. C.10. The likelihood profile for δ_S is not parabolic and the 95% confidence level range is [-0.6, 0.6] rad.

The measurement of the phase ϕ_s in the $B_s^0 \to J/\psi(e^+e^-)\phi$ decays is consistent with the one measured in $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ decays [39] and with the SM prediction [18].

4.7 Studies of systematic effects

The possible sources of systematic uncertainties have to be investigated in the analysis. Some of them are already taken into account in the errors of the fitted parameters. The other

Table 4.21 Results of the unbinned maximum likelihood fit to the selected $B_s^0 \to J/\psi(e^+e^-)\phi$ candidates including all acceptance and resolution effects. The uncertainty is statistical.

Parameter	Fit result and error
$A_\perp^2$	0.257 ± 0.022
A_0^2	0.508 ± 0.018
$\delta_\parallel$ [rad]	$3.06\ ^{+0.22}_{-0.21}$
$\delta_\perp$ [rad]	2.51 ± 0.44
F_S	$0.021\ ^{+0.032}_{-0.02}$
δ_S [rad]	$0.04\ ^{+0.55}_{-0.48}$
ϕ_s [rad]	$-0.18\ ^{+0.37}_{-0.39}$
$\lvert\lambda\rvert$	$0.83\ ^{+0.084}_{-0.226}$

Table 4.22 Statistical correlation matrix from nominal fit.

	$A_\perp^2$	A_0^2	$\delta_\parallel$	$\delta_\perp$	F_S	δ_S	Δm_s	ϕ_s	$\lvert\lambda\rvert$
$A_\perp^2$	1.00	-0.36	0.04	0.08	-0.14	0.01	0.01	0.10	-0.08
A_0^2		1.00	0.03	-0.11	0.15	0.02	-0.00	-0.07	0.12
$\delta_\parallel$			1.00	0.07	0.13	-0.04	-0.02	-0.04	0.17
$\delta_\perp$				1.00	-0.52	-0.20	0.13	-0.11	-0.57
F_S					1.00	-0.02	-0.04	-0.16	0.66
δ_S						1.00	-0.01	0.05	0.18
Δm_s							1.00	-0.11	-0.06
ϕ_s								1.00	-0.13
$\lvert\lambda\rvert$									1.00

possible uncertainties occur, for example from the background description or the decay time and angular acceptances. All systematic uncertainties are described below and a summary is reported in Table 4.24.

4.7.1 Factorization of $m(e^+e^-K^+K^-)$ with decay time and angles

Since the nominal fit is performed using sWeights, it is essential to verify that the control variable ($m(e^+e^-K^+K^-)$) used to determine the sWeights is uncorrelated with the variables used in the maximum likelihood fit. The mass fits to the sWeighted data sample are performed for different bins of the decay time and angles. A non zero dependence of the mass resolution with $\cos\theta_e$ has been observed. To assign a systematic uncertainty the sWeights are recomputed using best mass fit in each bin of $\cos\theta_e$ and a full fit is repeated with new sWeighted data sample. The variations in the physics parameters between these fits and the nominal fit are used as a systematic uncertainty (App. C.11).

4.7.2 $m(e^+e^-K^+K^-)$ mass model

In order to estimate the systematic uncertainty for the mass model, the reconstructed $m(e^+e^-K^+K^-)$ distribution is fitted using an alternative function for the signal candidates,

a double Ipatia [117]. An exponential and Gaussian functions are used for the combinatorial and partially reconstructed background events, respectively. The result from this fit is shown in App. C.12. The bias between the physics parameters determined using the sWeights from the alternative mass fit compared to the nominal value is taken as a systematic uncertainty (Table 4.24).

4.7.3 Λ_b background

The only identified peaking background is the $\Lambda_b \to J/\psi(ee)pK^-$ decays as described in Sec. 4.1.5. The fraction of such events is estimated using MC sample to be at most 1.03%. The Λ_b contribution is evaluated for the data sample by fitting the reconstructed $m(e^+e^-K^+K^-)$ distribution with mass model where the $\Lambda_b \to J/\psi pK$ decays are defined using a double CB function. The shape of the Λ_b events is fixed to MC fit. As in the signal decay mode, the peaking background of the $B^0 \to J/\psi K^*$ decay used to determine the decay time acceptance (Sec. 4.3.1) is considered, namely the $\Lambda_b \to J/\psi(ee)p\pi^-$ decays where the p has been misidentified as a K^+. In order to estimate its contribution in the reconstructed B^0 mass distribution, the $\Lambda_b \to J/\psi p\pi^-$ simulated events are reconstructed with $B^0 \to J/\psi K^*$ hypothesis and the full selection procedure is applied (Sec. 4.3.1). Taking into account its branching fraction from Ref [100] it is found to be 0.85% of the events under the B^0 meson peak. The contribution of the $\Lambda_b \to J/\psi p\pi$ decays is evaluated for the data sample by fitting the reconstructed $e^+e^-K^+\pi^-$ mass distribution with mass model where the Λ_b candidates are described using a double CB function. The details of this study are discussed in App. C.13. The shift between the physics parameters determined using the sWeights from the mass fit and decay time acceptance with Λ_b contribution compared to the nominal value is assigned as a systematic uncertainty.

4.7.4 Angular acceptance

A systematic error associated with angular acceptance is evaluated by coherently varying the normalization weights within their uncertainties, taking into account the correlations between them. The nominal fit is repeated 460 times each with a new set of weights that are shifted from their nominal values by a random Gaussian variation, using the covariance matrix as an input. The resulting RMS of the distribution for each physics parameter is assigned as a systematic error for that parameter (Fig. 4.32).

4.7.5 Decay time resolution

A systematic contribution was evaluated for effect of the finite decay time resolution by fitting the data with enlarged constraint on the decay time resolution parameters. The nominal constraint on the free parameters of the decay time resolution is determined as a

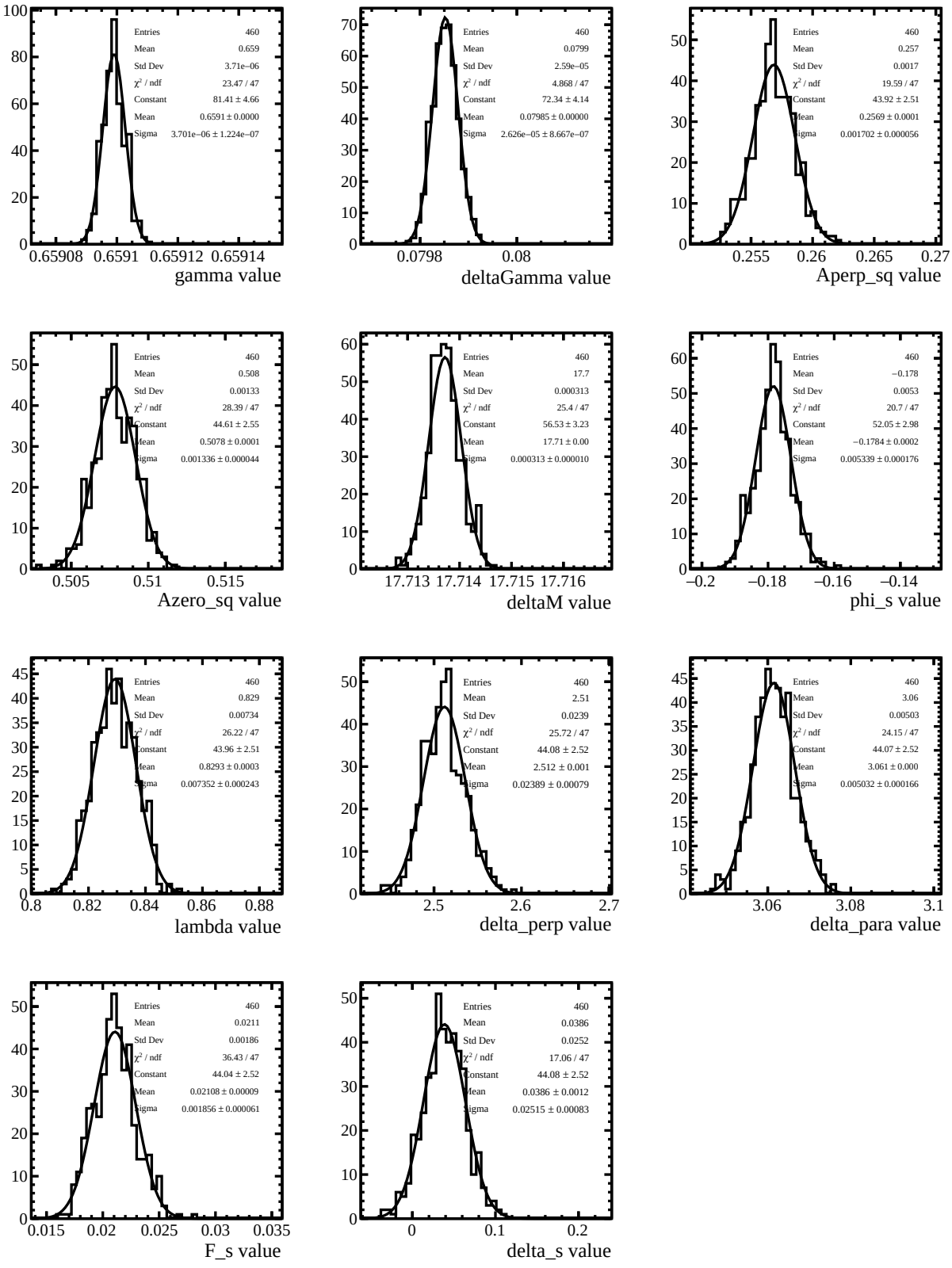


Fig. 4.32 Distributions of the physics parameters from 460 fits to data, varying the values of the angular acceptance weights within their uncertainties (taking into account correlations between weights). The RMS of the distributions is taken as a systematic uncertainty.

double difference between value parameters from the fit to the electron and muon mode of the $B_s^0 \to J/\psi\phi$ simulated samples as discussed in Sec. 4.3.2. The fit procedure was repeated with the constraint set as 3 and 4 times this difference (Table 4.23). The largest variation between the physics parameters obtained with these fits and the nominal values is assigned as a systematics uncertainty (App. C.14). In order to test the maximum scale of the dependence of the results on the decay time resolution parameters, the nominal fit is repeated with fixed parameters of the decay time resolution. This resulted in a negligible variation to all parameters except $\delta_\perp$ and λ, for which the bias is 0.014 rad and 0.002, respectively. This corresponds to a change of less than 3% of its statistical uncertainty. No systematic uncertainty is assigned for this variation.

Table 4.23 Different constraints on the free fit parameters of the decay time resolution model for data samples.

Parameters	Value	Constraint×2	Constraint×3	Constraint×4
$b^{'}$	-3.99	13.14	19.72	26.29
$b^{''}$	-2.23	15.64	23.46	31.27
$c^{'}$	1.4523	0.263	0.394	0.526
$c^{''}$	0.3925	0.479	0.718	0.958
f	0.243	0.090	0.135	0.180

In addition, the input fit values of the decay time resolution parameters are varied based on results obtained from a fit to the MC samples in repeated realisations of the final fit (Table 4.13). No changes in fit result is observed and therefore no systematic uncertainty is assigned.

The sensitivity to the constraint on the $\Delta\Gamma_s$ and Γ_s parameters has been studied by allowing these parameters to float in the fit. The difference between the fit results without constraint on $\Delta\Gamma_s$ and Γ_s and the nominal is taken as a systematic and reported in App. C.14. The obtained value of the lifetime parameters is $\Delta\Gamma_s = 0.117 \pm 0.049$ ps^{-1} and $\Gamma_s = 0.615 \pm 0.016$ ps^{-1} that gives a similar sensitivity compared to the $B_s^0 \to \psi(2S)\phi$ decay [42].

4.7.6 Decay time acceptance

The decay time acceptance is accounted for in the maximum likelihood fit by using the functional form given in Fig. 4.21. The acceptance values in each bin are fixed in the nominal fit. A systematic uncertainty is estimated by generating a new decay time acceptance histogram where each bin value has been changed via a random Gaussian shift (where the width of the Gaussian is given by the uncertainty of the value in that bin). The fit is repeated with the new histogram and the process is repeated 460 times. The RMS of the extracted physics parameters distribution is assigned as a systematic uncertainty (Fig. 4.33).

As there are multiple components that feed into the evaluation of the decay time acceptance, a systematic uncertainty is assigned for each. First, a systematic due to the model used in

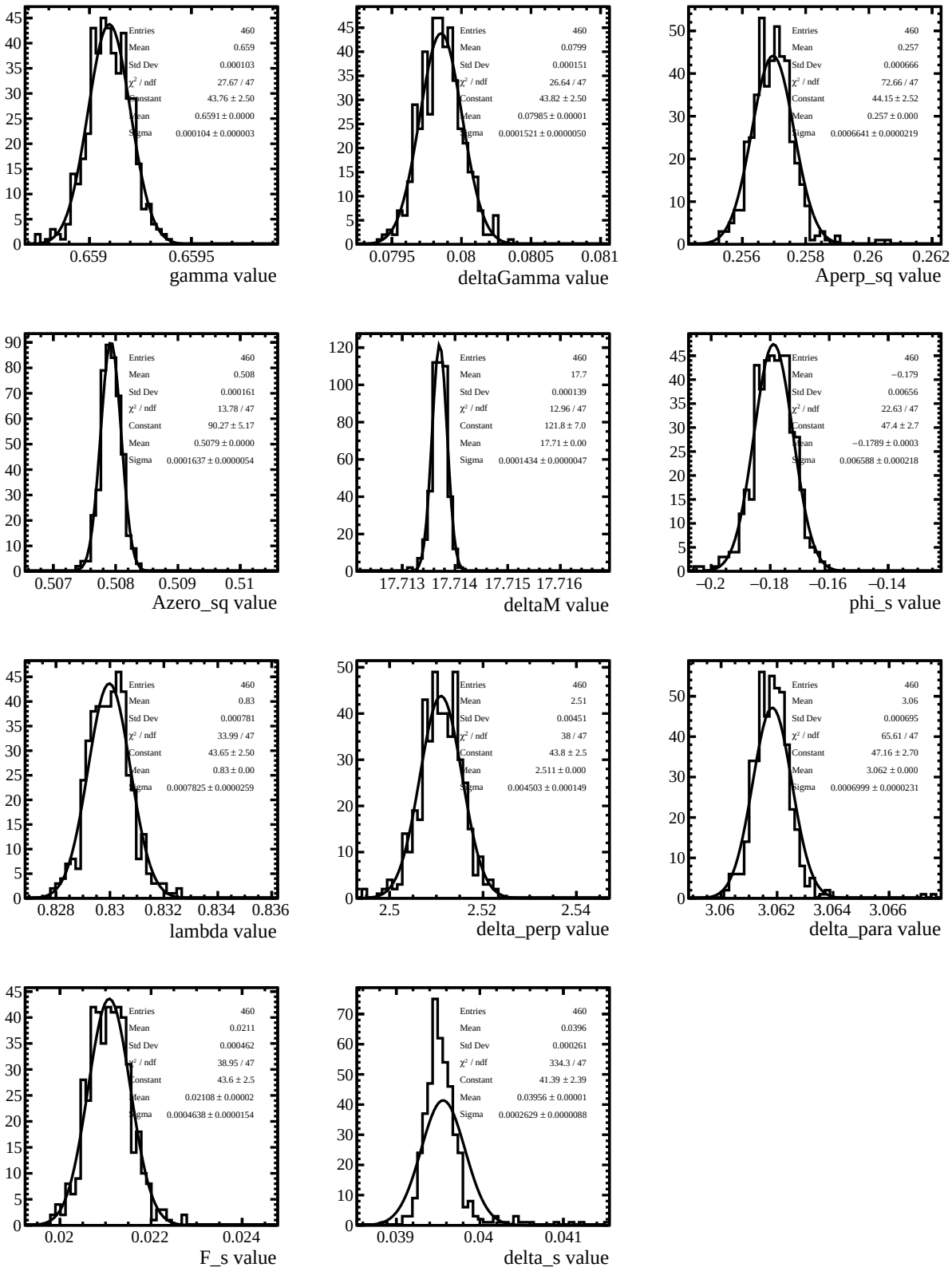


Fig. 4.33 Distributions of the physics parameters from 460 fits to data, varying the bins of the decay time acceptance histogram within their uncertainties. The RMS of the distributions is taken as a systematic uncertainty.

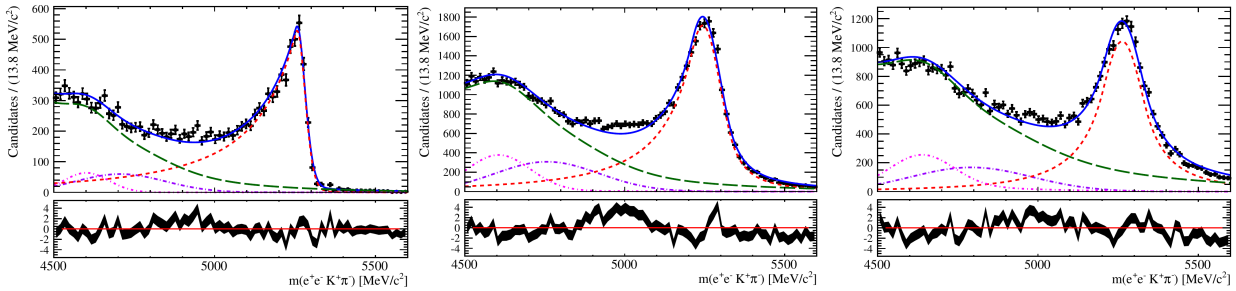


Fig. 4.34 Fit to the resonstructed $m(e^+e^-K^+\pi^-)$ distribution of selected $B^0 \to J/\psi K^*$ decays using a double Ipatia function for the signal candidates, an exponential for the combinatorial background and a double and single Gaussian for partially reconstructed background. The dataset is divided into three Bremsstrahlung categories: (left) without, (middle) one and (right) more reconstructed photons coming from the electron radiation.

the fit of the $e^+e^-K^+\pi^-$ mass distribution is considered. The mass distribution is fit with an alternative model, double Ipatia function (Fig. 4.34), and a new decay time acceptance histogram is calculated. The difference between the fit results using the new decay time acceptance histogram and the nominal is taken as a systematic uncertainty. The variation within 0.001 rad to $\delta_\perp$ and 0.002 rad to ϕ_s is observed and is negligible for all other parameters.

Secondly, when evaluating the decay time acceptance of the B^0 mode, the ratio of the decay time distribution in data sample is taken with simulated events, generated with the known B^0 lifetime [20]. In order to estimate a systematic for this, the B^0 acceptance histogram is re-evaluated where the simulated events are generated with the B^0 lifetime shifted by $\pm 1\sigma$. No changes in fit result is observed in this test and therefore no systematic uncertainty is assigned.

The sensitivity to the BDT selection has been studied using three approaches. First, the BDT criterium is adjusted around the nominal optimal point for the signal sample. In each case the time and angular acceptance is recomputed, including the control-signal correction factor r (Sec. 4.3.1), and the maximum likelihood fit repeated. The maximal difference between the fit results using the new BDT criterium and the nominal is taken as a systematic uncertainty. This is results in a significant variation for $\delta_\perp$ (0.082 rad) and ϕ_s (0.166 rad) as shown in Fig. 4.35. There is no evidence of a trend nor of a variation outside of statistical expectation. Second, the BDT selection on the signal mode is kept at the nominal optimal point while the BDT criterium on the control mode is adjusted. No parameters were affected with respect to their statistical errors and therefore no systematic uncertainty is assigned related to the BDT selection. Finally, to test the maximum scale of the dependence of the results upon the simulation, the nominal fit is repeated when ignoring the correction factor r from the decay time acceptance. This resulted in a variation to all parameters and therefore a systematic uncertainty is assigned for this extreme variation. No bias in fit result is observed when changing the binning of the efficiency histogram from 40 to 20 bins and no systematic uncertainty is assigned.

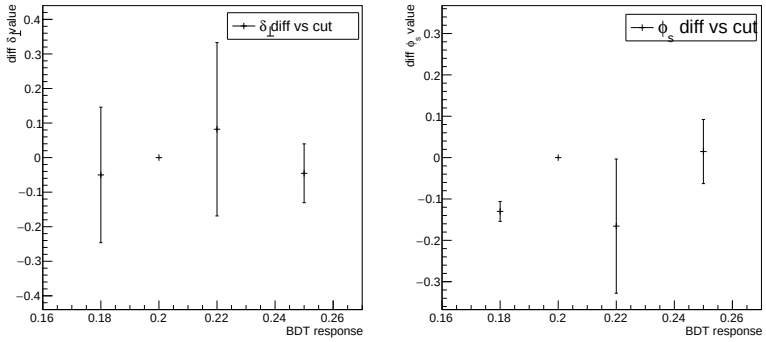


Fig. 4.35 Variation in the fit results for (left) $\delta_\perp$ and (right) ϕ_s between the nominal fit and the fit performed with a different BDT selection for signal channel. The nominal BDT criterium is > 0.2. The y-axis range corresponds to the nominal statistical uncertainty for $\delta_\perp$ or ϕ_s. The errors for each point are calculated as the quadrature difference of the uncertainty between the two fit results, in order to remove statistical correlations.

4.7.7 Tagging

The tagging parameters are allowed to float in the fit (Gaussian constrained within their uncertainty). The systematic uncertainty from the flavour tagging is evaluated by fixing and freeing the tagging parameters in the fit and taking a difference in quadrature of the obtained uncertainties on each physics parameter. The contributions are 0.001 rad to the statistical uncertainty on ϕ_s and $\delta_\perp$ and are negligible for all other parameters.

4.7.8 Length and momentum scale

The uncertainty on the LHCb length scale is estimated to be at most 0.020% [118], which translates directly into an uncertainty on Γ_s, $\Delta\Gamma_s$ and Δm_s of 0.020% with other parameters being unaffected. The momentum scale uncertainty is at most 0.0025% [118]. As it affects both the reconstructed momentum and mass of the B_s^0 meson, it cancels to a large extent and the resulting effect on Γ_s and $\Delta\Gamma_s$ is negligible.

4.7.9 Contribution from B_c^+ decays

The LHCb has observed the $B_c^+ \to B_s^0\pi^+$ decay mode and has measured $\sigma(B_c^+)/\sigma(B_s^0) \times \mathcal{BR}(B_c^+ \to B_s^0\pi^+) = [2.37 \pm 0.31 \pm 0.11^{+0.17}_{-0.13}] \times 10^{-3}$ [119]. As such, a small fraction of $B_s^0 \to J/\psi\phi$ signal candidates may come from the decay of a B_c^+ meson. Since the B_c^+ meson has a non-zero lifetime ∼0.4 ps, it can cause an average positive shift in the reconstructed decay time of the B_s^0 meson, which can bias a measurement of Γ_s. The $\Delta\Gamma_s$ is a difference of (inverse) lifetimes so, on average, will not be affected. The B_c^+ fraction was estimated as 0.8% in Ref. [39] and simulated experiments were used to assess the impact of ignoring such a contribution to the extraction of the physics parameters. Only Γ_s was observed to be affected, with a bias on its central value of 0.2σ, which was assigned as a systematic

uncertainty. The same uncertainty is used here, where the assumption has been made that the ratio of efficiencies for selected $B_s^0 \to J/\psi(ee)\phi$ decays either promptly or via the decay of a B_c^+ meson is the same as that for $B_s^0 \to J/\psi(\mu\mu)\phi$ decays.

4.7.10 Angular resolution

The systematic effect of ignoring the angular resolution in the nominal fit to the data is determined using an ensemble of simulated data samples. Each data sample is generated using the nominal fit model, with the same number of signal events as in the data sample. Using the parameters in Table 4.18 the helicity angles are smeared and a new data sample is created. The nominal fit model is used to fit to both smeared and unsmeared data samples and the toy-by-toy differences in fit parameters are recorded. The distributions of these differences are shown in Fig. 4.36, normalized by the corresponding statistical uncertainty. The biases are small with the $\Delta\Gamma_s$ and polarization amplitude parameters affected. The observed biases in these distributions are assigned as a systematic uncertainty. In addition, all distributions have a finite width, w, indicating a statistical uncertainty related to ignoring the resolution in the fit. The approach used in Ref. [39] was to scale the statistical uncertainty of each parameter such that $\sigma \to \sigma\sqrt{1+w^2}$. However, in this case, the largest effect is to several polarization amplitude parameters where the statistical error increases by only 1-2% as reported in Table 4.24.

4.7.11 Fit bias

A possible bias of the fitting procedure is investigated by generating and fitting many simulated pseudo-experiments of equivalent size to the data sample, generated with physics parameters close to those obtained in the nominal fit. The resulting biases are small, and those which are not compatible with zero within two standard deviations are quoted as systematic uncertainties. The details of this study are reported in App. C.15.

4.7.12 Nuisance $\mathcal{CP}$ asymmetries

A measurement of the asymmetry that results from $\mathcal{CP}$ violation in the interference between $B_s^0 - \bar{B}_s^0$ meson mixing and decay is potentially affected by $\mathcal{CP}$ violation in the mixing, direct $\mathcal{CP}$ violation in the decay, production asymmetry and tagging asymmetry. As in Ref. [28], in this analysis the fit parameter $|\lambda|$ (Eq. 1.30) as one central quantity associated to $\mathcal{CP}$ violation is used, separate tagging calibrations are used for B_s^0 and $\bar{B}_s^0$ decisions, as well as separate normalizations of the PDF for each tagging decision. As a result, any residual effects due to tagging efficiency asymmetry and production asymmetry are shown to be negligible through simulation studies in the validation of the fit.

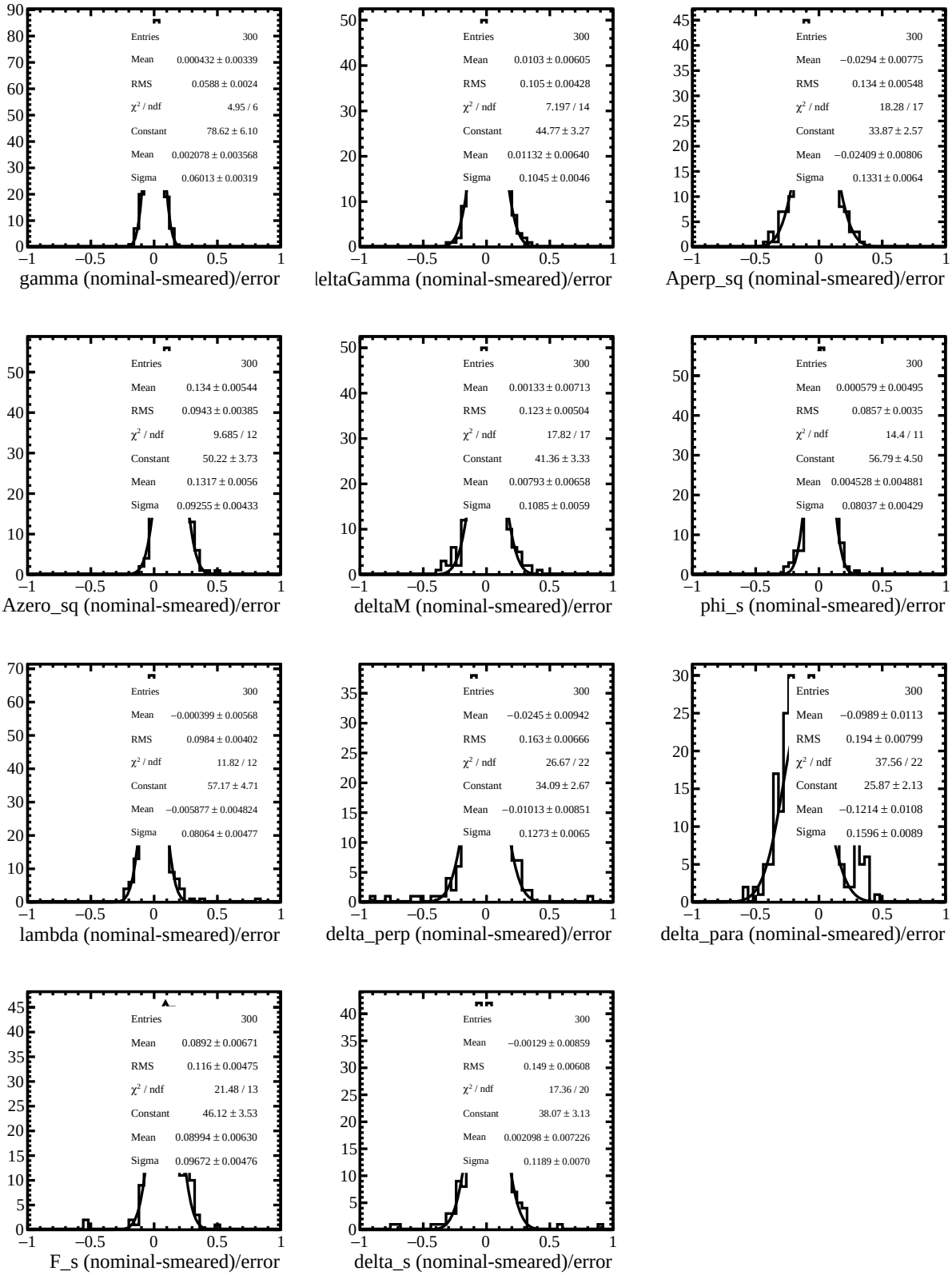


Fig. 4.36 Distributions of the difference of the fitted parameters divided by their statistical uncertainty obtained from a fit to a data sample first without (nominal) and then with (smeared) angular resolution applied to the dataset. The 300 toy experiments are used, each containing 12200 events.

4.7.13 Summary of systematic uncertainties

A summary of the systematic uncertainties evaluated in the last sections is given in the Table 4.24. The statistical uncertainties obtained from the fit are shown for comparison. Except for polarization amplitudes, $|A_\perp|^2$, $|A_0|^2$ and F_S, the statistical uncertainties are dominant for all physics parameters. For the $\mathcal{CP}$ violation parameters ϕ_s and $|\lambda|$, the main contribution to the systematic uncertainties is due to the sensitivity to the BDT selection and Λ_b peaking background. The sensitivity of the polarization amplitudes is limited by the determination of the decay time resolution.

Additional checks have been performed by repeating the nominal fit to data in bins of year of data taking, magnet polarity, number of primary vertices in each event and B_s^0 p_{T}. The results of these are documented in App. C.16. All subsets are consistent with the nominal fit to the full data set.

A closure test is performed by applying the full analysis chain on the simulated $B_s^0 \to J/\psi(ee)\phi$ sample (App. C.17). No bias is observed within the statistical precision of the available sample and no systematic uncertainty is assigned.

Table 4.24 Statistical and systematic uncertainties. Fields containing a dash (-) correspond to systematic uncertainties that are negligible.

Source	$A_\perp^2$	A_0^2	$\delta_\parallel$ [rad]	$\delta_\perp$ [rad]	F_S	δ_S [rad]	ϕ_s [rad]	$\lvert\lambda\rvert$
Stat. uncertainty	0.022	0.018	$^{+0.22}_{-0.21}$	0.44	$^{+0.032}_{-0.02}$	$^{+0.55}_{-0.48}$	$^{+0.37}_{-0.39}$	$^{+0.084}_{-0.226}$
Ang. resol. (scale factor)	1.009	1.004	1.013	1.008	1.005	1.007	1.003	1.003
Total stat. uncertainty	0.022	0.018	$^{+0.22}_{-0.21}$	0.44	$^{+0.032}_{-0.02}$	$^{+0.55}_{-0.48}$	$^{+0.37}_{-0.39}$	$^{+0.084}_{-0.227}$
Mass factorization	0.001	0.004	0.01	0.08	0.015	0.01	0.01	0.029
Mass model	–	–	0.01	0.04	0.002	–	0.01	0.011
Ang. acc. (stat.)	0.002	0.001	0.01	0.02	0.002	0.03	0.01	0.007
Time res. (Γ_s)	0.021	0.011	0.01	0.02	0.006	–	0.05	0.01
Time res. (constr.)	0.002	0.001	0.01	0.12	0.008	0.03	0.04	0.031
Time acc. (stat.)	0.001	–	–	0.01	–	–	0.01	0.001
Time acc. (B^0 data)	0.011	0.003	0.01	0.06	0.007	–	0.07	0.011
BDT selection	0.013	0.002	0.05	0.08	0.023	0.05	0.17	0.034
Λ_b background	0.003	0.03	0.07	0.03	0.019	0.01	0.15	0.044
Angular resolution	–	0.001	–	–	0.001	–	–	–
Fit bias	0.002	–	–	–	–	–	–	–
Quad. sum of syst.	0.027	0.032	0.09	0.18	0.036	0.07	0.25	0.073
Total uncertainty	0.035	0.037	$^{+0.24}_{-0.23}$	0.48	$^{+0.048}_{-0.041}$	$^{+0.55}_{-0.49}$	$^{+0.45}_{-0.46}$	$^{+0.111}_{-0.238}$

4.8 Final results

Using a dataset corresponding to an integrated luminosity of 3 fb^{-1} collected by the LHCb experiment in pp collisions during LHC Run 1, a tagged time dependent angular analysis of 12195±497 $B_s^0 \to J/\psi(e^+e^-)\phi(K^+K^-)$ decays is performed. The effective decay time resolution and effective tagging power are 44 fs and 4.93%, respectively. The analysis provides access to a number of different physics parameters including the $\mathcal{CP}$-violating phase of the B_s^0 meson system as well as the transversity amplitudes and strong phases of the decay. The final results are reported in Table 4.25.

Table 4.25 Results of the unbinned maximum likelihood fit to the selected $B_s^0 \to J/\psi(e^+e^-)\phi$ candidates including all acceptance and resolution effects. The tagging calibration parameters, $\Delta\Gamma_s$, Γ_s and Δm_s are Gaussian constrained in the fit. The first uncertainty is statistical while the second uncertainty is systematic.

Parameter	Fit result and errors
$A_\perp^2$	$0.257 \pm 0.022 \pm 0.027$
A_0^2	$0.508 \pm 0.018 \pm 0.032$
$\delta_\parallel$ [rad]	$3.06\ ^{+0.22}_{-0.21} \pm 0.09$
$\delta_\perp$ [rad]	$2.51 \pm 0.44 \pm 0.18$
F_S	$0.021\ ^{+0.032}_{-0.02} \pm 0.036$
δ_S [rad]	$0.04\ ^{+0.55}_{-0.48} \pm 0.07$
ϕ_s [rad]	$-0.18\ ^{+0.37}_{-0.39} \pm 0.25$
$\lvert\lambda\rvert$	$0.83\ ^{+0.084}_{-0.227} \pm 0.073$

The one-dimensional projections of the fitted PDF for decay time and angular distributions, obtained from the sFit technique, are shown in Fig. 4.37.

4.8.1 Interpretation of the results

The performed measurement of the $\mathcal{CP}$-violating parameters is the first determination of these quantities in the $B_s^0 \to J/\psi\phi$ decay with electron final state. The parameter $|\lambda|$ within two statistical deviations is consistent with unity, implying no evidence for $\mathcal{CP}$ violation in $B_s^0 \to J/\psi(e^+e^-)\phi(K^+K^-)$ decays. The ϕ_s result of the LHCb measurement using $B_s^0 \to J/\psi(e^+e^-)\phi$ decays is $\phi_s = -0.18^{+0.37}_{-0.39}\pm0.25$ rad that is consistent with previous LHCb measurements [39–43], the SM predictions [18], and show no evidence of $\mathcal{CP}$ violation in the interference between B_s^0 meson mixing and decay.

The result presented in this thesis can roughly be compared with the LHCb measurement using muon mode of the $B_s^0 \to J/\psi\phi$ [120] since both measurements use the same integrated luminosity, 3.0 fb^{-1}. The measured value of the phase ϕ_s will contribute to increased precision in the global average of the B_s^0 mixing parameters (Fig. 1.10).

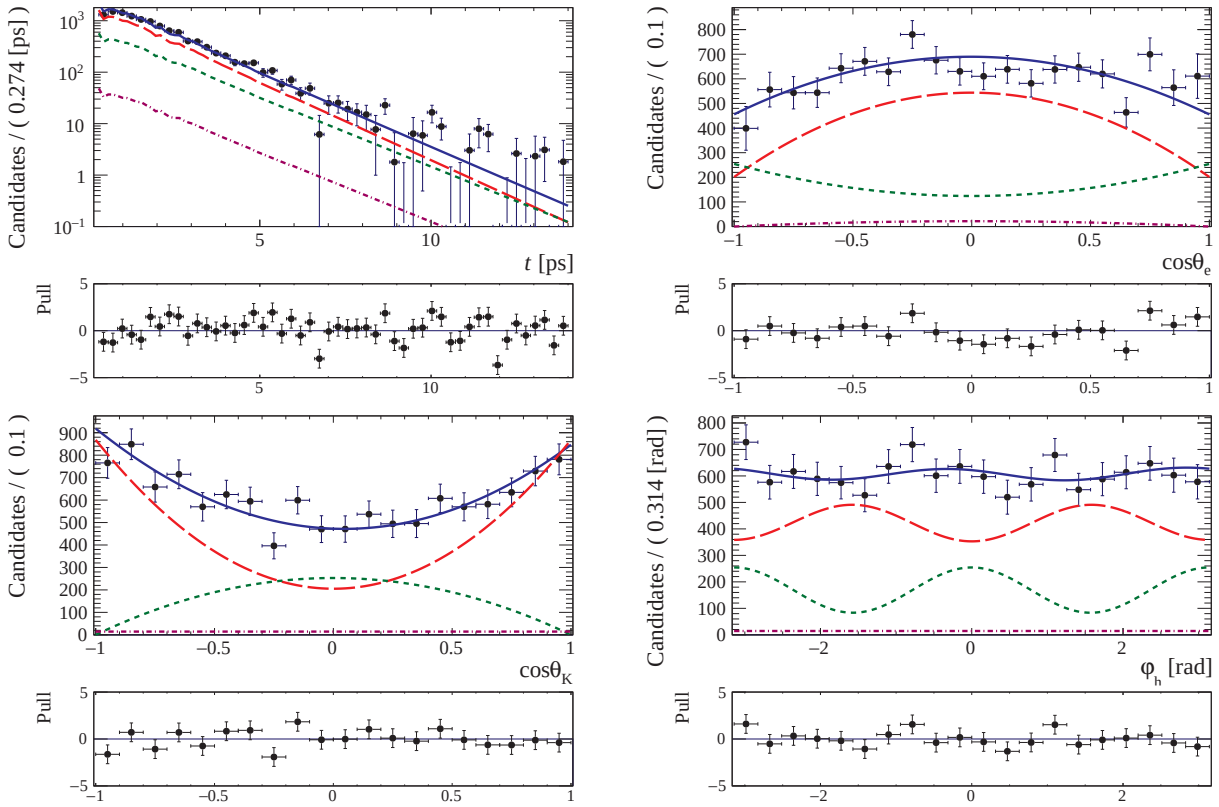


Fig. 4.37 Decay time and helicity-angle distributions for $B_s^0 \to J/\psi(e^+e^-)\phi$ reconstructed candidates (data points) with the one-dimensional projections of the PDF at the maximum likelihood fit. The solid blue line shows the total signal contribution, which is composed of $\mathcal{CP}$-even (long dashed red), $\mathcal{CP}$-odd (short-dashed green) and S-wave (dash-dotted purple) contributions.

Chapter 5

Summary and Outlook

The studies presented in this thesis cover two different aspects of the LHCb experiments.

The first one is the alignment of the tracking stations using 2011 and 2012 data and Monte Carlo simulation. The method to align the LHCb detector is presented which uses Kalman-fitted tracks coming from the standard track fit. The power of the procedure lies in the derivation of the global track covariance matrix after a Kalman-filter track fit. Two different alignment studies are performed. The first one is the magnetic field polarization study that shows no strong evidence of detector positions dependence on magnetic field polarizations. The study result is important since it allows to perform the most accurate measurements of physical observables, especially in verification of the SM predictions. The difference between magnetic field polarizations would cause a discrepancy of the particle reconstruction in the Up or Down polarity that would lead to an additional contribution to the detector and production asymmetry. A second study is focused on the variation of the alignment parameters within one full period of data taking. To investigate the time dependence of the alignment, each time period has been considered within two-four day periods. The study results show sizable variation of alignment parameters within the studied time periods, however, consistent with tracker detector resolutions. No discrepancy is observed between magnetic field polarizations for analyzed data within uncertainties. The variation of the alignment parameters does not have a significant impact on physical observables like resonance masses or their resolutions that are studied using $D^0 \to K^-\pi^+$ and $J/\psi \to \mu^+\mu^-$ data and simulation samples. The comparison between these two samples leads to a conclusion that the J/ψ sample is ill suited for alignment of detectors close to the beam pipe like the IT stations. This originates harder p_T distribution of produced J/ψ which leads to low population of tracks from those decays close to the beam pipe. For both J/ψ and D^0 data samples, a test is made to determine the impact of utilisation of a mass constraint in the alignment procedure, the obtained variation of the extracted mass and its resolution is within 0.15-0.5 MeV/c^2 which does not introduce a significant bias on physical observables.

The second main subject of the thesis is the measurement of the $\mathcal{CP}$-violating phase ϕ_s arising in the interference of $B_s^0 - \bar{B}_s^0$ meson mixing and decay amplitudes. The phase ϕ_s is

precisely predicted within the Standard Model. It is an excellent observable to search for contributions of New Physics beyond the SM. The measurement is performed using the decay channel $B_s^0 \to J/\psi\phi$ with electrons in the final state, reconstructed and selected at the LHCb experiment. The analysis principle relies on a maximum likelihood estimation, fitting the theoretical differential decay rates of the $B_s^0 \to J/\psi\phi$ decay to the experimentally measured decay time and angular distributions.

Using the data taken in the 2011 and 2012 running period of the LHC with a recorded integrated luminosity of $\mathcal{L}_{\text{int}} = 3.0$ fb^{-1}, a set of 12195±497 $B_s^0 \to J/\psi(e^+e^-)\phi(K^+K^-)$ signal candidates is selected. The background pollution of the sample is high due to Bremsstrahlung radiation (∼12% in the B_s^0 signal region). For the flavour tagging to determine the B_s^0 production flavour, two different types of tagging algorithms are used, exploiting the properties of the signal B_s^0 meson decay or the properties of the second produced B-hadron in the event. Combining all inputs from the tagging algorithms, the tagging efficiency is $\varepsilon_{tag} =$ (75.22±1.69)% with the average dilution $\mathcal{D}^2 = 0.0655\pm0.0026$. The overall tagging power is determined to be $\varepsilon_{eff} = (4.93\pm0.16)\%$ representing the statistical reduction of the data sample due to imperfect determination of the B_s^0 production flavour. The decay time resolution determines the capability of resolving the fast B_s^0 meson oscillation. It is extracted simultaneously in the fit to the data sample by taking advantage of its similar behaviour in the $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ sample, which is used as a constraint. The effective average decay time resolution is determined to be σ_t= 44.1 fs. The decay time acceptance correction is taken into account by a data driven method relying on the $B^0 \to J/\psi(e^+e^-)K^*(K^+\pi^-)$ decay mode and the known value of the B^0 lifetime. The detector acceptances influencing the measured angular distributions are determined with simulated events and investigated in detail.

An unbinned maximum likelihood fit is used to determine the physics observables from the decay time and angular distributions of selected $B_s^0 \to J/\psi\phi$ decays with electron final state, taking into account the tagging performance, the corrections to the decay time resolution and acceptance as well as the angular acceptance. The $\Delta\Gamma_s$, Γ_s and Δm_s are Gaussian constrained in the fit to value obtained from the muon mode of the $B_s^0 \to J/\psi\phi$ decay [39]. The measured results for the $B_s^0 - \bar{B}_s^0$ meson mixing parameters using electron mode of the $B_s^0 \to J/\psi\phi$ decays are

$$\begin{aligned} |\lambda| &= 0.83\ ^{+0.084}_{-0.227}\ \pm0.073, \\ \phi_s &= \text{-}0.18\ ^{+0.37}_{-0.39}\ \pm0.25 \text{ rad}, \end{aligned}$$

where the first uncertainty is statistical and the second systematic. The measured values are in an agreement with the prediction from the Standard Model for the $\mathcal{CP}$-violating phase $-2\beta_s = \text{-}0.0376^{+0.0007}_{-0.0008}$ rad [18]. No evidence for new physics affecting the $B_s^0 - \bar{B}_s^0$ mixing phase is found.

In addition, the measurement of the B_s^0 mixing parameters will be performed by combining the information of $B_s^0 \to J/\psi(e^+e^-)\phi$ and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ datasets. This will increase statistical power and expand the study of a possible systematic effects.

In the near future the statistical precision will be improved by analyzing the LHCb dataset collected in 2015-2018 with an integrated luminosity of 5.7 fb^{-1}. Extrapolating the number of selected $B_s^0 \to J/\psi(e^+e^-)\phi$ and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ signal candidates in the analyzed 3.0 fb^{-1} dataset to the combined 8.5 fb^{-1} dataset (Run 1 + Run 2), without taking into account possible improvements in selection and trigger efficiency, the expected statistical uncertainty of ϕ_s is $\sim$0.025 rad.

With the planned upgrade of the LHCb detector in 2021-2029 [121] and the expected integrated luminosity of $\mathcal{L}_{\text{int}} = 50.0$ fb^{-1}, the statistical precision of ϕ_s is extrapolated to be $\sim$0.009 rad [122]. Thanks to more data available and improvements in decay time resolution and acceptance determination, the systematic uncertainty from angular acceptance is expected to decrease. When achieving this precision, it will be important to understand the contribution of the suppressed Standard Model penguin diagrams to the $B_s^0 \to J/\psi(e^+e^-)\phi$ decay amplitude (Fig. 1.5). They contribute to the measured value of ϕ_s and could be misinterpreted as New Physics effects. A reliable theoretical prediction of the effect on ϕ_s is not possible, effects of up to $\sim$-0.1 rad can not be excluded [123]. The contribution of the penguin amplitudes can, however, be experimentally measured by exploiting flavour SU(3) symmetry of the strong interaction and analyzing the decay $B_s^0 \to J/\psi K^*$, where the penguin diagrams are not suppressed compared to the tree level process. The analysis of $B_s^0 \to J/\psi K^*$ decays was already performed at the LHCb [124]. The studies for the $B^0 \to J/\psi K^*$ decay, that is topological similar to $B_s^0 \to J/\psi\phi$, showed that effects in the order of $\sim$-0.01 rad are realistic.

In summary, the aspired precision of the phase ϕ_s will allow to either measure a significant deviation from the Standard Model prediction or put very strong constraints on New Physics scenarios.

References

[1] G. Altarelli, (2005), arXiv:hep-ph/0510281 [hep-ph] .

[2] O. Nachtmann, Braunschweig, F.r. Germany: Vieweg (1986) 474p (1986).

[3] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. **C40**, 100001 (2016).

[4] G. Aad *et al.* (ATLAS), Phys. Lett. **B716**, 1 (2012), arXiv:1207.7214 [hep-ex] .

[5] S. Chatrchyan *et al.* (CMS), Phys. Lett. **B716**, 30 (2012), arXiv:1207.7235 [hep-ex] .

[6] I. C. Brock and T. Schorner-Sadenius, *Physics at the Terascale* (Weinheim, Germany: Wiley-VCH, 476 p, 2011).

[7] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[8] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[9] L.-L. Chau and W.-Y. Keung, Phys. Rev. Lett. **53**, 1802 (1984).

[10] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[11] D. Boutigny *et al.* (BaBar), in *The BABAR physics book: Physics at an asymmetric B factory*, SLAC-R-0504 (1998).

[12] C. S. Wu *et al.*, Phys. Rev. **105**, 1413 (1957).

[13] T. D. Lee and C.-N. Yang, Phys. Rev. **104**, 254 (1956).

[14] R. L. Garwin, L. M. Lederman, and M. Weinrich, Phys. Rev. **105**, 1415 (1957).

[15] J. H. Christenson *et al.*, Phys. Rev. Lett. **13**, 138 (1964).

[16] B. Aubert *et al.* (BaBar), Phys. Rev. **D79**, 072009 (2009), arXiv:0902.1708 [hep-ex] .

[17] I. Adachi *et al.*, Phys. Rev. Lett. **108**, 171802 (2012), arXiv:1201.4643 [hep-ex] .

[18] J. Charles *et al.*, Phys. Rev. **D91**, 073007 (2015), arXiv:1501.05013 [hep-ph] .

[19] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **115**, 031601 (2015), arXiv:1503.07089 [hep-ex] .

[20] Y. Amhis *et al.*, (2016), arXiv:1612.07233 [hep-ex] .

[21] K. Anikeev *et al.* (2001) arXiv:hep-ph/0201071 [hep-ph] .

[22] G. Hiller and U. Uwer, in **Brock, Ian, Schoerner-Sadenius, Thomas (eds.): Physics at the terascale* 163-186* (2011).

[23] G. Luders, Annals Phys. **2**, 1 (1957), [Annals Phys.281,1004(2000)].

[24] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **110**, 221601 (2013), arXiv:1304.6173 [hep-ex] .

[25] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **114**, 041601 (2015), arXiv:1409.8586 [hep-ex] .

[26] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **117**, 061803 (2016), arXiv:1605.09768 [hep-ex] .

[27] M. Artuso, G. Borissov, and A. Lenz, Rev. Mod. Phys. **88**, 045002 (2016), arXiv:1511.09466 [hep-ph] .

[28] R. Aaij *et al.* (LHCb), Phys. Rev. **D87**, 112010 (2013), arXiv:1304.2600 [hep-ex] .

[29] A. S. Dighe *et al.*, Phys. Lett. **B369**, 144 (1996), arXiv:hep-ph/9511363 [hep-ph] .

[30] I. Dunietz, R. Fleischer, and U. Nierste, Phys. Rev. **D63**, 114015 (2001), arXiv:hep-ph/0012219 [hep-ph] .

[31] Y. Xie *et al.*, JHEP **09**, 074 (2009), arXiv:0908.3627 [hep-ph] .

[32] F. Azfar *et al.*, JHEP **11**, 158 (2010), arXiv:1008.4283 [hep-ph] .

[33] R. Harnik *et al.*, Phys. Rev. **D69**, 094024 (2004), arXiv:hep-ph/0212180 [hep-ph] .

[34] A. Lenz *et al.*, Phys. Rev. **D86**, 033008 (2012), arXiv:1203.0238 [hep-ph] .

[35] T. Aaltonen *et al.* (CDF), Phys. Rev. **D85**, 072002 (2012), arXiv:1112.1726 [hep-ex] .

[36] V. M. Abazov *et al.* (D0), Phys. Rev. **D85**, 032006 (2012), arXiv:1109.3166 [hep-ex] .

[37] G. Aad *et al.* (ATLAS), JHEP **08**, 147 (2016), arXiv:1601.03297 [hep-ex] .

[38] V. Khachatryan *et al.* (CMS), Phys. Lett. **B757**, 97 (2016), arXiv:1507.07527 [hep-ex] .

[39] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **114**, 041801 (2015), arXiv:1411.3104 [hep-ex] .

[40] R. Aaij *et al.* (LHCb), JHEP **08**, 037 (2017), arXiv:1704.08217 [hep-ex] .

[41] R. Aaij *et al.* (LHCb), Phys. Lett. **B736**, 186 (2014), arXiv:1405.4140 [hep-ex] .

[42] R. Aaij *et al.* (LHCb), Phys. Lett. **B762**, 253 (2016), arXiv:1608.04855 [hep-ex] .

[43] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **113**, 211801 (2014), arXiv:1409.4619 [hep-ex] .

[44] L. Evans and P. Bryant, JINST **3**, S08001 (2008).

[45] A. Team, "The four main LHC experiments," (1999).

[46] J. Nardulli, Ph.D. thesis, CERN-THESIS-2007-063 (2007).

[47] C. Elsässer (LHCb collaboration), "$\bar{b}b$ production angle plots," .

[48] R. Aaij *et al.* (LHCb), Eur. Phys. J. **C71**, 1645 (2011), arXiv:1103.0423 [hep-ex] .

[49] R. Aaij *et al.* (LHCb), JHEP **06**, 064 (2013), arXiv:1304.6977 [hep-ex] .

[50] A. A. Alves, Jr. *et al.* (LHCb), JINST **3**, S08005 (2008).

[51] LHCb Collaboration, Tech. Rep. CERN-LHCC-2003-030 (LHCb, 2003).

[52] "LHCb Silicon Tracker," http://lhcb.physik.uzh.ch/ST/public/material/index.php.

[53] R. Forty, Nucl. Instrum. Meth. **A384**, 167 (1996).

[54] F. Archilli *et al.*, JINST **8**, P10020 (2013), arXiv:1306.0249 .

[55] A. Puig, Tech. Rep. LHCb-PUB-2014-046 (CERN, 2014).

[56] "Brunel project," http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/.

[57] "DaVinci project," http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/.

[58] M. Clemencic *et al.*, J. Phys. Conf. Ser. **219**, 042006 (2010).

[59] A. Sciaba *et al.* (CERN), "GLITE 3.2 USER GUIDE MANUALS SERIES," .

[60] "Dirac project," http://lhcbdoc.web.cern.ch/lhcbdoc/dirac/.

[61] I. Belyaev *et al.* (LHCb), J. Phys. Conf. Ser. **331**, 032047 (2011).

[62] M. Clemencic and others. (LHCb), J. Phys. Conf. Ser. **331**, 032023 (2011).

[63] T. Sjöstrand, S. Mrenna, and P. Z. Skands, JHEP **05**, 026 (2006), arXiv:hep-ph/0603175 [hep-ph] .

[64] T. Sjöstrand *et al.*, Comput. Phys. Comm. **191**, 159 (2015), arXiv:1410.3012 [hep-ph] .

[65] D. J. Lange, Nucl. Instrum. Meth. **A462**, 152 (2001).

[66] P. Golonka and Z. Was, Eur. Phys. J. **C45**, 97 (2006), arXiv:hep-ph/0506026 [hep-ph] .

[67] S. Agostinelli *et al.* (GEANT4), Nucl. Instrum. Meth. **A506**, 250 (2003).

[68] "Boole project," http://lhcbdoc.web.cern.ch/lhcbdoc/boole/.

[69] "Moore project," http://lhcbdoc.web.cern.ch/lhcbdoc/moore/.

[70] "Alignment project," http://lhcbdoc.web.cern.ch/lhcbdoc/alignment/.

[71] T. Lastoviicka, CERN-LHCb-2007-002 (2008).

[72] R. Forty and M. Needham, CERN-LHCB-2007-023 (2007).

[73] O. Callot and S. Hansmann-Menzemer, CERN-LHCb-2007-015 (2007).

[74] M. Needham, CERN-LHCB-2007-129 (2007).

[75] E. Rodrigues, CERN-LHCb-2006-057 (2006).

[76] R. Fruhwirth, Nucl. Instrum. Meth. **A262**, 444 (1987).

[77] W. Baldini *et al.*, Tech. Rep. LHCb-2006-035 (CERN, 2006).

[78] S. Viret, C. Parkes, and D. Petrie, Tech. Rep. LHCb-2005-101 (CERN, 2005).

[79] S. Viret, C. Parkes, and M. Gersabeck, Tech. Rep. LHCb-2007-067 (CERN, 2007).

[80] M. Gersabeck, C. Parkes, and S. Viret, (2008), arXiv:0807.3532 .

[81] V. Blobel, Nucl. Instrum. Meth. **A566**, 5 (2006).

[82] G. Conti and F. Blanc, CERN-LHCB-2008-069 (2008).

[83] M. Needham, Tech. Rep. LHCb-2009-030. LPHE-Note-2009-02 (CERN, 2009).

[84] V. Batozskaya and K. Klimaszewski, Tech. Rep. LHCb-INT-2017-033 (CERN, 2017).

[85] "Sim08 Generator Statistics," http://lhcbdoc.web.cern.ch/lhcbdoc/STATISTICS/SIM08STAT/index.shtml.

[86] R. Aaij *et al.*, LHCb-CONF-2011-001 (2011).

[87] W. D. Hulsbergen, Nucl. Instrum. Meth. **A552**, 566 (2005), arXiv:physics/0503191 .

[88] A. Poluektov, Journal of Instrumentation **10**, P02011 (2015).

[89] "Urania project," https://gitlab.cern.ch/lhcb/Urania.

[90] B. P. Roe *et al.*, Nucl. Instrum. Meth. **A543**, 577 (2005), arXiv:physics/0408124 .

[91] D. J. Amit, K. Y. M. Wong, and C. Campbell, J. Phys. **A22**, 2039 (1989).

[92] A. Hoecker *et al.*, PoS **ACAT**, 040 (2007), arXiv:physics/0703039 .

[93] Y. Xie, (2009), arXiv:0901.3305 [hep-ex] .

[94] M. Pivk and F. R. Le Diberder, Nucl.Instrum.Meth. **A555**, 356 (2005), arXiv:physics/0402083 .

[95] J. Schmidhuber, Neural Networks **61**, 85 (2015).

[96] "NetMaker project," http://www.ire.pw.edu.pl/~rsulej/NetMaker/.

[97] W. A. Little and G. L. Shaw, Math. Biosci. **39**, 281 (1978).

[98] C. F. Gauss, *Theoria Motus Corporum Celestium* (1809) translated as *Theory of Motion of the Heavenly Bodies Moving about the Sun in Conic Sections* (trans. C. H. Davis), Boston, Little, Brown 1857. Reprinted: New York, Dover 1963.

[99] G. Breit and E. Wigner, Phys. Rev. **49**, 519 (1936).

[100] R. Aaij *et al.* (LHCb), Chin. Phys. **C40**, 011001 (2016), arXiv:1509.00292 [hep-ex] .

[101] J. E. Gaiser, Ph.D. thesis, SLAC (1982).

[102] J. Glaisher, Philosophical Magazine Series 4 **42** (1871).

[103] W. Heitler, *The Quantum Theory of Radiation*, Monographs on Physics (Oxford University Press, 1954).

[104] T. du Pree, M. H. M. Merk, and H. G. Raven, Ph.D. thesis, CERN-THESIS-2010-124 (2010).

[105] J. Olivero and R. Longbothum, Journal of Quantitative Spectroscopy and Radiative Transfer **17**, 233 (1977).

[106] H. G. Moser and A. Roussarie, Nucl. Instrum. Meth. **A384**, 491 (1997).

[107] M. Calvi, O. Leroy, and M. Musy, Tech. Rep. LHCb-2007-058 (CERN, 2007).

[108] "LHCb Flavour Tagging," https://twiki.cern.ch/twiki/bin/view/LHCbPhysics/FTAlgorithmsInfo.

[109] R. Aaij *et al.* (LHCb), Eur. Phys. J. **C72**, 2022 (2012), arXiv:1202.4979 [hep-ex] .

[110] G. Krocker *et al.*, LHCb-ANA-2011-103 (2011).

[111] M. Dorigo *et al.* (LHCb), https://indico.cern.ch/event/257864/session/0/contribution/37/material/slides/0.pdf.

[112] M. Dorigo *et al.* (LHCb), https://twiki.cern.ch/twiki/bin/viewauth/LHCbPhysics/FlavourTagging4Reco14.

[113] LHCB-ROADMAP3-001 (2009).

[114] G. Punzi, eConf **C030908**, WELT002 (2003), arXiv:physics/0401045 .

[115] F. Porter, Nucl. Instrum. Meth. **A368**, 793 (1996).

[116] B. A. Murtagh and M. A. Saunders, *MINOS 5.51 User's Guide*, Tech. Rep. SOL 83-20R (Stanford University, 2013).

[117] D. Martínez Santos and F. Dupertuis, Nucl. Instrum. Meth. **A764**, 150 (2014), arXiv:1312.5000 [hep-ex] .

[118] R. Aaij *et al.* (LHCb), New J. Phys. **15**, 053021 (2013), arXiv:1304.4741 [hep-ex] .

[119] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **111**, 181801 (2013), arXiv:1308.4544 [hep-ex] .

[120] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **108**, 101803 (2012), arXiv:1112.3183 [hep-ex] .

[121] I. Bediaga *et al.*, Tech. Rep. CERN-LHCC-2012-007 (LHCb collaboration, 2012).

[122] LHCb Collaboration, Tech. Rep. LHCb-PUB-2014-040. LHCb-INT-2013-024 (CERN, 2014).

[123] S. Faller, R. Fleischer, and T. Mannel, Phys. Rev. D **79**, 014005. 28 p (2008).

[124] R. Aaij *et al.* (LHCb), JHEP **11**, 082 (2015), arXiv:1509.00400 [hep-ex] .

[125] L. V. Kantorovich, Uspekhi Mat. Nauk **3**, 89 (1948).

[126] "SIMCOND project," https://gitlab.cern.ch/lhcb-conddb/SIMCOND.

[127] S. van der Meer, Tech. Rep. ISR-PO-68-31 (CERN, 1968).

[128] L. B. Smith, Commun. ACM **9**, 86 (1966).

[129] J. Wimberly, "EspressoPerformanceMonitor project," https://gitlab.cern.ch/lhcb-ft/EspressoPerformanceMonitor.

Appendix A

Mixing phenomenology

The time dependent decay rate for the decay of a produced B_q^0 to the final state $\bar{f}$ is

$$\frac{\mathrm{d}\Gamma(B_q^0 \to \bar{f})}{\mathrm{d}t\mathcal{N}_{\bar{f}}} = |\langle \bar{f}|B_q^0(t)\rangle|^2 = \left| g_+(t)A_{\bar{f}} + \frac{q}{p} g_-(t)\bar{A}_{\bar{f}} \right|^2$$

$$= \frac{1}{2}|A_{\bar{f}}|^2 e^{-\Gamma t}\left[(1+|\lambda_{\bar{f}}|^2)\cosh\frac{\Delta\Gamma}{2}t + (1-|\lambda_{\bar{f}}|^2)\cos\Delta m t - 2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\Re(\lambda_{\bar{f}}) - 2\sin(\Delta m t)\Im(\lambda_{\bar{f}})\right]. \tag{A.1}$$

Similarly the decay rate for $\bar{B}_q^0$ to $\bar{f}$ is given by

$$\frac{\mathrm{d}\Gamma(\bar{B}_q^0 \to \bar{f})}{\mathrm{d}t\mathcal{N}_{\bar{f}}} = |\langle \bar{f}|\bar{B}_q^0(t)\rangle|^2 = \left| \frac{p}{q} g_-(t)A_{\bar{f}} + g_+(t)\bar{A}_{\bar{f}} \right|^2$$

$$= \frac{1}{2}\left|\frac{p}{q}\right|^2 |A_{\bar{f}}|^2 e^{-\Gamma t}\left[(1+|\lambda_{\bar{f}}|^2)\cosh\frac{\Delta\Gamma}{2}t - (1-|\lambda_{\bar{f}}|^2)\cos\Delta m t - 2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\Re(\lambda_{\bar{f}}) + 2\sin(\Delta m t)\Im(\lambda_{\bar{f}})\right] \tag{A.2}$$

where $\mathcal{N}_{\bar{f}}$ is a normalization factor. The decay rates to final state $\bar{f}$ are the $\mathcal{CP}$ conjugated decay rates to final state f by Eqs. 1.23-1.24 where substituting $A_f \to A_{\bar{f}}, \bar{A}_f \to \bar{A}_{\bar{f}}$ and $\lambda_f \to \lambda_{\bar{f}} = \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_{\bar{f}}}$.

Appendix B

Alignment of the Tracking Stations

B.1 Kalman Filter Track Fit

The Kalman filter track fit proceeds in three steps: prediction, filter and smoothing.

Prediction

At the start of the fit, a first estimate of the track state $\vec{x}_0$ is needed. This is provided by a first fit in the track finding procedure. An initial estimate of the covariance matrix C_0 is also given with increased values for the errors on the diagonal, in order not to be biased by the estimate of $\vec{x}_0$. Starting from this, the prediction relation for the track state and the covariance matrix are

$$\vec{x}_k^{k-1} = F_k \vec{x}_{k-1} + \vec{w}_k, \tag{B.1}$$

$$C_k^{k-1} = F_k C_{k-1} F_k^T + Q_k, \tag{B.2}$$

where F_k is called the transport matrix. The $\vec{w}_k$ and matrix Q_k are the process-noise terms coming from the multiple scattering, which has the effect of increasing the predicted error on the state vector $\vec{x}_k^{k-1}$. The superscript indicates how many measurements have been incorporated in the determination of $\vec{x}_k$: $k-1$ means a predicted state, k is a filtered state.

From this predicted state, a prediction of the residual r and its covariance R can be made, by defining the residual as the distance between the measurement m and the state vector $h(\vec{x})$ projected in the measurement plane:

$$r_k^{k-1} = m_k - h_k(\vec{x}_k^{k-1}), \tag{B.3}$$

$$R_k^{k-1} = V_k + H_k C_k^{k-1} H_k^T, \tag{B.4}$$

where H_k is the measurement matrix and V_k the measurement variance. The aim of the Kalman fit is to minimize the predicted contributions

$$(\chi_+^2)_k^{k-1} = r_k^{k-1} (R_k^{k-1})^{-1} r_k^{k-1} \tag{B.5}$$

of the current measurement to the total χ^2 with respect to the track parameters.

Filter

The filter step adds the information of the current measurement k to the predicted state based on the $k-1$ first measurements added to the track. The filtered state vector and its covariance matrix are obtained by the filter equation:

$$\vec{x}_k = \vec{x}_k^{k-1} + K_k r_k^{k-1}, \tag{B.6}$$

$$C_k = (\mathbb{1} - K_k H_k) C_k^{k-1}, \tag{B.7}$$

where K_k is a 5×1 gain matrix given by

$$K_k = C_k^{k-1} H_k^T (V_k + H_k C_k^{k-1} H_k^T)^{-1} = C_k^{k-1} H_k^T (R_k^{k-1})^{-1}. \tag{B.8}$$

The complete derivation of this Kalman gain matrix is given in Ref. [76]. This also leads to the expression of the filtered residual and its covariance matrix:

$$r_k^{k-1} = m_k - h_k(\vec{x}_k) = (1 - H_k K_k) r_k^{k-1}, \tag{B.9}$$

$$R_k = (1 - H_k K_k) V_k = V_k - H_k C_k H_k^T. \tag{B.10}$$

The contribution of this measurement to the total χ^2 becomes

$$(\chi_+^2)_k = r_k R_k^{-1} r_k. \tag{B.11}$$

Smoothing

After the best estimate of the state vector $\vec{x}_n$ is obtained at the last measurement by predicting and filtering each of the previous states, the procedure is reversed in order to propagate this full information back to all the previous states. The smoothed state vector and covariance matrix become

$$\vec{x}_k^n = \vec{x}_k + A_k(\vec{x}_{k+1}^n - \vec{x}_{k+1}^k), \tag{B.12}$$

$$C_k^n = C_k + A_k(C_{k+1}^n - C_{k+1}^k) A_k^T, \tag{B.13}$$

where

$$A_k = C_k F_{k+1}^T (C_{k+1}^k)^{-1} \tag{B.14}$$

is the 5×5 smoother gain matrix. The smoothed residual and corresponding covariance matrix finally become

$$r_k^n = m_k - h_k(\vec{x}_k^n), \tag{B.15}$$

$$R_k^n = V_k - H_k C_k^n H_k^T \tag{B.16}$$

and the total χ^2 to minimize in the fitting procedure is a sum of all smoothed contributions:

$$(\chi^2_+)^n_k = r^n_k (R^n_k)^{-1} r^n_k. \tag{B.17}$$

χ^2 minimization

The fitting procedure consists of finding the best track parameters by minimizing a sum of the χ^2_+ given in Eq. B.17 and which add up to a total χ^2 of the form

$$\chi^2 = r^T V^{-1} r. \tag{B.18}$$

The minimization is expressed by

$$\frac{d\chi^2}{dx} \equiv 0. \tag{B.19}$$

Eq. B.19 can be rewritten by using an initial estimate x_0 of the track parameters and a linear expansion of the measurement model around it:

$$h(x) = h(x_0) + H(x - x_0),$$

where $H = \frac{\partial h(x)}{\partial x}\Big|_{x_0}$ is the projection matrix. Then Eq. B.19 becomes

$$0 \equiv \frac{d\chi^2}{dx}\Big|_{x_0} = -2H^T V^{-1}[m - h(x_0) - H(x - x_0)]. \tag{B.20}$$

The solution of Eq. B.20 is

$$x = x_0 - CH^T V[m - h(x_0)], \tag{B.21}$$

where the covariance matrix of the track parameters is given by

$$C = 2\left(\frac{d^2\chi^2}{dx^2}\Big|_{x_0}\right)^{-1} = 2\left(2H^T V^{-1}\frac{dh(x)}{dx}\Big|_{x_0}\right)^{-1} = (H^T V^{-1} H)^{-1}. \tag{B.22}$$

If H depends on x (non-linearity), iterations are needed in order to reach some predefined convergence criterion which is defined by a minimal change in the track χ^2. The numerical method of successive approximations of real zeros of a real function is called the Newton-Raphson method [125]. In this case, Eq. B.21 becomes

$$x = x_0 - \left(\frac{d^2\chi^2}{dx^2}\Big|_{x_0}\right)^{-1} \frac{d\chi^2}{dx}\Big|_{x_0}. \tag{B.23}$$

B.2 Minimum χ^2 Formalism for Alignment

The minimization of a sum of the track χ^2 with respect to α, the dependence of x on α being taken into account through the total derivative, is given by:

$$\frac{d}{d\alpha} = \frac{\partial}{\partial\alpha} + \frac{dx}{d\alpha}\frac{\partial}{\partial x}, \tag{B.24}$$

where

$$\frac{dx}{d\alpha} = -\frac{\partial^2\chi^2}{\partial\alpha\partial x}\left(\frac{\partial^2\chi^2}{\partial x^2}\right)^{-1} \tag{B.25}$$

which is a consequence of the minimization of the χ^2 with respect to the track parameters x:

$$\frac{d}{d\alpha}\frac{\partial\chi^2}{\partial x} = 0. \tag{B.26}$$

Using a first order linearization of the residual $r = m - h(x,\alpha)$ around the expansion point $(x(\alpha_0),\alpha_0)$ and rewriting its derivative with respect to α as

$$A_{ij} \equiv \frac{\partial r_i}{\partial\alpha_j} \tag{B.27}$$

the derivative of the track parameters with respect to α given in Eq. B.25 becomes

$$\frac{dx}{d\alpha} = -(-2A^TV^{-1}H)(2H^TV^{-1}H)^{-1} = A^TV^{-1}HC.$$

Using this result, the total derivative given in Eq. B.24 can be written as

$$\frac{d}{d\alpha} = \frac{\partial}{\partial\alpha} + A^TV^{-1}HC\frac{\partial}{\partial x}. \tag{B.28}$$

The first and second order derivatives of the track χ^2 used in Eq. 3.5 become

$$\frac{d\chi^2}{d\alpha} = \frac{\partial\chi^2}{\partial\alpha} + A^TV^{-1}HC\frac{\partial\chi^2}{\partial x} = 2A^TV^{-1}(V - HCH^T)V^{-1}r \tag{B.29}$$

and

$$\frac{d^2\chi^2}{d\alpha^2} = \frac{d}{d\alpha}\left(\frac{d\chi^2}{d\alpha}\right) = 2A^TV^{-1}(V - HCH^T)V^{-1}A. \tag{B.30}$$

The difference with respect to the equations derived for the track χ^2 minimization (Eqs. B.20 and B.22) is the addition of the term HCH^T, which is the covariance matrix of the track parameters in the measurement space. The residual covariance matrix $R = V - HCH^{-1}$ is equivalent to that discussed for the Kalman filter in Sec. 3.1.3.

In the case where the track parameters are the best estimate for the given alignment

parameters (i.e. x satisfies Eq. B.20), Eq. B.29 simplifies to

$$\frac{d\chi^2}{d\alpha} = 2A^T V^{-1} r. \tag{B.31}$$

Hence, the equations to solve in order to extract the alignment parameters are Eqs. B.30 and B.31.

B.3 Alignment Studies of the Tracking Stations

B.3.1 Event selection

The alignment study is performed using the D^0 and J/ψ sample where the candidates for a D^0 meson are reconstructed from K^- and π^+ particles while J/ψ candidates are formed from muon pairs of opposite sign tracks. The D^0 meson is required to have the log-likelihood difference of particle identification between the kaon and pion hypothesis larger than 0 and between the pion and kaon hypothesis smaller than 0 (Sec. 2.3.2) and a mass within 20 MeV/c^2 of the nominal D^0 mass [3]. The J/ψ candidate is required that the $\Delta\ln\mathcal{L}_{\mu K}$ is larger than 5 while the $\Delta\ln\mathcal{L}_{\mu e}$ is larger than 8 and a mass within 35 MeV/c^2 of the nominal J/ψ mass [3]. The full list of applied selection criteria is reported in Table B.1.

Table B.1 Particle selection criteria for $D^0 \to K^-\pi^+$ and $J/\psi \to \mu^+\mu^-$ candidates.

Decay mode	Cut parameter	Value
$D^0 \to K\pi$	$\Delta\ln\mathcal{L}_{K\pi}$	$> 0.$
	$\Delta\ln\mathcal{L}_{\pi\pi}$	$< 0.$
	χ^2_{vtx}	$< 9.$
	$\lvert m(K\pi) - m(D^0)\rvert$	$< 20.$ MeV/c^2
$J/\psi \to \mu\mu$	$\Delta\ln\mathcal{L}_{\mu K}$	$> 5.$
	$\Delta\ln\mathcal{L}_{\mu e}$	$> 8.$
	χ^2_{vtx}	$< 7.$
	$\lvert m(\mu\mu) - m(J/\psi)\rvert$	$< 35.$ MeV/c^2

In case of a data sample, the number of analysed events is selected such that the number of vertices for alignment is larger than $2{\cdot}10^4$. The numbers of selected events, vertices and tracks used in the alignment for each year in two polarizations are given in Table B.2.

The simulated sample is generated 2011 run conditions with a centre-of-mass energy of $\sqrt{s}$ =7 TeV [84]. The sample contains an equal amount of events for both magnet polarities (Table B.3).

Table B.2 Number of events, vertices and tracks in a data sample.

Polarization	N_{events}	N_{vertex}	N_{track}
2011, $D^0 \to K\pi$			
Down	$2{\cdot}10^6$	31814	63739
Up	$2{\cdot}10^6$	31868	63859
2012, $D^0 \to K\pi$			
Down	$3{\cdot}10^5$	24504	49150
Up	$2{\cdot}10^6$	28251	56640
2011, $J/\psi \to \mu\mu$			
Down	10^4	29345	59204

Table B.3 Number of events, vertices and tracks in 2011 MC sample.

Polarization	N_{events}	N_{vertex}	N_{track}
Down	$1.5{\cdot}10^6$	10306	20652
Up	$1.5{\cdot}10^6$	10253	20548

B.3.2 Alignment convergence study

2011 data

Two time periods of the data with two magnet field polarities are considered for the convergence study. The fill numbers, the time periods in which the data is stored, are shown in Table B.4.

Table B.4 Time periods investigated in the convergence study with 2011 data sample.

Period	Polarization	Fill number
28.04 - 10.06	Down	1856
10.06 - 06.07	Up	1889

The shift is shown in Figs. B.1-B.2 where black triangles are the official values taken from DB, blue crosses are values with artificial misalignment, red circles and green boxes are values after first and second iteration, respectively. The average value for mean and root mean square (RMS) of residual distributions is presented in Table B.5.

Table B.5 Average mean and RMS value of the residual distributions for IT1Boxes with 2011 data sample.

	official	Iter0	Iter1	Iter2
Down polarization				
Mean	0.00287	0.05205	0.01643	0.00154
RMS	0.06295	0.06716	0.06366	0.06289
Up polarization				
Mean	-0.00282	0.05105	0.01405	0.00015
RMS	0.06211	0.0651	0.06252	0.06221

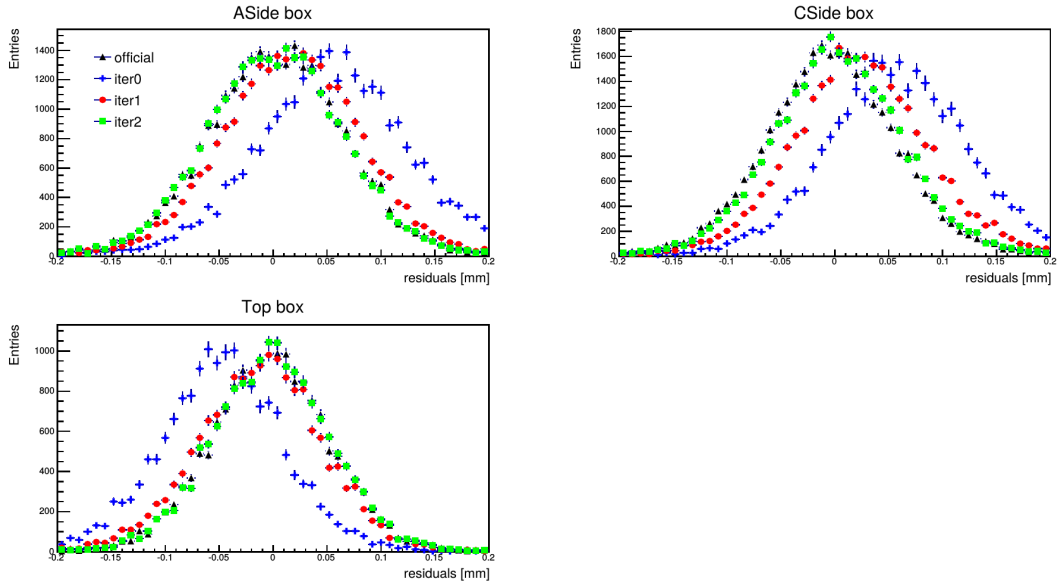


Fig. B.1 Residual distributions with 2011 data sample for Down polarization.

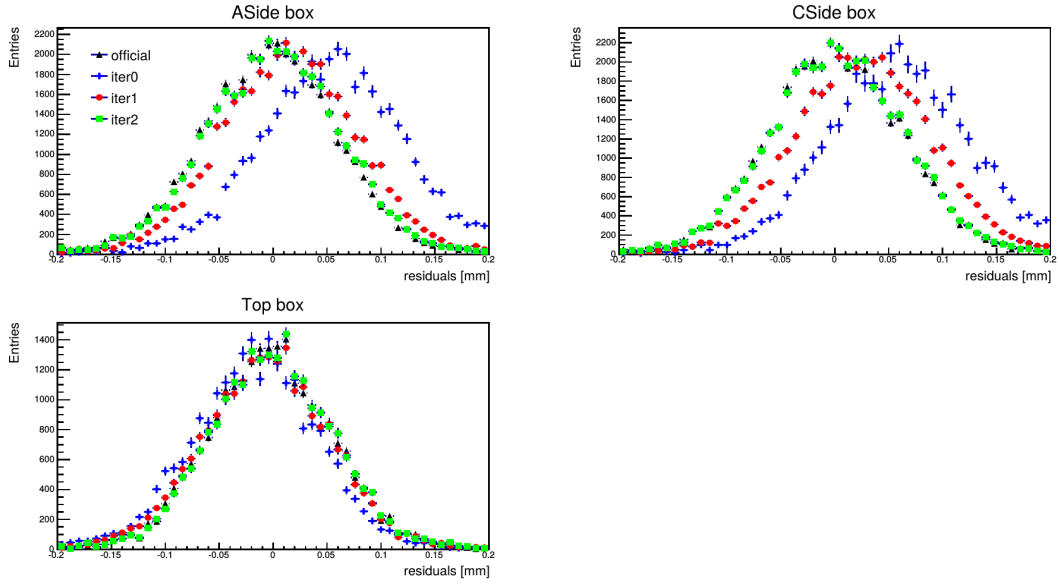


Fig. B.2 Residual distributions with 2011 data sample for Up polarization.

The results of the Tx alignment in the local reference frame of IT1Boxes are reported in Table B.6 where the alignment parameters are obtained for three scenarios: before iterations (official), artificial shift (start) and after alignment (after 3 iterations). The difference (Δ) of local parameters between each iteration and artificial misalignment is also included in Table B.6.

The study result shows that the alignment with 3 iterations has converged since the difference between 2 and 3 iteration parameters is very small. The procedure converges for both offsets of 100 and 200 μm. The aligned track χ^2 is consistent with track χ^2 for

Table B.6 Local Tx alignment parameters and the average track χ^2 for IT1Boxes with 2011 data sample.

IT/Station1	ASide [mm]	Bottom [mm]	CSide [mm]	Top [mm]	$\langle\chi^2_{\text{track}}\rangle$
Down polarization					
official	-2.224	-2.236	1.349	-1.673	35.829
start	-2.424	-2.136	1.549	-1.773	37.045
Δiter0	0.219±0.007	-0.050±0.009	-0.207±0.008	0.085±0.008	
Δiter1	0.044±0.005	0.012±0.007	-0.026±0.006	-0.014±0.007	
Δiter2	0.005±0.005	0.003±0.007	0.004±0.006	-0.003±0.007	
after 3 iterations	-2.156	-2.169	1.317	-1.707	35.802
Δ to official	0.068	0.067	-0.03	-0.034	
Up polarization					
official	-2.16	-2.189	1.338	-1.658	35.407
start	-2.357	-2.167	1.538	-1.663	36.358
Δiter0	0.133±0.007	-0.047±0.008	-0.180±0.008	-0.014±0.008	
Δiter1	0.043±0.005	0.004±0.007	-0.035±0.006	-0.022±0.007	
Δiter2	-0.003±0.005	-0.002±0.007	0.005±0.006	-0.001±0.007	
after 3 iterations	-2.184	-2.21	1.325	-1.702	35.423
Δ to official	-0.024	-0.02	-0.013	-0.04	
Δ Up/Down official	0.067	0.047	0.011	0.015	
Δ Up/Down Iter2	0.028	0.04	0.008	0.005	

official values. Furthermore, the residual distribution has proper shape. The small difference between realigned values and the official ones can be attributed to the inherent precision of the procedure which is studied on simulation sample discussed below.

In addition, the difference in detector position between data samples with opposite magnet polarizations is verified. This difference is reduced using three iterations where the fastest movement to official positions is observed in ASidebox of IT1 station. The discrepancy between Up and Down polarizations is observed for both official and for aligned parameter values. This result indicates that a part of observed differences in the opposite magnet fields could originate from intrinsic resolution of the alignment procedure. The time dependence of alignment result is investigated to verify this assumption (Sec. 3.4.3).

2011 MC simulation

The same study of the procedure resolution verification is performed with MC sample. The residual distributions for artificially shifted and realigned conditions are shown in Figs. B.3-B.4 where the same colour definition is used for 2011 data sample. The typical value for the average mean and RMS of the residual distributions is reported in Table B.7.

In order to obtain meaningful results, it is crucial to use proper survey constraints for all detector stations. For simulation the exact positions can be used. The survey constraints are extracted from `SIMCOND`, special condition database for simulation samples [126].

The IT boxes positions are reproduced within 20 μm while the difference between two magnet polarities is up to 25 μm. The result for 2011 simulation sample is in agreement with 2011 data result as reported in Table B.8.

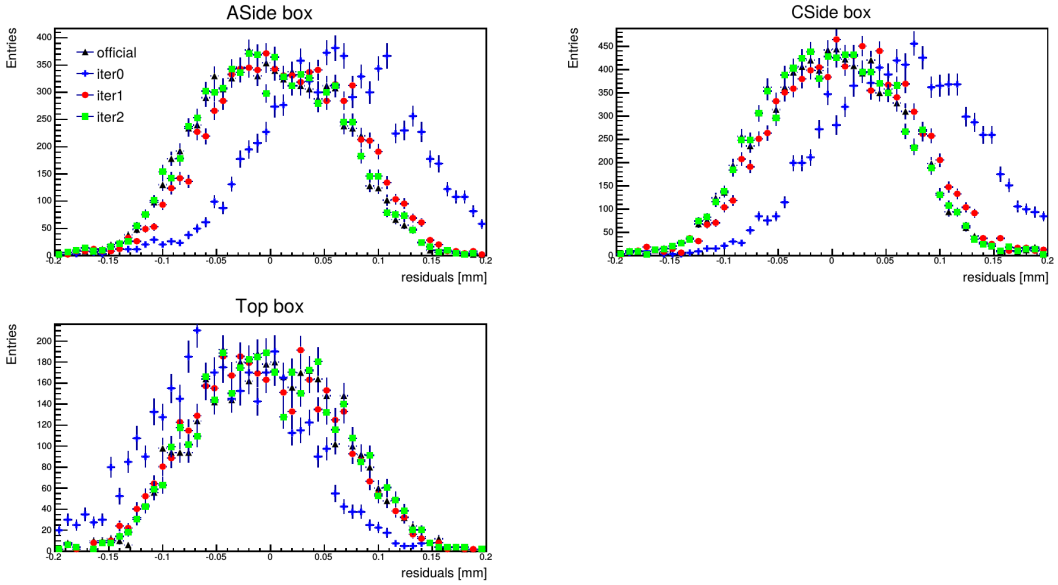


Fig. B.3 Residual distributions with 2011 MC sample for Down polarization.

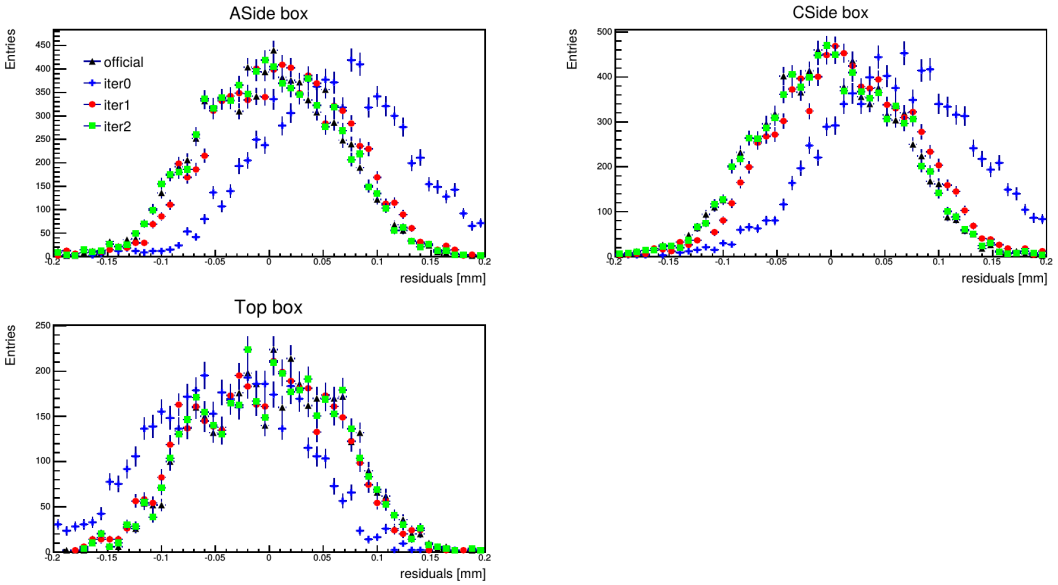


Fig. B.4 Residual distributions with 2011 MC sample for Up polarization.

2012 data

The alignment convergence study is also verified with 2012 data sample with six time periods: two groups of consecutive time periods with opposite magnetic field (Table B.9). For the first group the official alignment is resulted in a large difference (up to 200 μm) for the Tx alignment parameter. The second group exhibits small difference ($\sim$50 μm) in the same degree of freedom.

The result of the alignment study is as follows. In case of the first group, the difference

Table B.7 Average mean and RMS value of the residual distributions for IT1Boxes with 2011 MC sample.

	official	Iter0	Iter1	Iter2
Down polarization				
Mean	-0.00136	0.06017	0.0115	-0.00055
RMS	0.06308	0.06548	0.06318	0.06309
Up polarization				
Mean	-0.00018	0.05831	0.01014	-0.00040
RMS	0.06178	0.06482	0.06188	0.06179

Table B.8 Local Tx alignment parameters, their values after realignment with modified survey constraints, differences between magnet polarities and $\langle\chi^2_{\text{track}}\rangle$ for IT1Boxes with 2011 MC sample.

IT/Station1	ASide [mm]	Bottom [mm]	CSide [mm]	Top [mm]	$\langle\chi^2_{\text{track}}\rangle$
Down polarization					
official	0	0	0	0	34.881
start	-0.2	0.1	0.2	-0.1	35.766
after 3 iterations	-0.007	0.015	-0.009	0.016	34.894
Δ to official	-0.007	0.015	-0.009	0.016	
Up polarization					
official	0	0	0	0	34.549
start	-0.2	0.1	0.2	-0.1	35.517
after 3 iterations	0.016	-0.011	-0.018	-0.007	34.581
Δ to official	0.016	-0.011	-0.018	-0.007	
Δ Up/Down official	0	0	0	0	
Δ Up/Down Iter2	0.023	0.026	0.009	0.023	

Table B.9 Time periods investigated in the convergence study with 2012 data sample.

Period	Polarization	Fill number
Group 1		
23.10 - 07.11	Up	3236
07.11 - 03.12	Down	3300
Group 2		
01.07 - 19.07	Up	2848
19.07 - 24.07	Up	2858
24.07 - 09.08	Down	2880
09.08 - 27.08	Up	2976

between official and realigned values after three alignment iterations is below 60 μm that is in line with result for 2011 data sample. The track χ^2 after alignment is in agreement with official track χ^2 value. The difference of alignment parameters between two magnet polarities is also calculated. For the ASide or Bottom boxes of IT1 station this discrepancy is decreased in comparison with the official values while a slight increase of the difference is observed for Top and CSide boxes. The difference in polarizations for all boxes follows a similar trend throughout the whole year (Fig. B.5). However, only for this pair of the time period (group

1) the difference in CSide and Top boxes of IT1 station is inverted (Table B.10). As it could indicate a real effect in the alignment procedure, an additional pair of the time periods is studied.

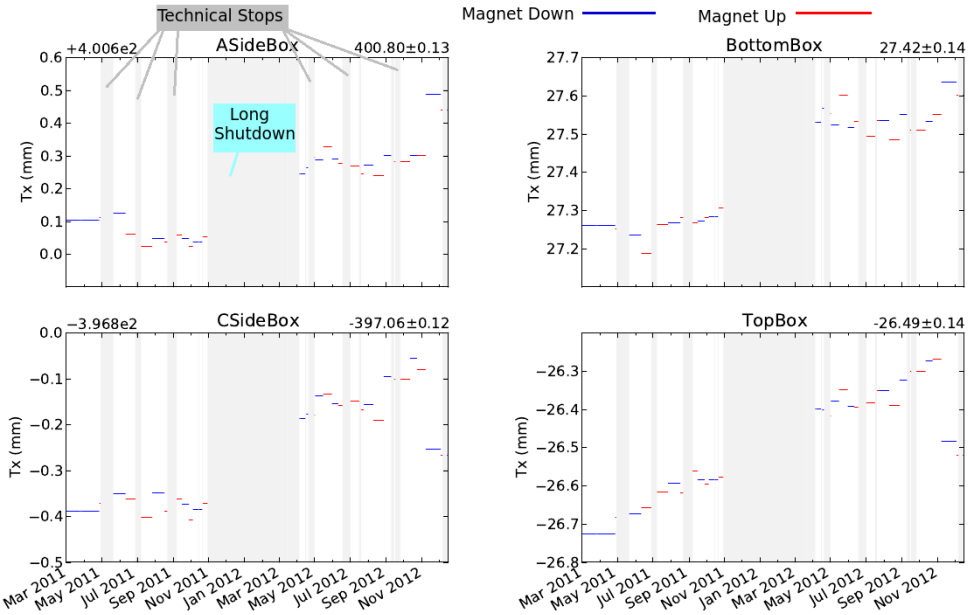


Fig. B.5 Official alignment positions of IT1Boxes as a time function where a blue line is a time period with Down magnet polarization, a red line shows a time period with Up magnet polarization.

Table B.10 Local Tx alignment parameters, their values after realignment, differences between magnet polarities and $\langle\chi^2_{\text{track}}\rangle$ for IT1Boxes with 2012 data sample (Group 1).

IT/Station1	ASide [mm]	Bottom [mm]	CSide [mm]	Top [mm]	$\langle\chi^2_{\text{track}}\rangle$
Down polarization					
official	-2.589	-2.637	1.446	-1.484	36.547
start	-2.789	-2.537	1.646	-1.584	37.317
after 3 iterations	-2.544	-2.578	1.436	-1.499	36.618
Δ to official	0.045	0.059	-0.01	-0.015	
Up polarization					
official	-2.402	-2.552	1.62	-1.268	36.534
start	-2.602	-2.452	1.82	-1.368	37.248
after 3 iterations	-2.391	-2.552	1.645	-1.222	36.634
Δ to official	0.011	0.0	0.025	0.046	
Δ Up/Down official	0.187	0.085	0.174	0.216	
Δ Up/Down Iter2	0.153	0.026	0.209	0.277	

The alignment procedure for the second group of 2012 data sample is converged after 3 iterations as performed for the first group of 2012 and 2011 data samples. The difference between consecutive time periods of the official and realigned parameters is shown in Table B.11. The variation between time periods after realignment is reduced in comparison with official values and is $\sim$40 μm.

Table B.11 Difference of the Tx alignment parameters in the official DB and after convergence study between magnet polarities for IT1Boxes with 2012 data sample (Group 2).

IT/Station1	ASide [mm]	Bottom [mm]	CSide [mm]	Top [mm]
Δ Up(Jul)/Up(Aug) official	0.029	0.007	0.046	0.007
Δ Up(Jul)/Up(Aug) Iter2	0.003	0.025	0.005	0.04
Δ Up(Jul)/Down official	0.001	0.041	0.007	0.032
Δ Up(Jul)/Down Iter2	0.004	0.001	0.017	0.019
Δ Down/Up(Aug) official	0.03	0.048	0.034	0.039
Δ Down/Up(Aug) Iter2	0.001	0.024	0.012	0.021

B.3.3 Time dependence study of alignment parameters

Survey constraints study

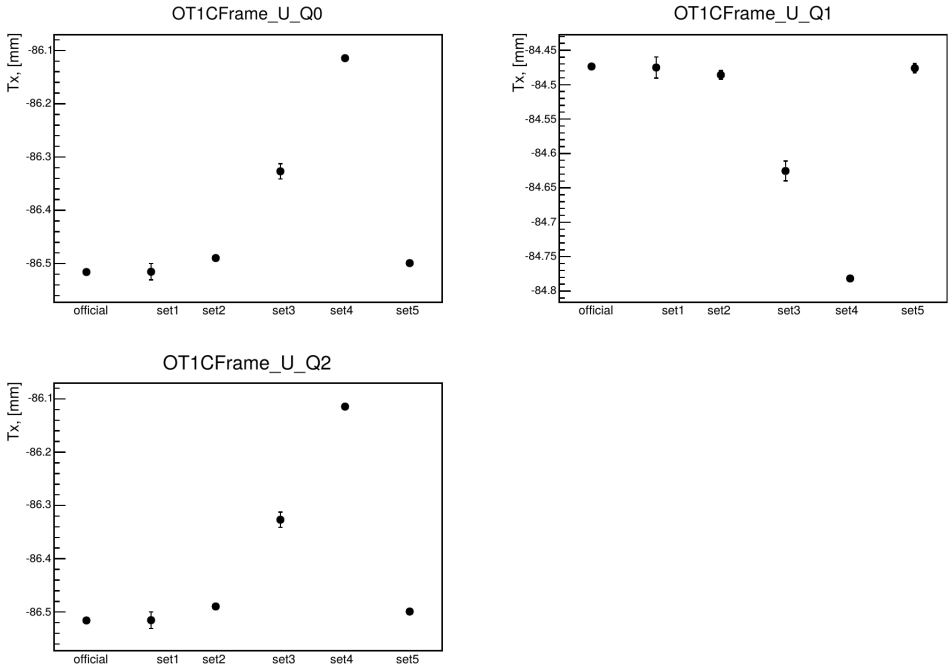


Fig. B.6 Tx alignment parameter fluctuation of OT1CFrame.

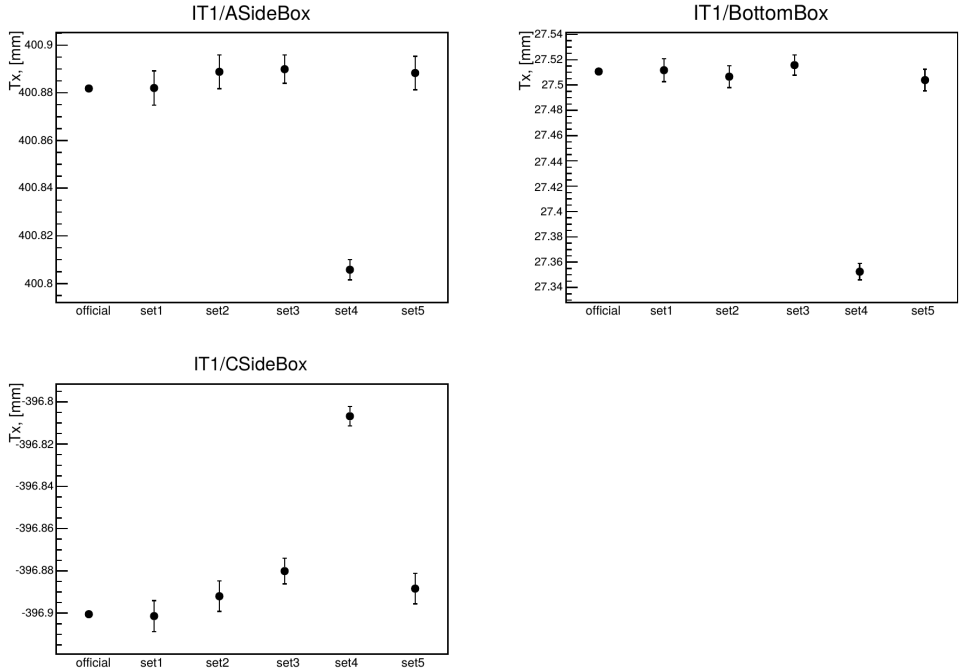


Fig. B.7 Tx alignment parameters fluctuation of IT1Boxes.

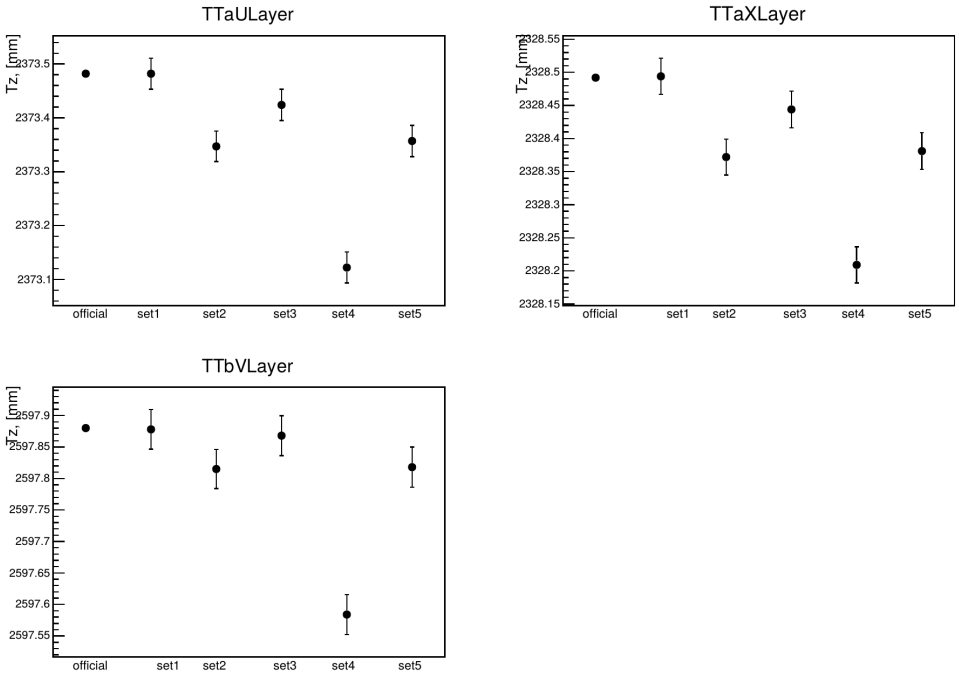


Fig. B.8 Tz alignment parameters fluctuation of TTLayers.

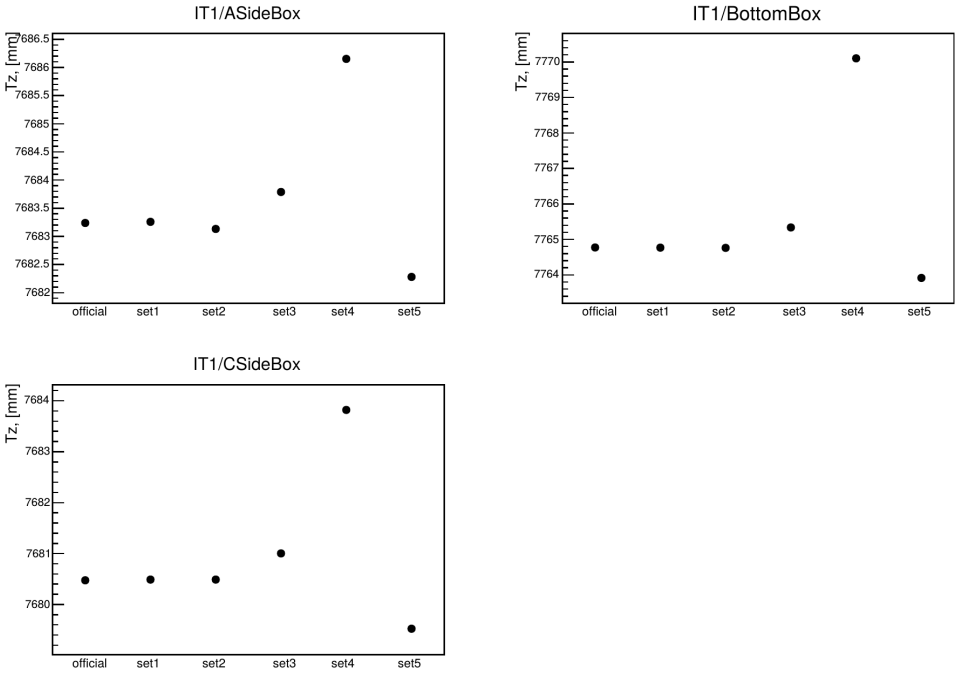


Fig. B.9 Tz alignment parameters fluctuation of IT1Boxes.

Time dependence study

2011 $D^0 \to K^-\pi^+$ data

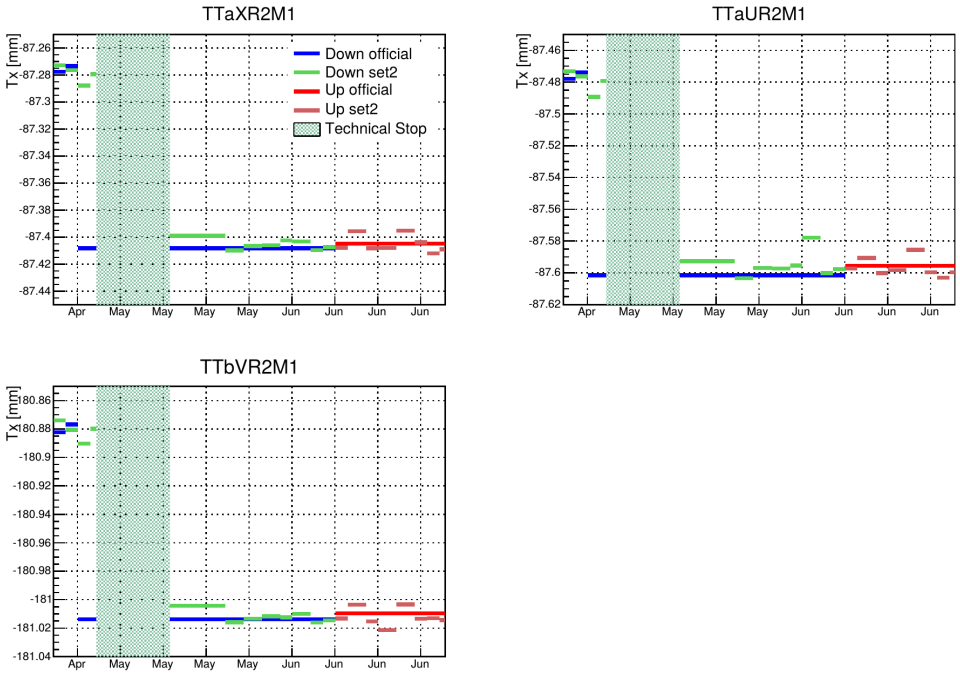


Fig. B.10 Tx alignment parameter fluctuation of TTModules with 2011 $D^0 \to K^-\pi^+$ data sample.

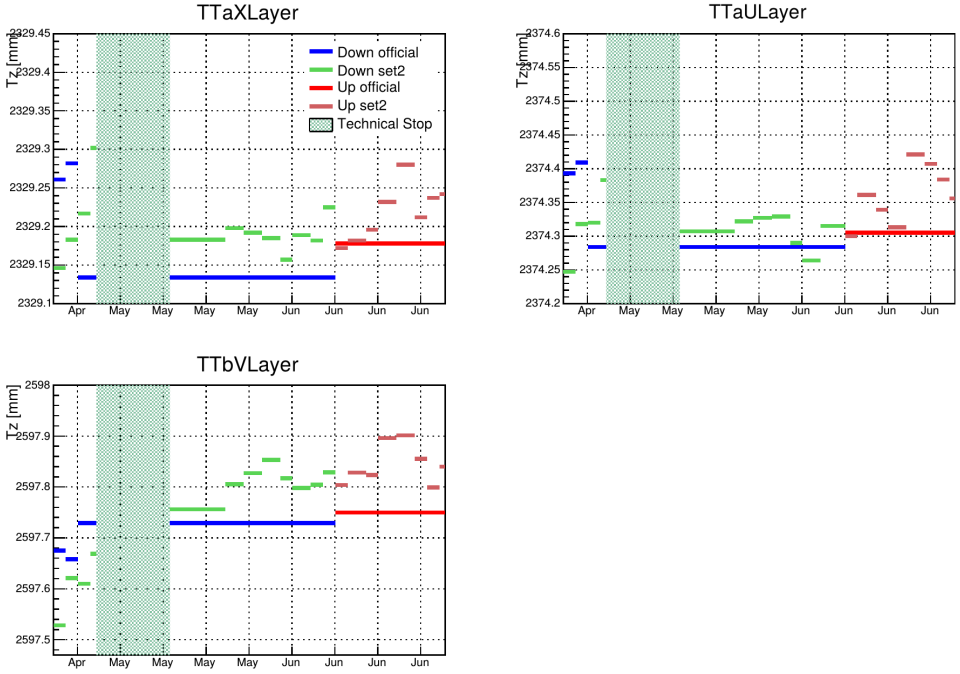


Fig. B.11 Tz alignment parameter fluctuation of TTLayers with 2011 $D^0 \to K^-\pi^+$ data sample.

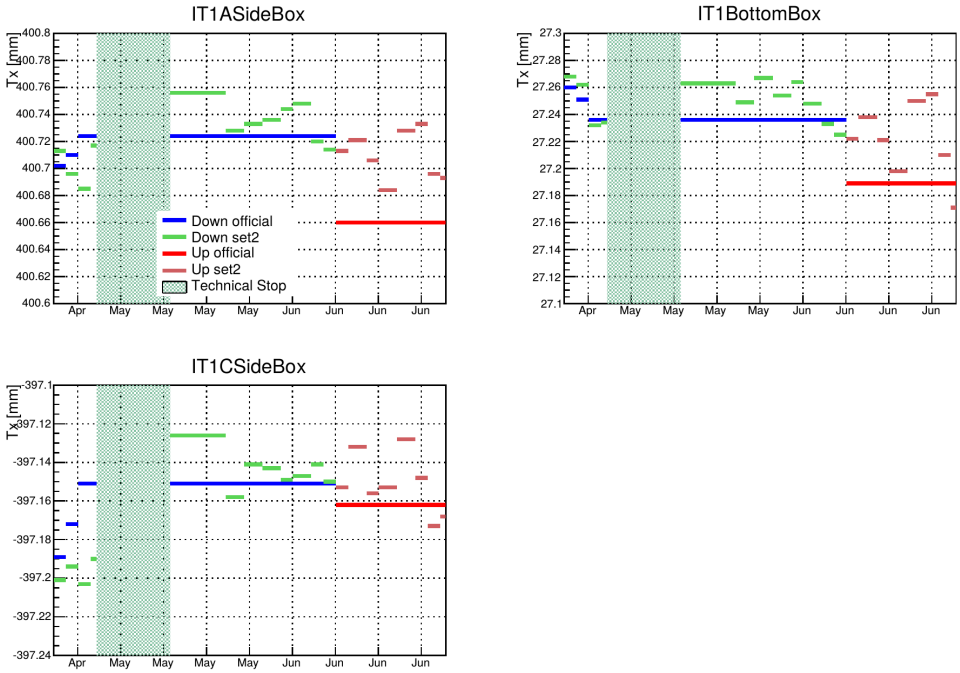


Fig. B.12 Tx alignment parameter fluctuation of IT1Boxes with 2011 $D^0 \to K^-\pi^+$ data sample.

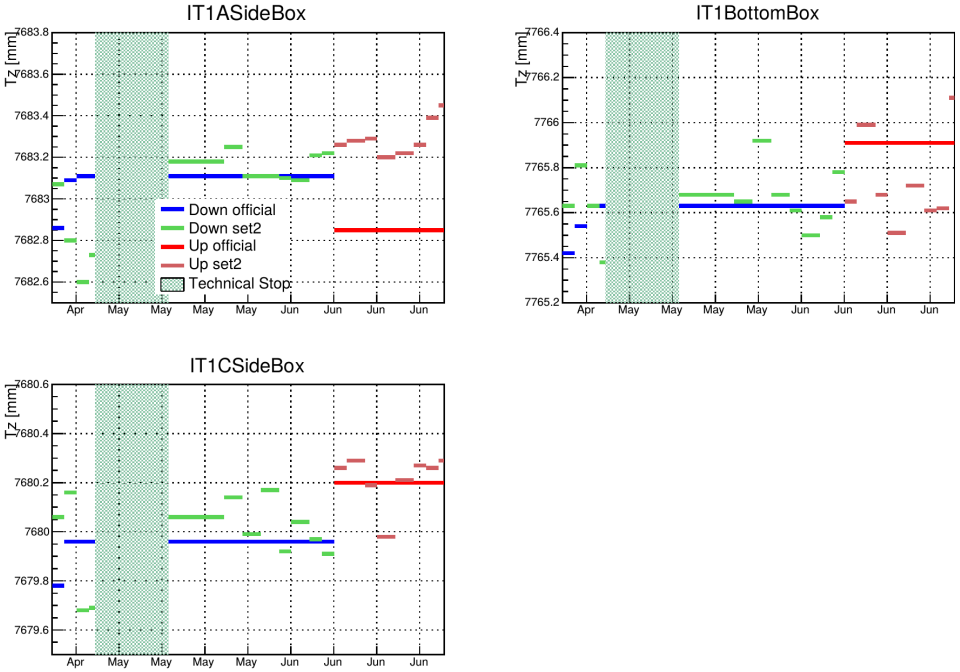


Fig. B.13 Tz alignment parameters fluctuation of IT1Boxes with 2011 $D^0 \to K^-\pi^+$ data sample.

2012 $D^0 \to K^-\pi^+$ data

The data sample is split into six consecutive time periods with opposite magnetic field (Table B.12) where VdM fills are the LHC fills for which a Van der Meer scan was performed

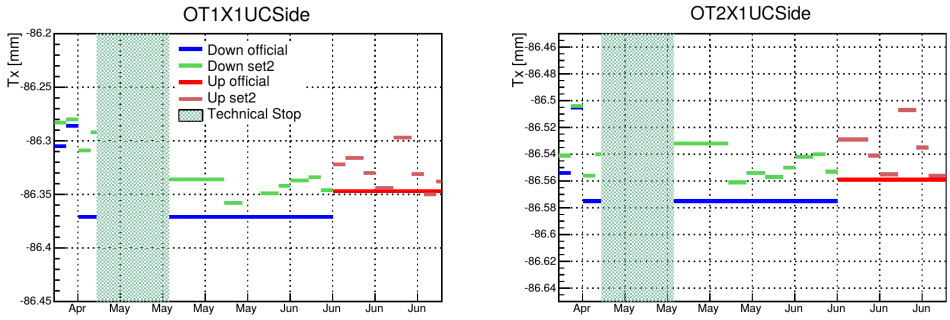


Fig. B.14 Tx alignment parameter fluctuation of OTCFrames with 2011 $D^0 \to K^-\pi^+$ data sample.

Table B.12 Time periods of 2012 data sample (July-October).

Period	Polarization	Fill number
01.07 - 17.07	Up	2848
17.07 - 19.07 VdM fills		
19.07 - 24.07	Up	2858
24.07 - 09.08	Down	2880
09.08 - 27.08	Up	2976
27.08 - 09.09	Down	3021
09.09 - 16.09 Technical Stop		
16.09 - 11.10	Up	3113 3114 3114 (another runs) 3121

to measure and optimize the luminosity of colliding beams [127]. The last time period is divided into four parts: 3113 fill, two data sets of 3114 fill and 3121 fill where 3114 fill data is used in the survey constraints study. The alignment procedure is performed for the set1 and set2 configurations that provided the best results in the study of OT station survey constraints.

Since the set2 configuration was available only for 3114 fill in time period after technical stop, additional time period (24.10.2012 - 08.11.2012) with Up magnet polarity is considered (Table B.13). For this study the alignment procedure is repeated for each sub-period of time using the set2 configuration of the OT survey constraints. These constraints significantly affect the IT and OT stations that are discussed below in more details.

Table B.13 Time periods of 2012 data sample (October-November).

Part	Day	Fill number
1	25.10	3220
2	28.10	3236
3	30.10	3242
4	02.11	3259
5	05.11	3265

The fluctuations of the Tx and Tz alignment parameters in 2012 data sample are presented in Figs. B.15-B.18 for IT and in Figs. B.19-B.22 for OT stations where the official parameters are a red and blue lines for Up and Down magnet polarizations, values of the set1 configuration are indicated by a brown and green lines for July-October period, values of the set2 configuration are marked by a pink and light blue lines, a technical stop and VdM time periods are indicated by a green area. Moreover, the survey position of IT and OT stations is shown by a dark line (Figs. B.17-B.22).

The observed fluctuations in Tx and Tz alignment parameters for IT1 boxes is 40 μm and 400 μm, respectively, which are consistent with fluctuations between three configurations.

For the OT station the observed difference between considered configurations is consistent for first and second stations. Since the third station is fixed by the survey constraints, the alignment parameters value of two configurations are the same as a survey position.

The sensitivity of the physical observables to the alignment quality is also investigated. The value of the D^0 meson mass and resolution over October-November time period for the set2 configuration is given in Fig. B.23 where the official parameter value is a red square, the set2 value is indicated by a pink triangle. The vertical error bars correspond to the uncertainty of the fit due to the mass spectrum. The distribution of the D^0 mass resolution is not sensitive to the alignment procedure, the points of two configurations are indistinguishable. The D^0 meson mass distribution is more sensitive to the variation of the alignment parameters, however the values of the time dependence study using 2012 data sample are within the D^0 meson mass uncertainty of the nominal mass value [3].

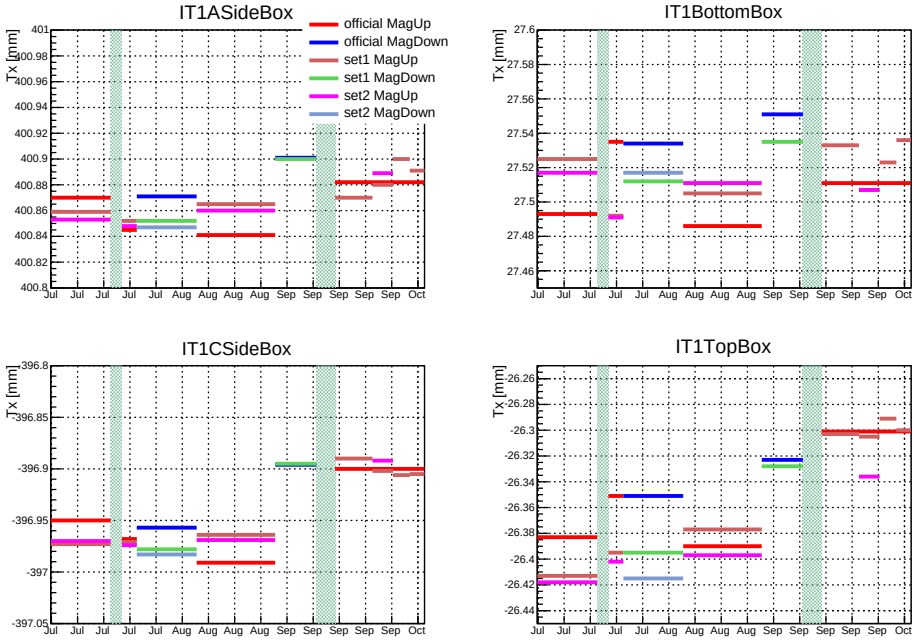


Fig. B.15 Tx alignment parameter fluctuation of IT1Boxes with 2012 $D^0 \to K^-\pi^+$ data sample (July-October).

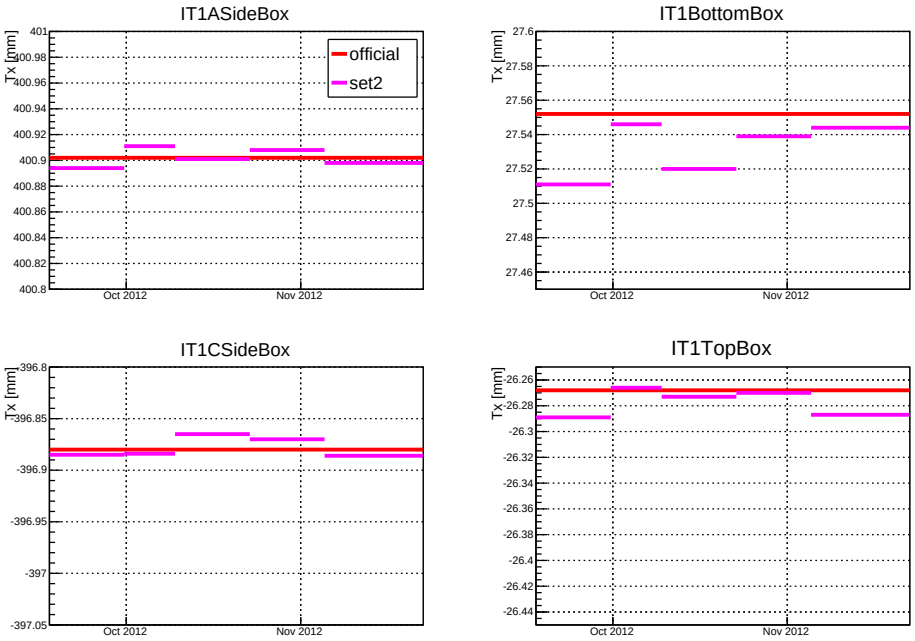


Fig. B.16 Tx alignment parameter fluctuation of IT1Boxes with 2012 $D^0 \to K^-\pi^+$ data sample (October-November).

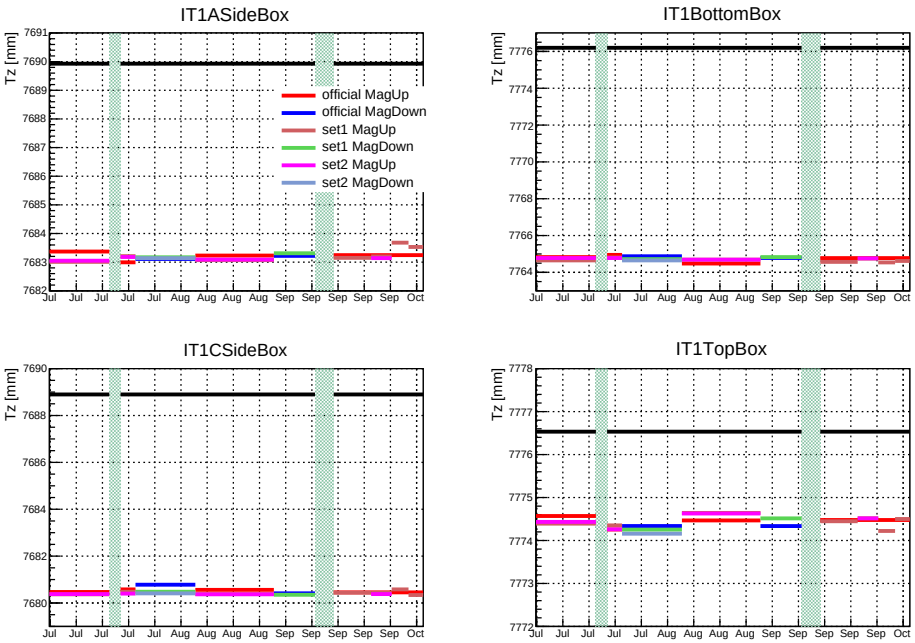


Fig. B.17 Tz alignment parameter fluctuation of IT1Boxes with 2012 $D^0 \to K^-\pi^+$ data sample (July-October).

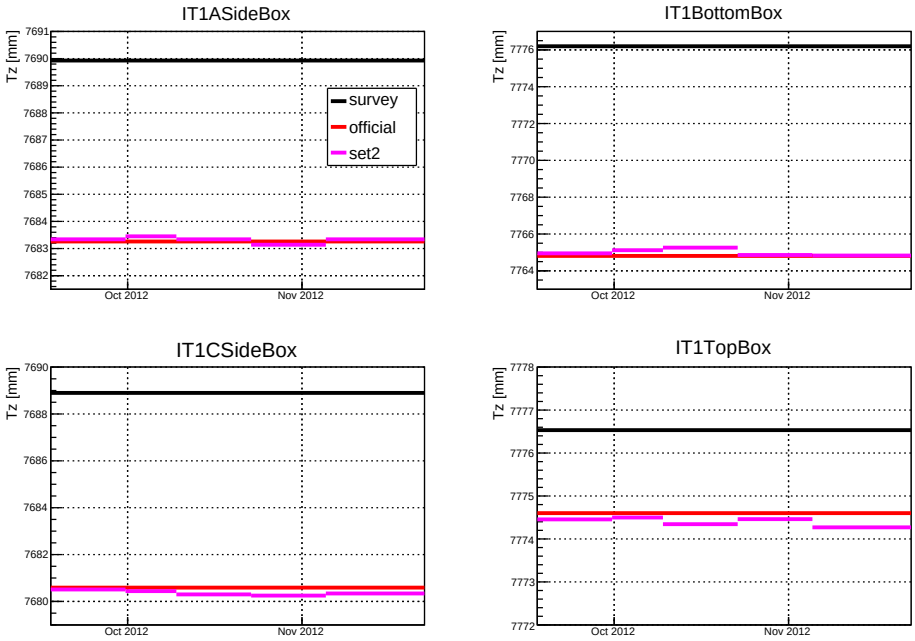


Fig. B.18 Tz alignment parameter fluctuation of IT1Boxes with 2012 $D^0 \to K^-\pi^+$ data sample (October-November).

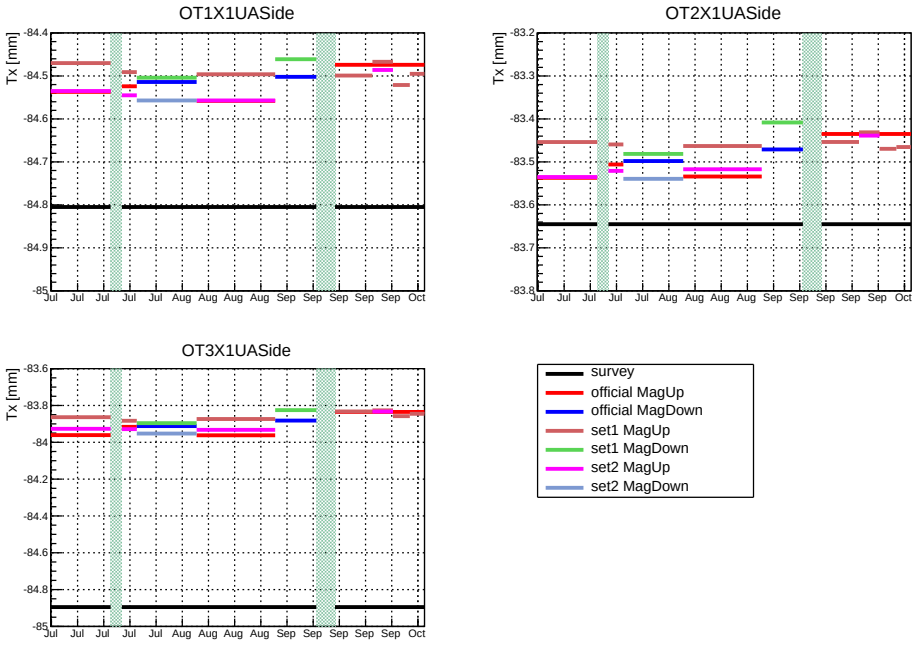


Fig. B.19 Tx alignment parameter fluctuation of OTCFrames with 2012 $D^0 \to K^-\pi^+$ data sample (July-October).

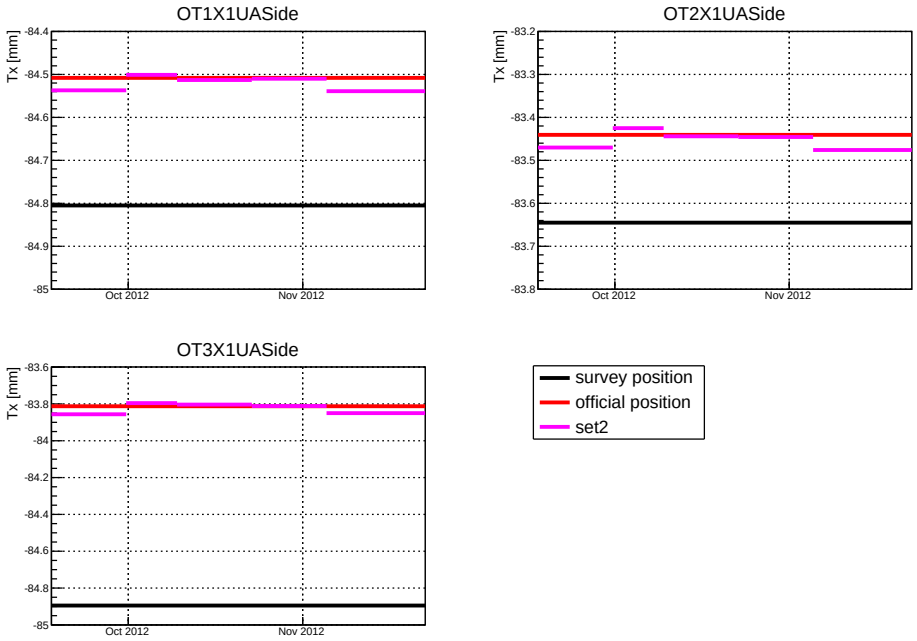


Fig. B.20 Tx alignment parameter fluctuation of OTCFrames with 2012 $D^0 \to K^-\pi^+$ data sample (October-November).

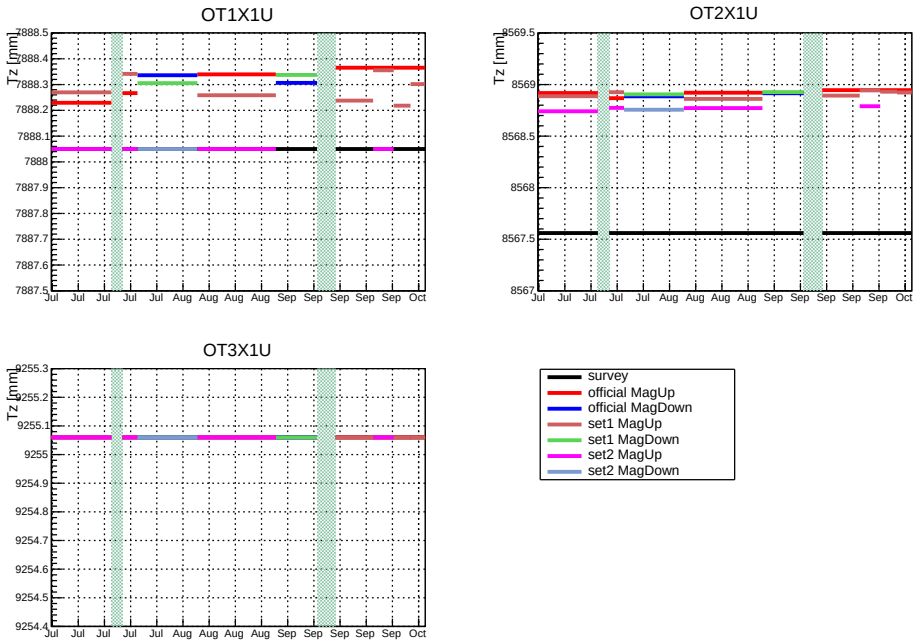


Fig. B.21 Tz alignment parameter fluctuation of OTCFrameLayers with 2012 $D^0 \to K^-\pi^+$ data sample (July-October).

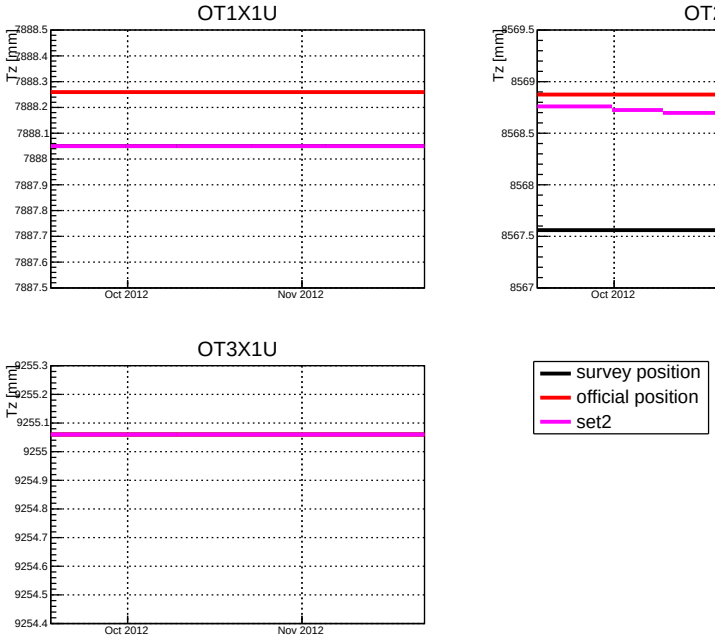


Fig. B.22 Tz alignment parameter fluctuation of OTCFrameLayers with 2012 $D^0 \to K^-\pi^+$ data sample (October-November).

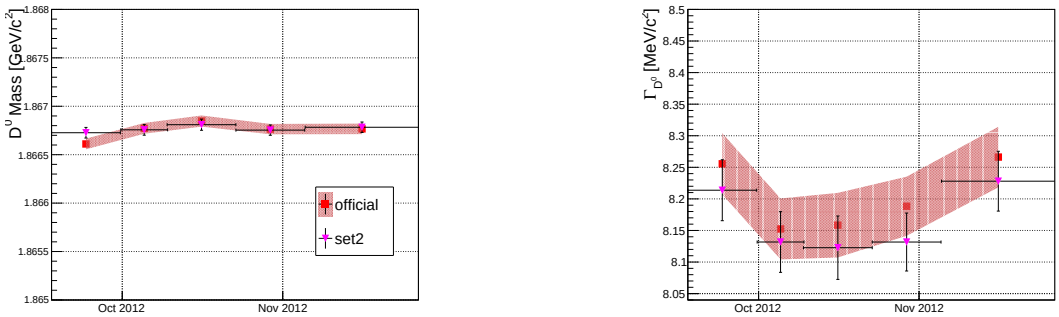


Fig. B.23 Distibution of the D^0 meson (left) mass and (right) resolution in 2012 data sample.

2011 $J/\psi \to \mu^+\mu^-$ data

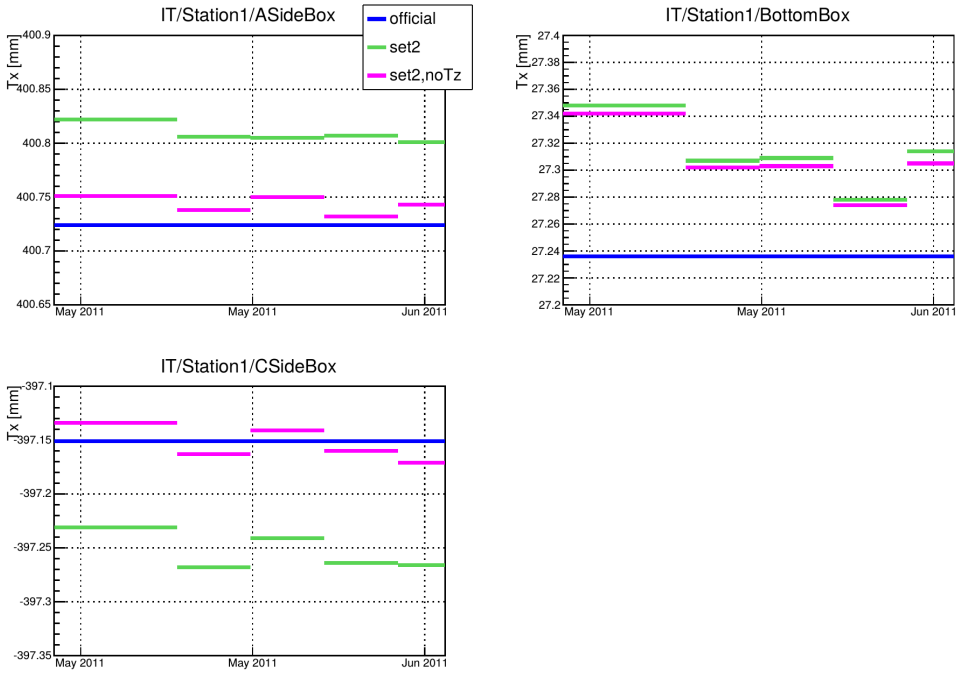


Fig. B.24 Tx alignment parameter fluctuation of IT1Boxes with 2011 $J/\psi \to \mu^+\mu^-$ data sample.

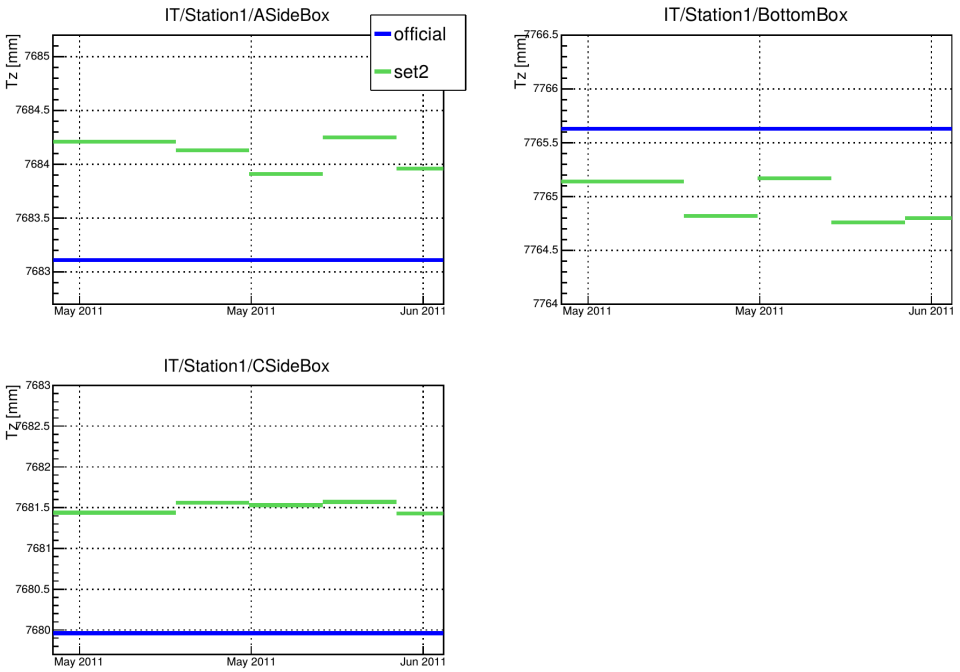


Fig. B.25 Tz alignment parameter fluctuation of IT1Boxes with 2011 $J/\psi \to \mu^+\mu^-$ data sample.

Alignment of the J/ψ and D^0 data

The time dependence study is repeated for the IT detector using set1 configuration to investigate the discrepancy between the D^0 and J/ψ data samples. The set1 configuration is used to exclude any bias due to OT station survey constraints. For the $D^0/J/\psi$ comparison 3114 fill of 2012 data is used.

The alignment constant fluctuations in Tx and Tz DoFs of TTLayers, OTCFrames and ITBoxes are presented in Figs. B.26-B.29 where the first point is an official alignment constant, the second and third points are indicated the D^0 and J/ψ alignment parameters, respectively. The error bars correspond to the parameter uncertainty returned by the alignment procedure. The observed difference between two samples of TT, IT and OT stations are reported in Table B.14.

Table B.14 Time dependence comparison of the 2012 D^0 and J/ψ data samples.

Detector	Alignable	DoFs	Difference [μm] set2	
			D^0	J/ψ
σ_{det} ~50 μm				
TT	Layers	Tz	5 - 80	5 - 120
	Modules	Tx	5 - 20	<10
σ_{det} ~50 μm				
IT	Boxes	Tz	10 - 500	500 - 1500
		Tx	5 - 50	40 - 120
σ_{det} ~200 μm				
OTCFrs	T(1|2)X1U(A|C)Side T3X1U(A|C)Side	Tx	10 - 50	5 - 50
OTCFrLrs	T1X1U	Tz	200	10
	T2X1U		150	30 - 80
	T3X1U		10^{-6}	10^{-6}

A discrepancy in the IT station position is observed between results obtained with D^0 and J/ψ data samples. One possible explanation of this discrepancy is a difference in track distribution for the IT detector between considered samples. Another one would be a difference in track angles which could decrease sensitivity of the IT detector to Tz alignment for the J/ψ data sample. The illumination maps of the IT and OT detectors are extracted from both samples to investigate the hypotheses with track angle difference. The tracks from the D^0 or J/ψ meson decay candidates are extrapolated to the z position of detector layers. The distribution of expected xy position of a track in the first, middle and the last layers of the tracker detectors is plotted in Figs. B.30-B.33. The two-dimensional distributions are not biased by the detector efficiency. In case of the OT stations, no significant difference is observed between D^0 and J/ψ data samples. For the IT detector a clear region around the beam hole is seen where no μ tracks from the J/ψ meson decays are reconstructed. It is clear that the number of hits observed in the IT detector is significantly reduced, especially for Top and Bottom boxes (Fig. B.34). Lower statistics for this detector makes it more susceptible to

statistical fluctuations in the alignment procedure.

In addition, the difference in the pseudorapidity η distribution between D^0 and J/ψ data samples is observed (Fig. B.36). The average track angles for the J/ψ meson decay are smaller compared to the $D^0 \to K\pi$ decay. This is lead to smaller sensitivity in the Tz variation of the detector stations.

Taking into account the above observations the D^0 data sample is better suited for alignment procedure of the IT station than J/ψ data sample.

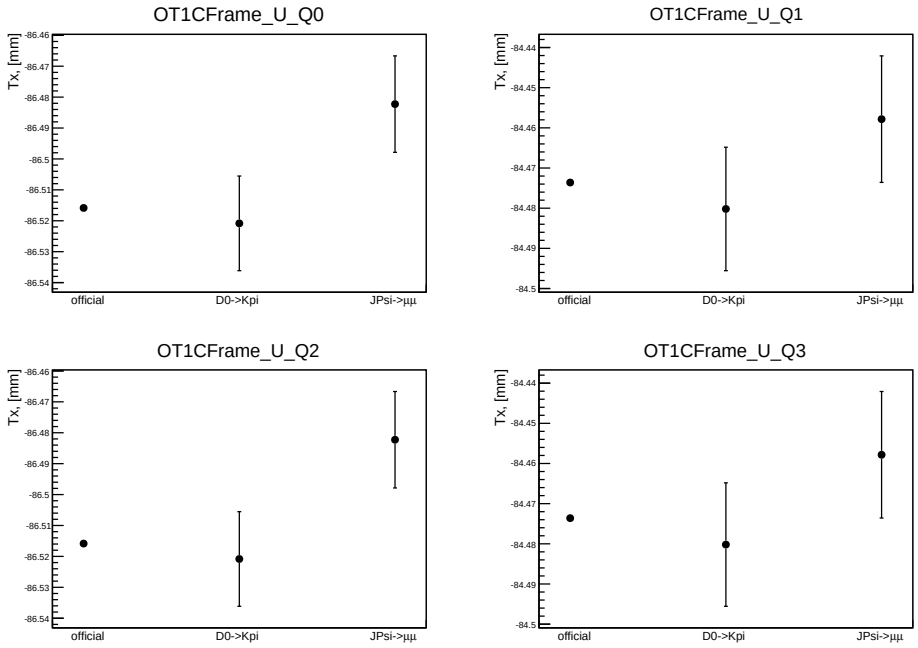


Fig. B.26 Tx alignment parameter fluctuation of OT1CFrames with 2012 D^0 and J/ψ data samples.

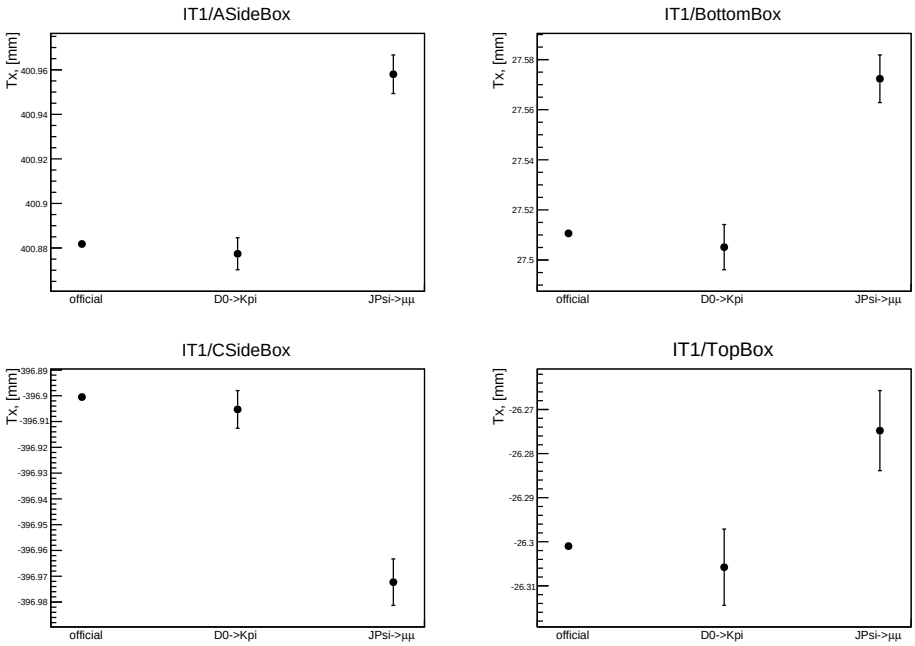


Fig. B.27 Tx alignment parameter fluctuation of IT1Boxes with 2012 D^0 and J/ψ data samples.

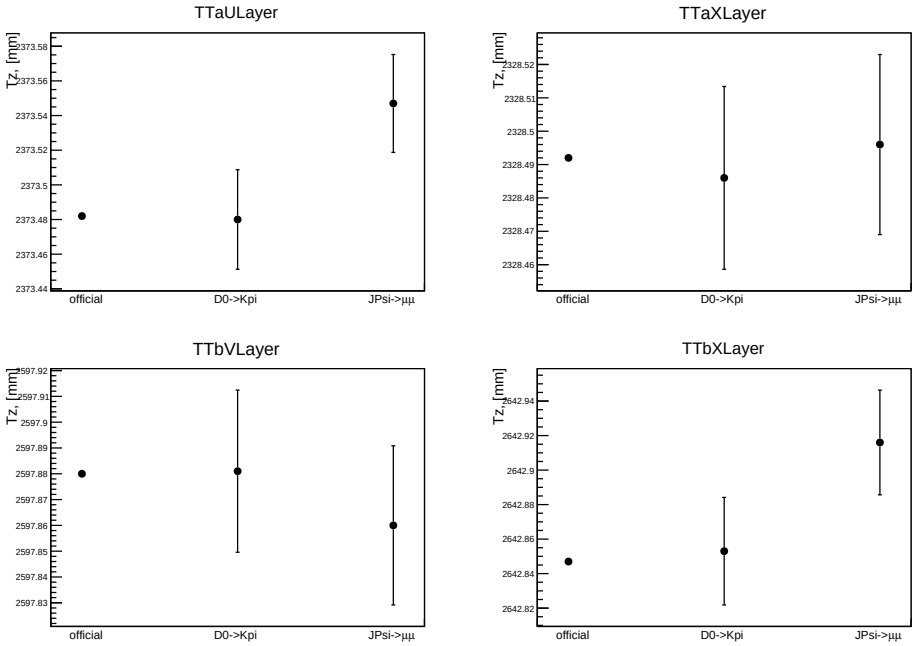


Fig. B.28 Tz alignment parameter fluctuation of TTLayers with 2012 D^0 and J/ψ data samples.

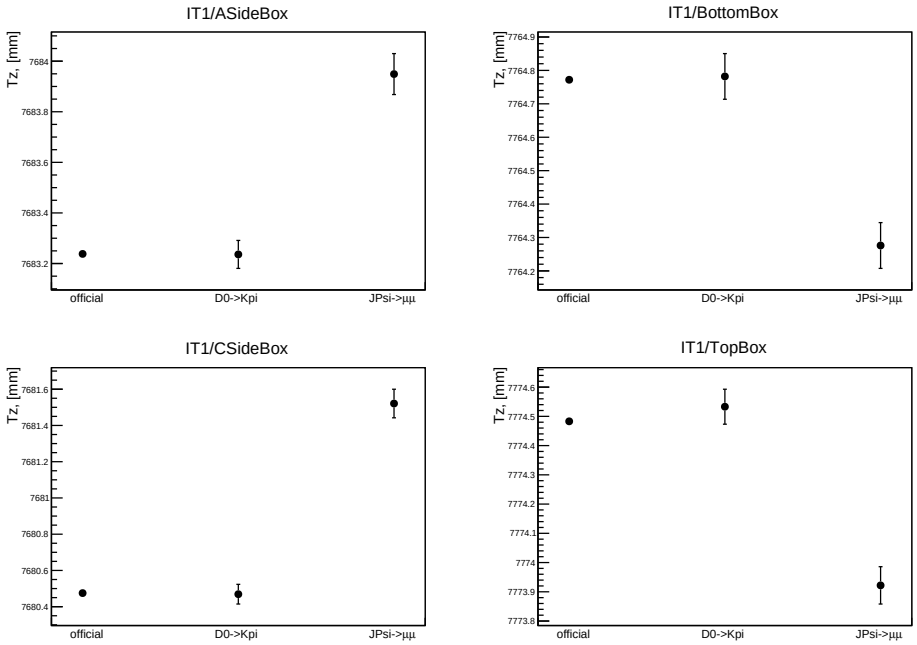


Fig. B.29 Tz alignment parameter fluctuation of IT1Boxes with 2012 D^0 and J/ψ data samples.

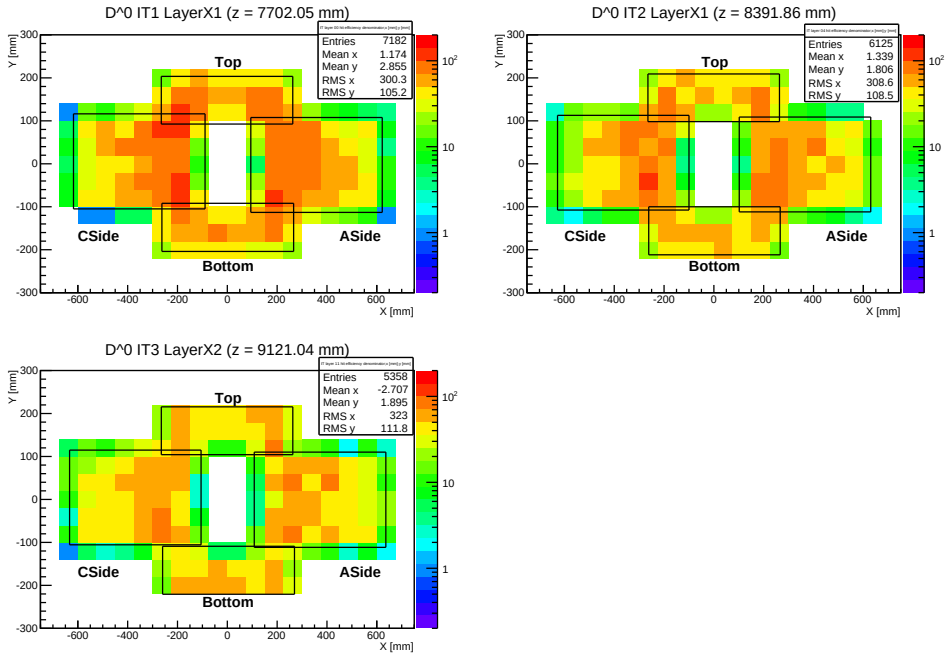


Fig. B.30 IT expected hits from extrapolated $D^0 \to K\pi$ tracks.

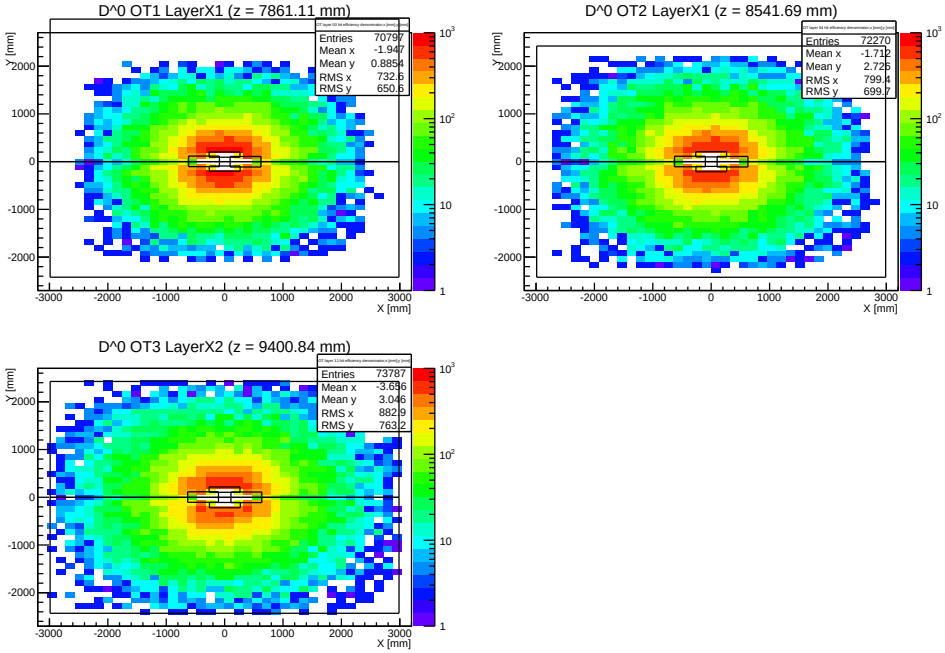


Fig. B.31 OT expected hits from extrapolated $D^0 \to K\pi$ tracks.

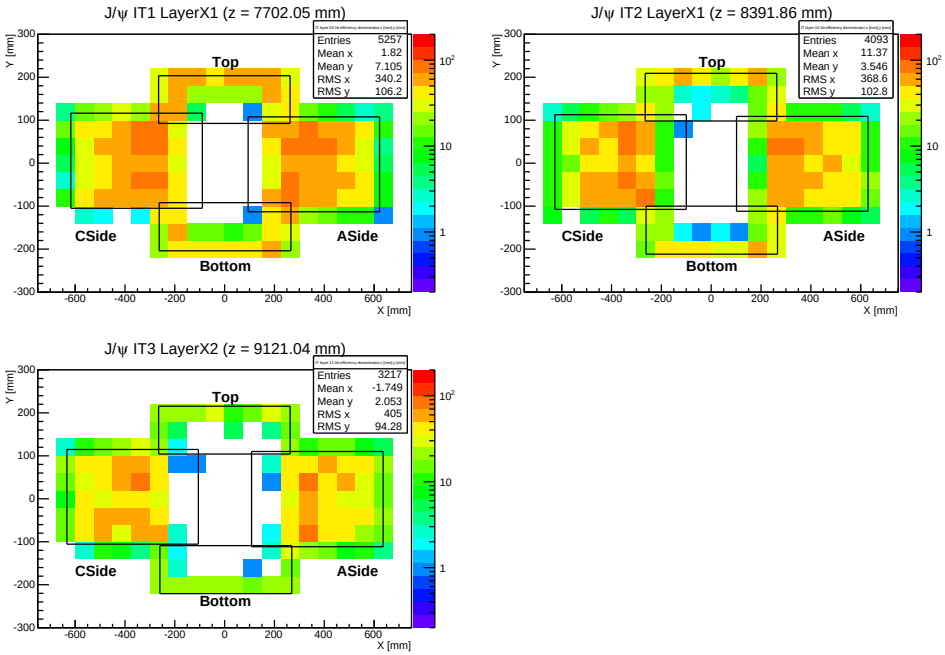


Fig. B.32 IT expected hits from extrapolated $J/\psi \to \mu\mu$ tracks.

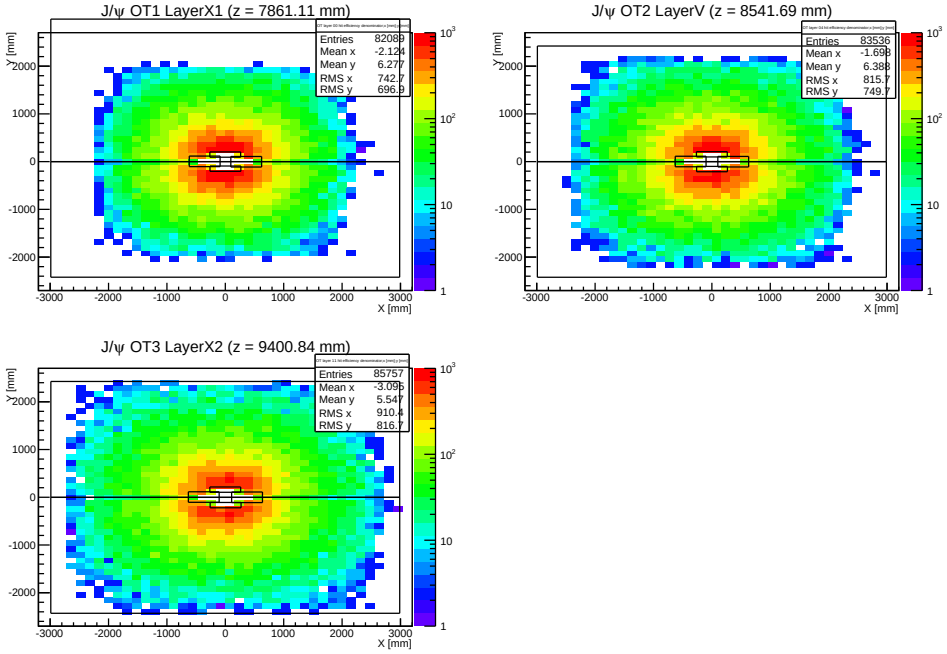


Fig. B.33 OT expected hits from extrapolated $J/\psi \to \mu\mu$ tracks.

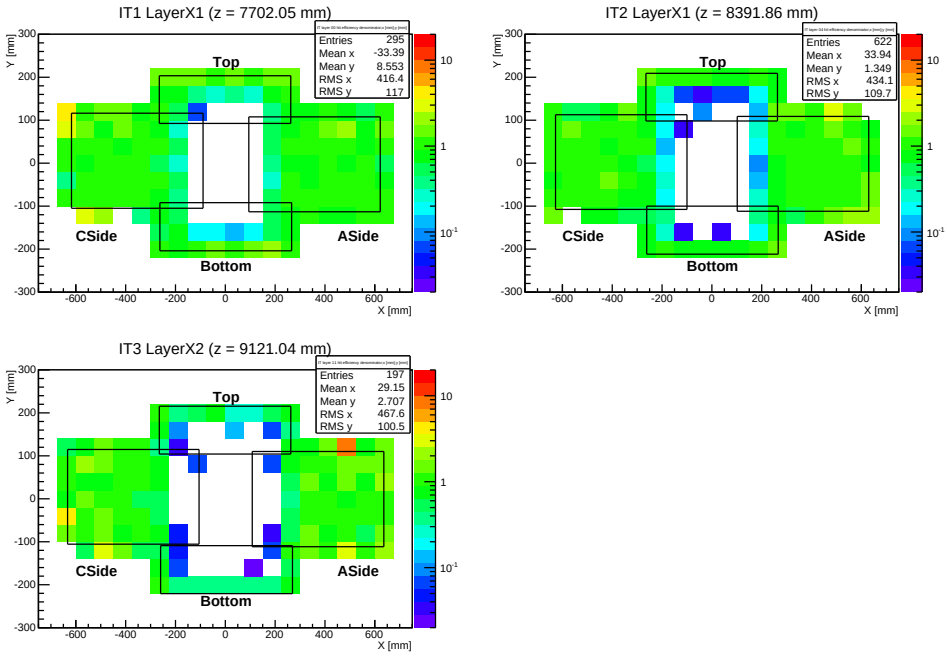


Fig. B.34 2012 $J/\psi/D^0$ data sample ratio of IT expected hits from extrapolated tracks.

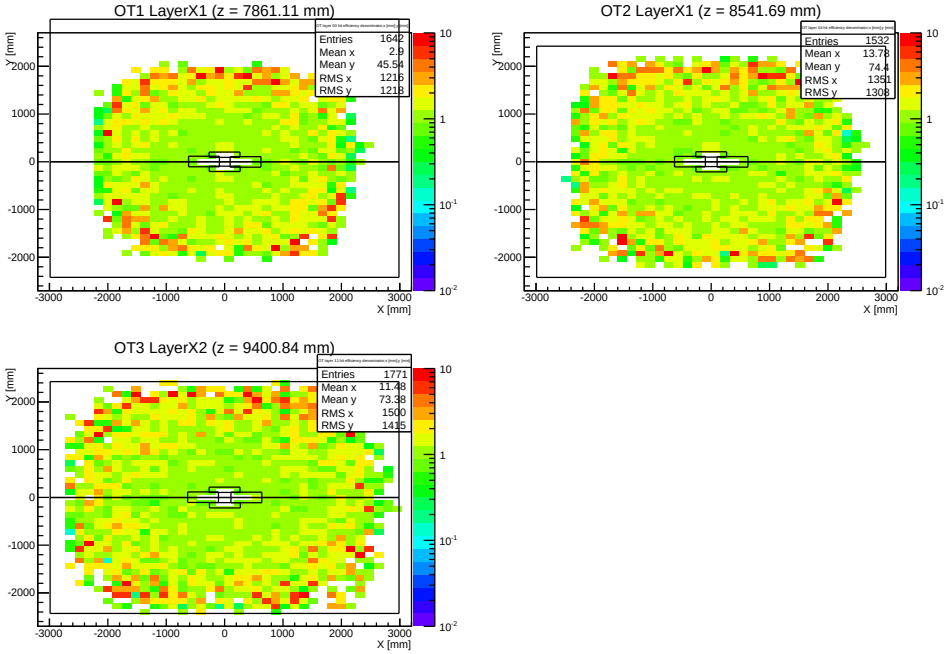


Fig. B.35 2012 $J/\psi/D^0$ data sample ratio of OT expected hits from extrapolated tracks.

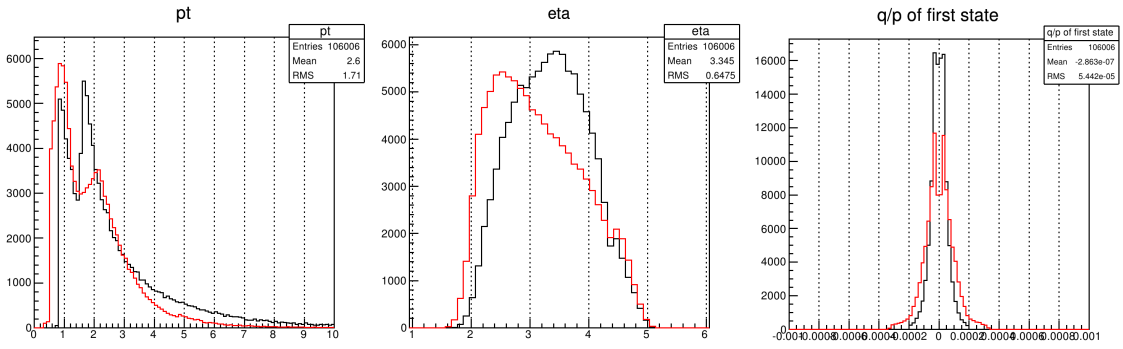


Fig. B.36 $D^0 \to K\pi$ and $J/\psi \to \mu\mu$ track properties: transverse momentum (left), pseudorapidity (middle) and ratio of the charge and momentum (right).

Appendix C

Analysis of the $B_s^0 \to J/\psi(e^+e^-)\phi$ decay

C.1 Trigger selection

A list of the L0, HLT1 and HLT2 trigger lines selection is given in Tables C.1-C.4[1].

Table C.1 Selection requirements of the L0 trigger lines.

Line	SPD mult	$E_{\mathrm{T}}^{\mathrm{hadron}}$ [GeV]	$E_{\mathrm{T}}^{\mathrm{electron}}$ [GeV]
`L0Electron`			>2.5
`L0Hadron`		>3.5	

Table C.2 Selection requirements of the `Hlt1TrackAllL0` trigger line.

L0	VELO miss hits	p [GeV/c]	p_{T} [GeV/c]	hits
`L0_DECISION_PHYSICS`	<3	>10	>1.7	>16

Table C.3 Selection requirements of the `Hlt2Topo(E)(2,3,4)BodyBBDT` trigger lines.

Line	χ^2_{track}/ndf	p [GeV/c]	p_{T} [GeV/c]	PID	χ^2_{IP}
`Hlt2Topo(2,3,4)BodyBBDT`	<3	>5	>0.5		>4
`Hlt2TopoE(2,3,4)BodyBBDT`	<3	>5	>0.5	∈(-2,5)	>4

C.2 BDT training

The section includes the details of the BDT training used for the final selection step (Sec. 4.1.4). The list of BDT input variables is reported in Table 4.4. The input variable distributions and correlation matrices for the signal and background samples used to train BDT are shown in Figs. C.1 and C.2, respectively.

[1]DOCA is the distance of the closest approach cut.

Table C.4 Selection requirements of the `Hlt2IncPhi` trigger line.

	Forward tracks	<120
$K^{\pm}$	$\chi^2_{\text{track}}/\text{ndf}$	<5
	χ^2_{IP}	>6
	p_{T}	>0.8 GeV/c
	PIDK	>0
ϕ	$\chi^2_{\text{vtx}}/\text{ndf}$	<20
	DOCA	<0.2 mm
	p_{T}	>1.8 GeV/c
	$\lvert M - m(\phi)\rvert$	<30 MeV/c^2

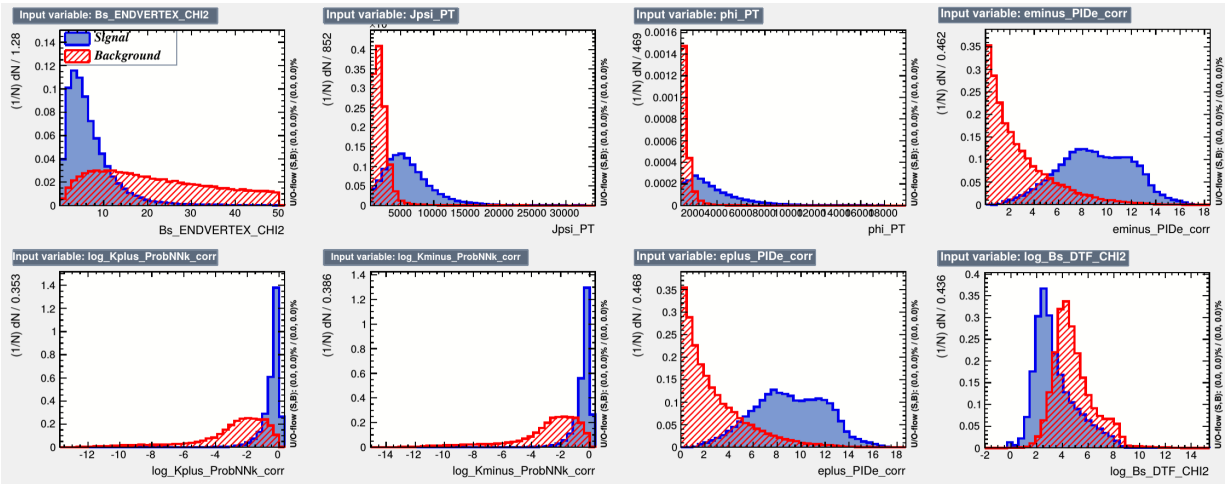


Fig. C.1 Input variable distributions for the signal (blue) and background (red) samples used in the BDT training.

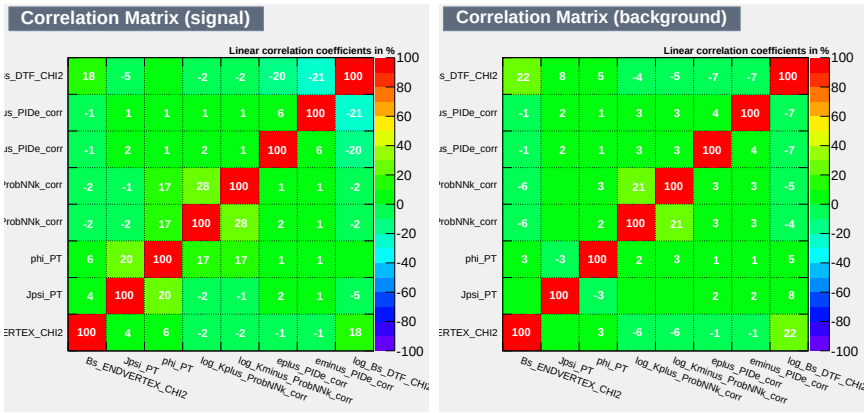


Fig. C.2 Correlation matrices for the (left) signal and (right) background samples.

C.3 BDT training for peaking background

The section includes the details of the BDT training used for the peaking background selection (Sec. 4.1.5). The list of BDT input variables is reported in Table 4.6. The input variable distributions and correlation matrices for the signal and background samples used to train

BDT are shown in Figs. C.3 and C.4, respectively.

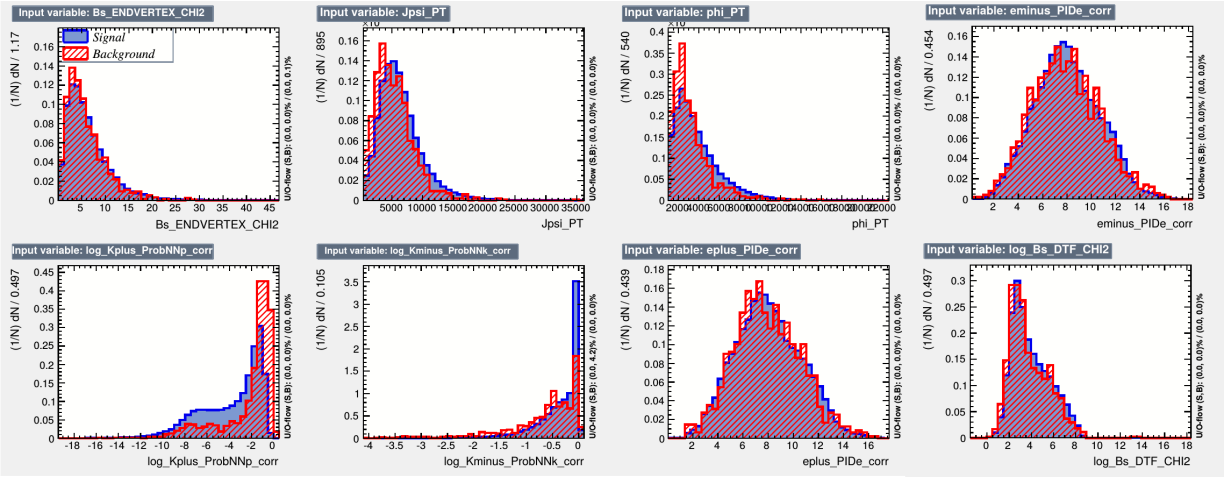


Fig. C.3 Input variable distributions for the signal (blue) and background (red) samples used in BDT training for the peaking background studies.

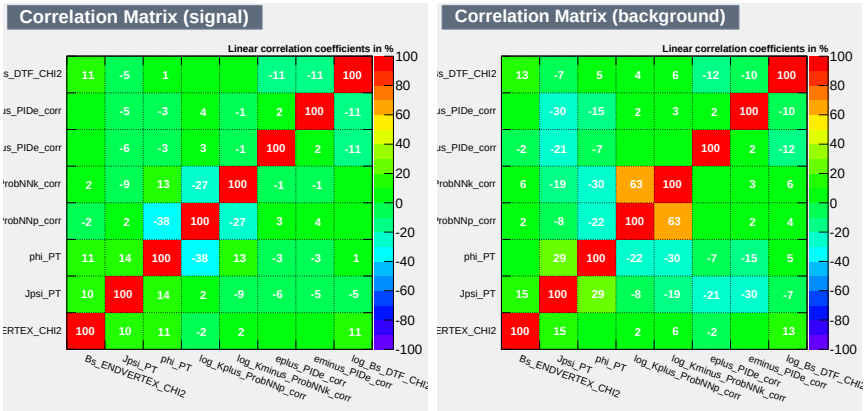


Fig. C.4 Correlation matrices for the (left) signal and (right) background samples used in the BDT training for the peaking background studies.

C.4 $J/\psi K^+K^-$, J/ψ and ϕ mass fit (update)

Fig. C.5 shows the combined fit to the $m(J/\psi K^+K^-)$, $m(e^+e^-)$ and $m(K^+K^-)$ distributions. The signal shape for the $J/\psi K^+K^-$ mass distribution is a sum of Gaussian and double Ipatia functions [117]. An exponential function is used to model the background events. The e^+e^- mass fit is performed using a sum of two Crystal Ball functions with a common mean for the signal candidates and an exponential function for the background events. The data sample is divided into three Bremsstrahlung categories for a better description of left tail of the $m(e^+e^-)$ distribution. In case of the ϕ meson, a sum of Gaussian and Voigtian functions with a common mean [105] is used as a mass model for the signal candidates. A Chebychev polynomial function [128] describes the background events.

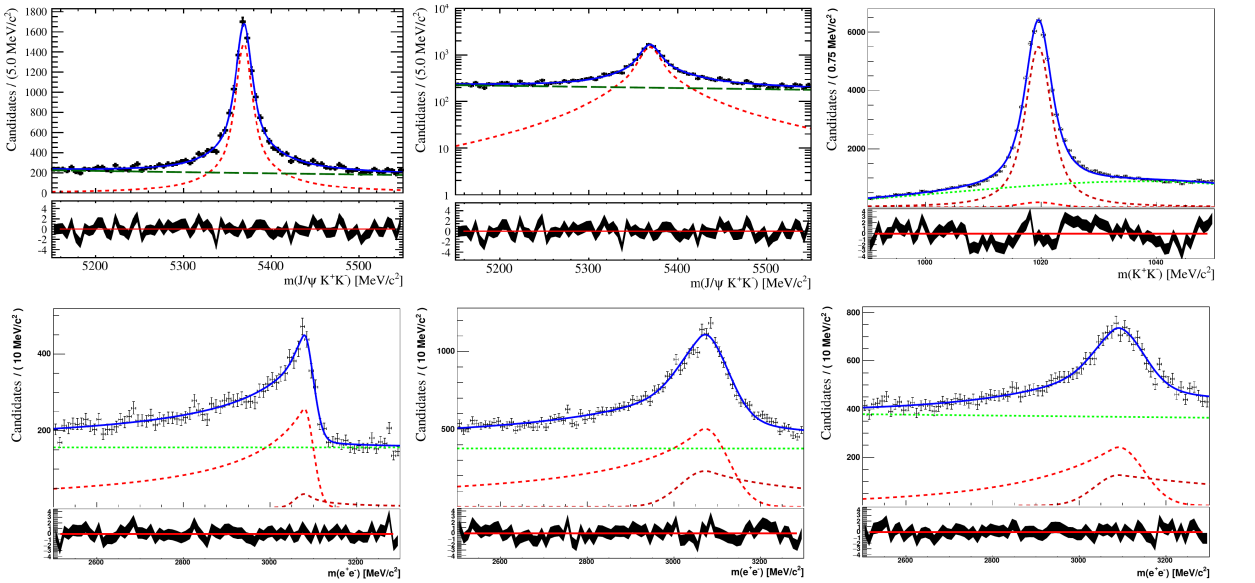


Fig. C.5 (top) Fit to the $m(J/\psi K^+K^-)$ and $m(K^+K^-)$ distribution for selected $B_s^0 \to J/\psi\phi$ candidates. The middle figure shows the $m(J/\psi K^+K^-)$ distribution in a logarithmic scale. (bottom) Fit to the $m(e^+e^-)$ distribution divided into three Bremsstrahlung categories: (left) without, (middle) one and (right) more reconstructed photons coming from the electron radiation. In all cases, the solid blue line shows the total fit. The signal and combinatorial background components are given by red and green lines, respectively.

C.5 $B^0 \to J/\psi K^*$ decay analysis

C.5.1 BDT training

The section includes the details of the BDT training of the control mode used for the final step of selection (Sec. 4.3.1). The list of the BDT input variables is shown in Table 4.10. The input variable distributions and correlation matrices for the signal and background samples are shown in Figs. C.6 and C.7, respectively.

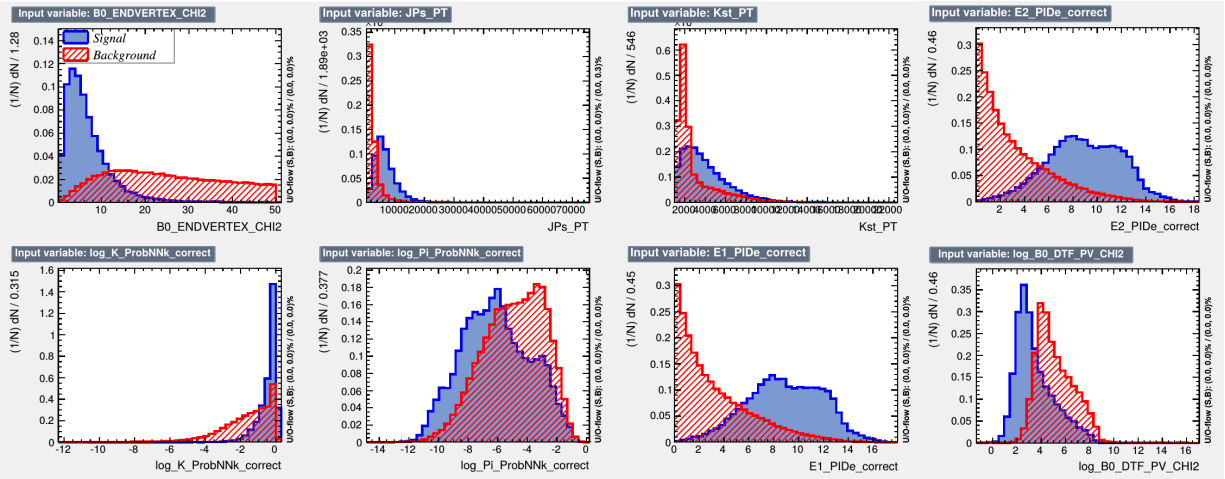


Fig. C.6 Input variable distributions for the signal (blue) and background (red) samples used in the BDT training of the control decay mode.

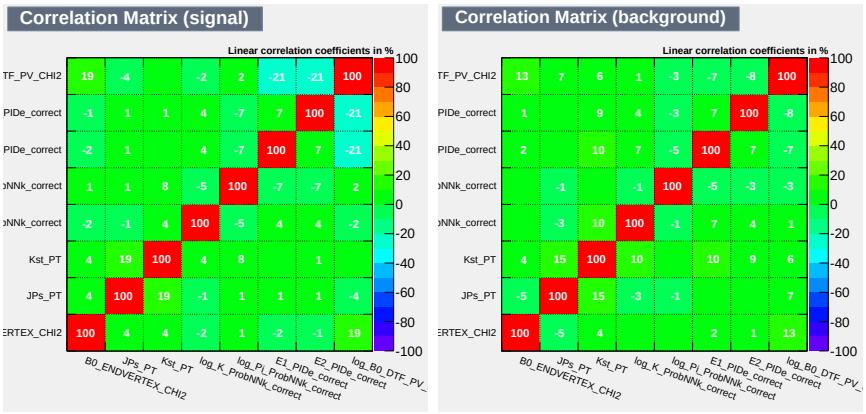


Fig. C.7 Correlation matrices for the (left) signal and (right) background samples used in the BDT training of the control decay mode.

C.5.2 MC reweighting

The comparison of the ProbNNK(K^+, π^-) and PIDe($e^\pm$) distributions of the selected $B^0 \to J/\psi K^*$ candidates for sWeighted data and simulated data with and without MC reweighting (Fig. C.8).

C.5.3 Comparison of the $B_s^0 \to J/\psi\phi$ and $B^0 \to J/\psi K^*$ distributions

The section shows the comparison of the background subtracted distribution of variables that have been used to train the BDT for the signal and control channels (Fig. C.9). A good agreement is observed for distributions except ProbNNK variable which can be attributed to different particles in the final state: K^- for the $B_s^0 \to J/\psi\phi$ decay and π^- for the $B^0 \to J/\psi K^*$ decay.

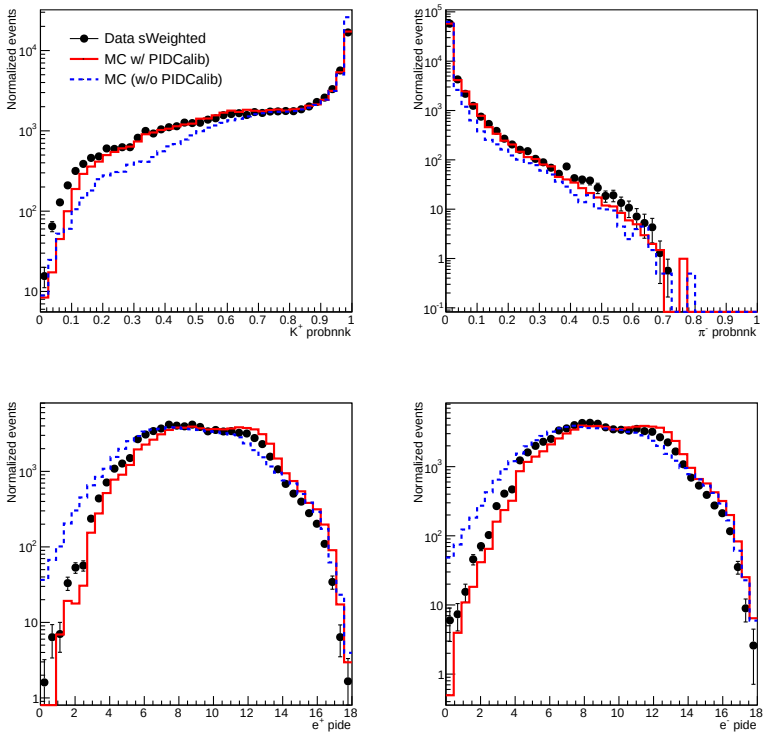


Fig. C.8 ProbNNK(K^+, π^-) and PIDe($e^\pm$) distributions of the selected $B^0 \to J/\psi(e^+e^-)K^*$ candidates for sWeighted data (black points), MC (blue dashed line) and resampled MC (red line). The distributions are in a logarithmic scale.

C.5.4 Partially reconstructed background

The $B^0 \to J/\psi X$ simulated data sample is used to distinguish the different type of partially reconstructed background in the $m(e^+e^-K^+\pi^-)$ distribution in data sample in the same way as it is performed for B_s^0 meson decay (Sec. 4.2.1). The mass fit is applied to the simulated events that give the largest contribution in the partially reconstructed background: $B^0 \to \psi(2S)K^*$ and $B^0 \to \chi_{c1}(1P)K^*$ decay modes. Their MC samples are also divided into three Bremsstrahlung categories and the mass fit is shown in Fig. C.10. The B^0 mass range (4500,5600) MeV/c^2 is applied to better describe the distribution shape. The obtained fit parameters are reported in Table C.5 where the mean μ, width σ and fraction f are used in the mass fit performed on data sample (Table 4.11). In order to constrain the number of the partially reconstructed background contribution in data, the fraction of the $B^0 \to \chi_{c1}(1P)K^*$ events to the number of partially reconstructed background[2] is calculated from the simulated data sample and used as a Gaussian constraint for the mass fit on the data sample, $f_{B^0 \to \chi_{c1}(1P)K^*}$.

[2]The number of partially reconstructed background is a sum of the $B^0 \to \psi(2S)K^*$ and $B^0 \to \chi_{c1}(1P)K^*$ events.

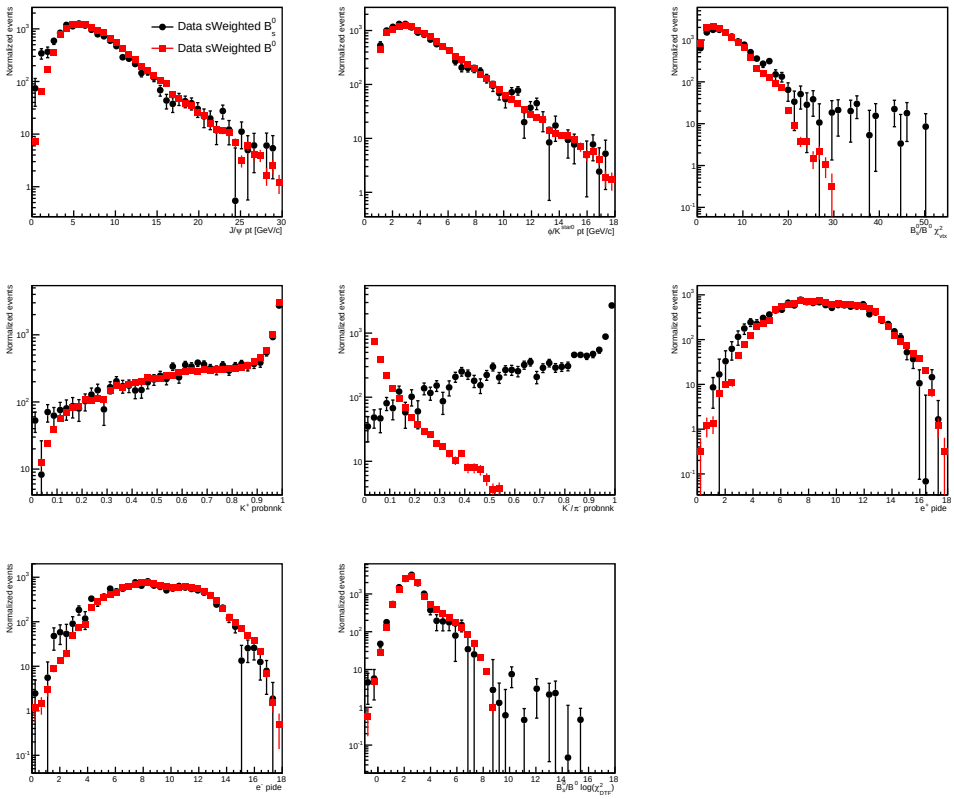


Fig. C.9 Comparison of the variable distributions used to train the BDT for the $B_s^0 \to J/\psi\phi$ (black circle) and $B^0 \to J/\psi K^*$ (red square) data.

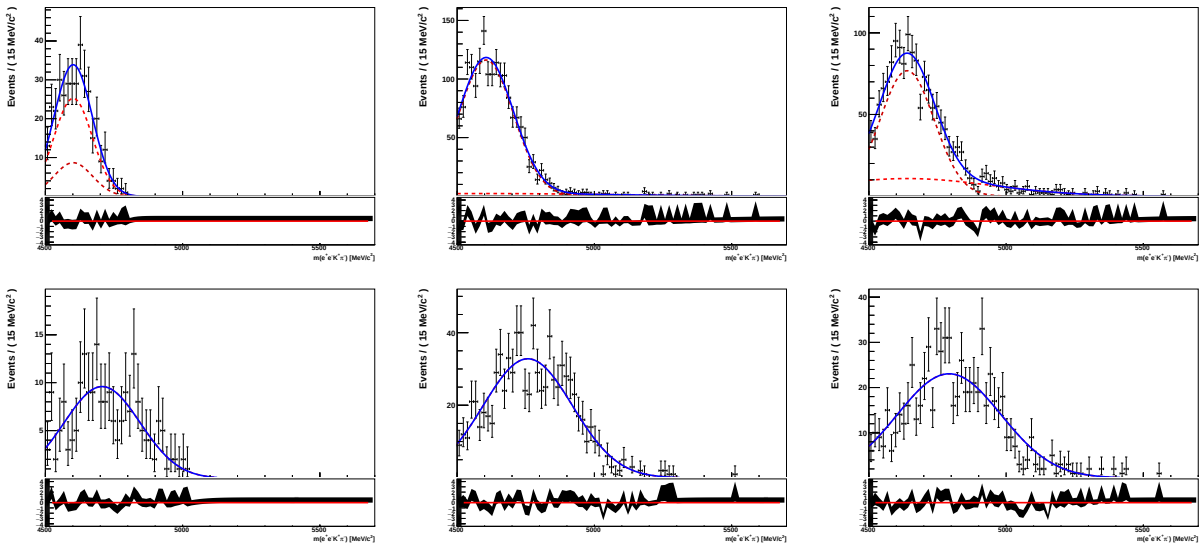


Fig. C.10 Distribution of the $m(e^+e^-K^+\pi^-)$ for the (top) $B^0 \to \psi(2S)K^*$ and (bottom) $B^0 \to \chi_{c1}(1P)K^*$ simulated data sample divided into three Bremsstrahlung categories: (left) without, (middle) one and (right) more photons coming from the electron radiation. The blue line shows the total fit which is shown the contribution of two Gaussian functions (red line) in case of the $B^0 \to \psi(2S)K^*$ decay.

Table C.5 Results of the fit to the $m(e^+e^-K^+\pi^-)$ distribution in the $B^0 \to \psi(2S)K^*$ and $B^0 \to \chi_{c1}(1P)K^*$ simulated data sample divided into three Bremsstrahlung categories. The shape is modelled by a double or single Gaussian function. The last line shows a fraction of the $B^0 \to \chi_{c1}(1P)K^*$ events with respect to the number of partially reconstructed background.

Parameter	0γ	1γ	2γ
	$B^0 \to \psi(2S)K^*$		
μ[MeV/c^2]	4601±6	4606±4	4639±4
σ_1[MeV/c^2]	70±13	433±58	310±20
σ_2[MeV/c^2]	70±109	99±4	99±5
f_1	0.7±0.5	0.05±0.01	0.24±0.04
N_{evt}	653±19	1745±42	1546±39
	$B^0 \to \chi_{c1}(1P)K^*$		
μ[MeV/c^2]	4707±13	4757±7	4792±9
σ[MeV/c^2]	139±10	163±5	189±7
N_{evt}	207±14	841±29	685±26
$f_{B^0 \to \chi_{c1}(1P)K^*}$	0.361±0.027	0.325±0.013	0.307±0.013

C.6 Angular acceptance

The acceptance function is written as:

$$\frac{\mathcal{S}^{\text{obs}}(\Omega|t)}{\mathcal{S}^{\text{phys}}(\Omega|t)} = \frac{\varepsilon(t,\Omega)s(t,\Omega)}{\int \mathrm{d}\Omega\varepsilon(t,\Omega)s(t,\Omega)} \frac{\int \mathrm{d}\Omega s(t,\Omega)}{s(t,\Omega)} = \frac{\varepsilon(t,\Omega)}{\int \mathrm{d}\Omega\varepsilon(t,\Omega)\mathcal{S}^{\text{phys}}(\Omega|t)} = \frac{\varepsilon(t,\Omega)}{\langle\varepsilon\rangle(t)},$$

$$\varepsilon(t,\Omega) = \langle\varepsilon\rangle(t)\frac{\mathcal{S}^{\text{obs}}(\Omega|t)}{\mathcal{S}^{\text{phys}}(\Omega|t)},$$

where $\mathcal{S}^{\text{phys}}$ is the PDF describing the underlying physics distribution according to which simulated events are generated and $\mathcal{S}^{\text{obs}}$ is the PDF describing the observed distribution of simulated events, including detector and selection effects. The function s describes the differential rate for the signal decay. The angular mean of the total acceptance as a function of time is given by $\langle\varepsilon\rangle(t) \equiv \int \mathrm{d}\Omega\varepsilon(t,\Omega)\mathcal{S}^{\text{phys}}(\Omega|t)$.

If the acceptance function factorizes into a decay time part and an angular part, $\varepsilon(t,\Omega) = \varepsilon_t(t) \times \varepsilon_a(\Omega)$, the mean acceptance reduces to:

$$\langle\varepsilon\rangle(t) = \int \mathrm{d}\Omega\varepsilon(t,\Omega)\mathcal{S}^{\text{phys}}(\Omega|t) = \varepsilon_t(t) \times \int \mathrm{d}\Omega\varepsilon_a(\Omega)\mathcal{S}^{\text{phys}}(\Omega|t) = \varepsilon_t(t) \times \langle\varepsilon_a\rangle(t),$$

where $\langle\varepsilon_a\rangle(t)$ is the mean angular acceptance as a function of decay time. The acceptance function is now given by:

$$\varepsilon(t,\Omega) = \varepsilon_t(t) \times \varepsilon_a(\Omega) = \varepsilon_t(t) \times \langle\varepsilon_a\rangle(t)\frac{\mathcal{S}^{\text{obs}}(\Omega|t)}{\mathcal{S}^{\text{phys}}(\Omega|t)}.$$

To derive an expression that can be used to determine the acceptance weights, the constant

E_a is defined:

$$E_a \equiv \left[\int \mathrm{d}t \frac{\mathcal{S}^{\mathrm{obs}}(t)}{\langle \varepsilon_a \rangle (t)} \right]^{-1}.$$

Note that E_a reduces to the mean angular acceptance $\langle \varepsilon_a \rangle \equiv \int \mathrm{d}\Omega \varepsilon_a(\Omega) \mathcal{S}^{\mathrm{phys}}(\Omega)$ if the physics PDFs for time and angles factorize; $\mathcal{S}^{\mathrm{obs}}(t,\Omega) = \mathcal{S}^{\mathrm{obs}}(t) \times \mathcal{S}^{\mathrm{obs}}(\Omega)$ and $\mathcal{S}^{\mathrm{obs}}(\Omega|t) = \mathcal{S}^{\mathrm{obs}}(\Omega)$. The ratios of the acceptance weights and E_a can be written as:

$$\frac{1}{E_a}\xi_i \equiv \int \mathrm{d}t \frac{\mathcal{S}^{\mathrm{obs}}(t)}{\langle \varepsilon_a \rangle (t)} \int \mathrm{d}\Omega \varepsilon_a(\Omega) f_i(\Omega)$$

$$= \int \mathrm{d}t \mathrm{d}\Omega \mathcal{S}^{\mathrm{obs}}(t) \frac{\mathcal{S}^{\mathrm{obs}}(\Omega|t)}{\mathcal{S}^{\mathrm{phys}}(\Omega|t)} f_i(\Omega) = \int \mathrm{d}t \mathrm{d}\Omega \mathcal{S}^{\mathrm{obs}}(t,\Omega) \frac{f_i(\Omega)}{\mathcal{S}^{\mathrm{phys}}(\Omega|t)}.$$

This quantity is estimated by a MC integral, which is a sum over the selected events in the simulation:

$$E\left(\frac{1}{E_a}\xi_i\right) = \frac{1}{N^{\mathrm{obs}}} \sum_{k=1}^{N^{\mathrm{obs}}} \frac{f_i(\Omega_k)}{\mathcal{S}^{\mathrm{phys}}(\Omega_k|t_k)}$$

where N^{obs} is the number of observed events. Since E_a is a constant factor, common to all weights, it gives a constantterm in the log-likelihood and may be ignored there.

The K^+ and B_s^0 momentum and ϕ mass distributions for data and simulation samples are compared per year as shown in Fig. C.11. No difference is observed between both samples.

The correlations between the angular acceptance weights for full simulated sample is shown in Table C.6.

Table C.6 Correlations between the angular acceptance weights for a sum of 2011 and 2012 simulated sample.

k	1(00)	2(‖‖)	3(⊥⊥)	4(‖⊥)	5(0‖)	6(0⊥)	7(SS)	8(S‖)	9(S⊥)	10(S0)
1(00)	1	-0.67	-0.69		0.27		-0.06			
2(‖‖)		1	0.40		0.21		0.24			
3(⊥⊥)			1		0.31		0.18			
4(‖⊥)				1		-0.08				
5(0‖)					1		0.30			
6(0⊥)						1				
7(SS)							1			
8(S‖)								1		0.14
9(S⊥)									1	
10(S0)										1

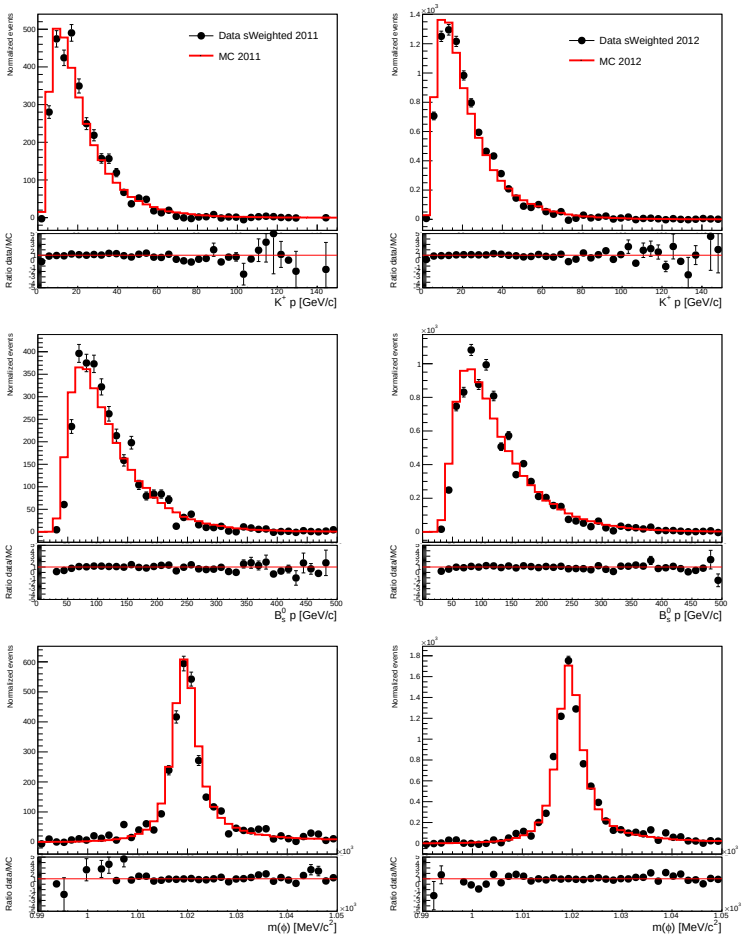


Fig. C.11 Distribution comparison of the $B_s^0 \to J/\psi\phi$ signal events in data (blue) and simulation (red) for 2011 and 2012.

C.6.1 Factorization of angular and decay time acceptance

One of the assumptions of the analysis is the factorization of the angular acceptance and the decay time acceptance. This assumption is verified by comparing the acceptance normalization weights obtained in five equally populated bins of the decay time (Fig. C.12). No dependence of the angular acceptance on the decay time is observed.

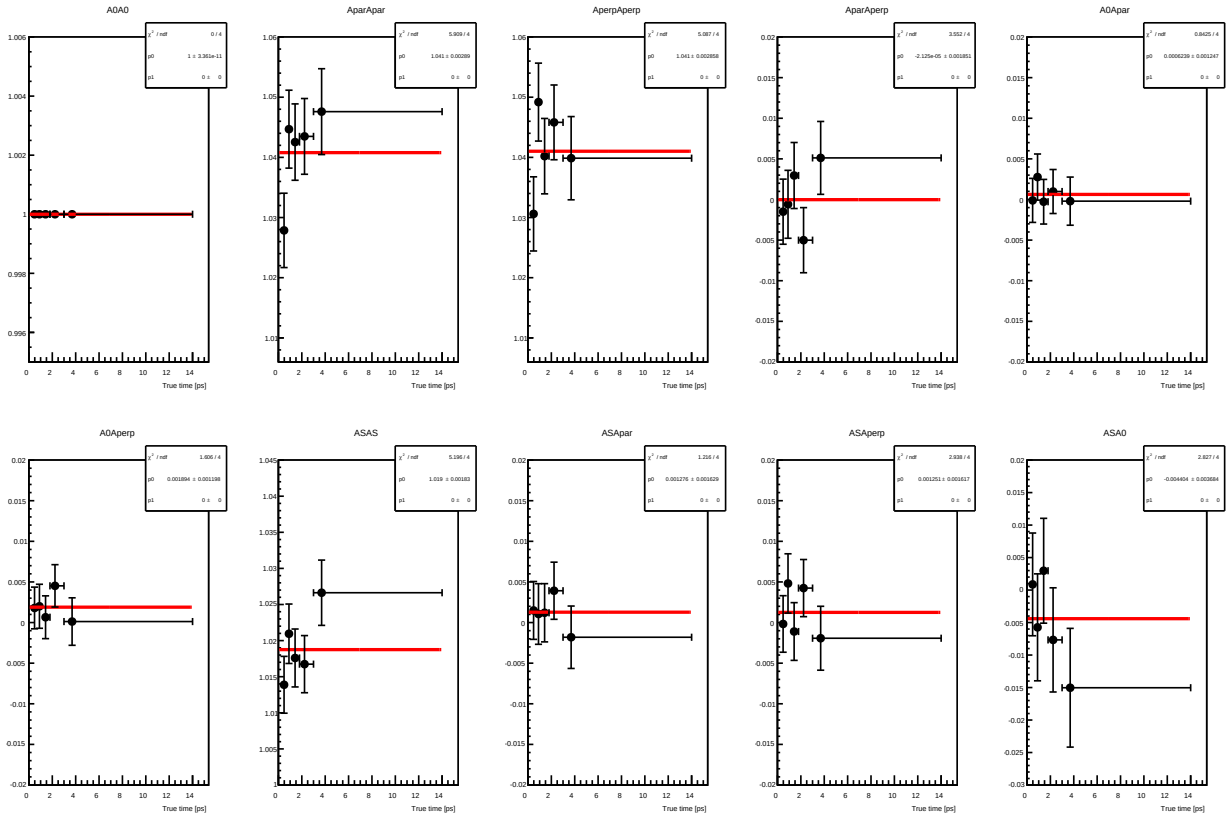


Fig. C.12 Normalization weights from the simulated signal sample evaluated as a function of the decay time.

C.7 Angular resolution

The two-dimensional distributions of the helicity angles are shown in Fig. C.13.

C.8 Time-angles distribution of $B_s^0 \to J/\psi(e^+e^-)\phi$ and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$

The comparison of the decay time and angular distributions of the $B_s^0 \to J/\psi(e^+e^-)\phi$ and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ decay modes is shown in Fig. C.14. The distributions of both modes are in an agreement except for the lower decay time.

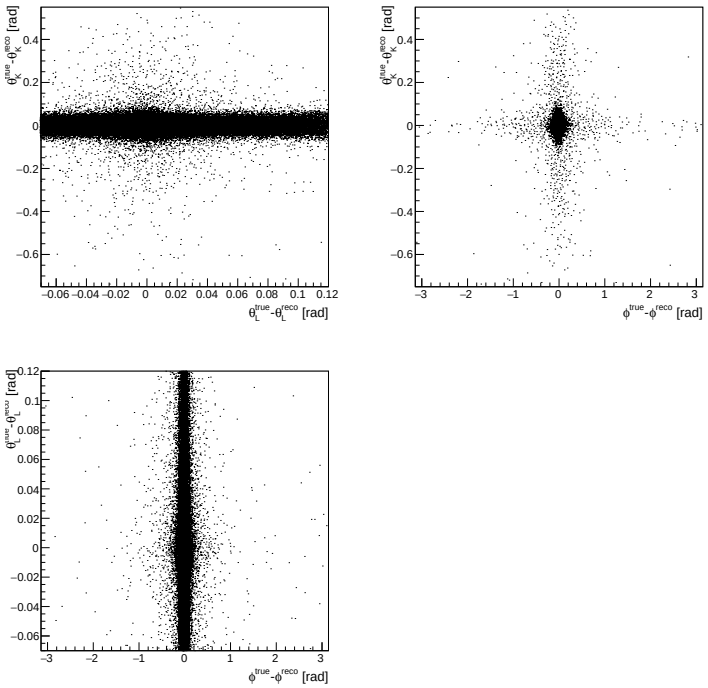


Fig. C.13 Two-dimensional distributions of the helicity angle resolutions obtained from simulated sample. The units are radians in all cases.

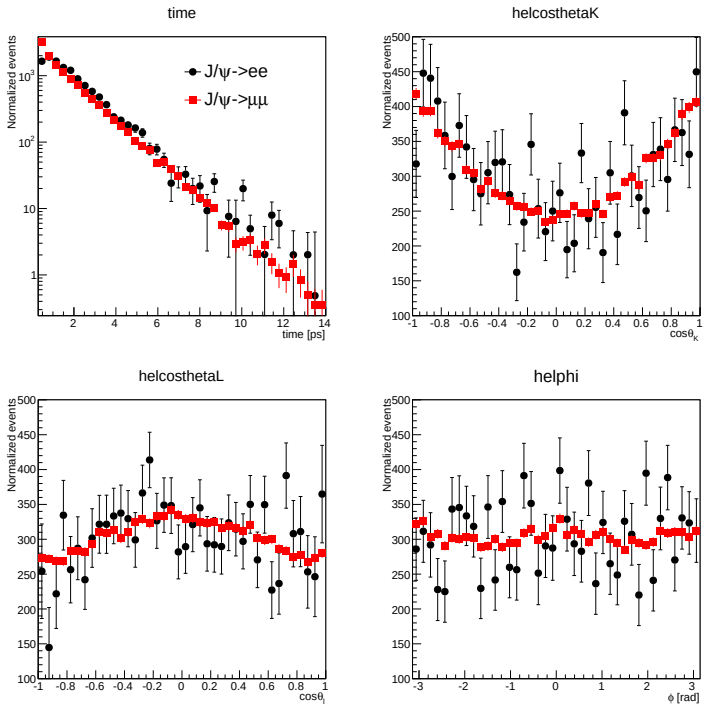


Fig. C.14 Distribution of the decay time and helicity angles for sWeighted $B_s^0 \to J/\psi(e^+e^-)\phi$ (black circle) and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ (red square) data samples.

C.9 SS calibration for flavour tagging

In order to reduce the systematic uncertainties due to transfer of SS neural network kaon (SSKNNet) tagger calibration, the kinematic variables from the $B_s^0 \to J/\psi\phi$ to $B_s^0 \to D_s^-\pi^+$ decay is reweighted using sWeighted data samples. The comparison is performed using five kinematic distributions: B_s^0 transverse momentum p_{T}, pseudorapidity η, azimuthal angle ϕ, the number of tracks and PVs as shown in Fig. C.15. The largest difference

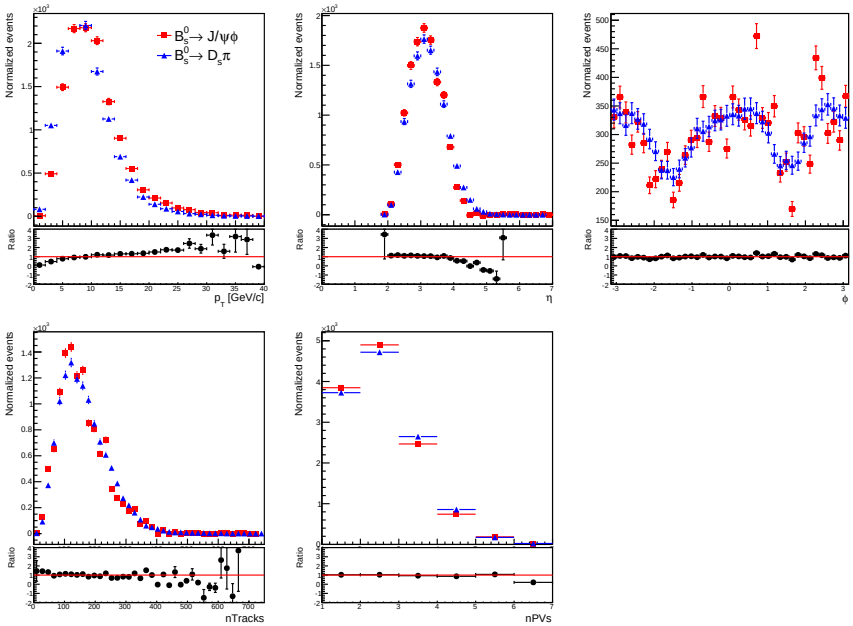


Fig. C.15 Comparison of kinematic distributions for sWeighted $B_s^0 \to J/\psi\phi$ to $B_s^0 \to D_s\pi$ data before reweighting.

in the kinematic variables is observed for the transverse momentum and η distributions. The two-dimensional reweighting is calculated using these two variables (Fig. C.16). The two-dimensional weight obtained from the ratio of the background subtracted $B_s^0 \to J/\psi\phi$ to $B_s^0 \to D_s\pi$ data is applied to the calibration data sample. The reweighted kinematic distributions are shown in Fig. C.17 where a better agreement between both modes is observed. Then, the SSKNNet calibration parameters are recalculated from the reweighted $B_s^0 \to D_s\pi$ data using `EspressoPerformanceMonitor`[3] tool [129]. The corrected probability ω_{SS} before and after two calibrations is shown in Fig. C.18. The previously used, called old, and new SSKNNet calibration parameters are listed in Table C.7. Finally, the obtained new SSKNNet parameters are used to determine the B_s^0 flavour of the analysed decay (Table 4.19).

The old SS tagging efficiency and tagging power of the $B_s^0 \to J/\psi\phi$ data sample are (63.47±1.12)% and (1.88±0.05)%, respectively. In case of new SS calibration, the tagging efficiency and tagging power are (63.47±1.12)% and (1.95±0.05)% which corresponds to a 0.07% increase of tagging power.

[3] `EspressoPerformanceMonitor` (`EPM`) tool provides functionality to calibrate and combine tagging algorithms.

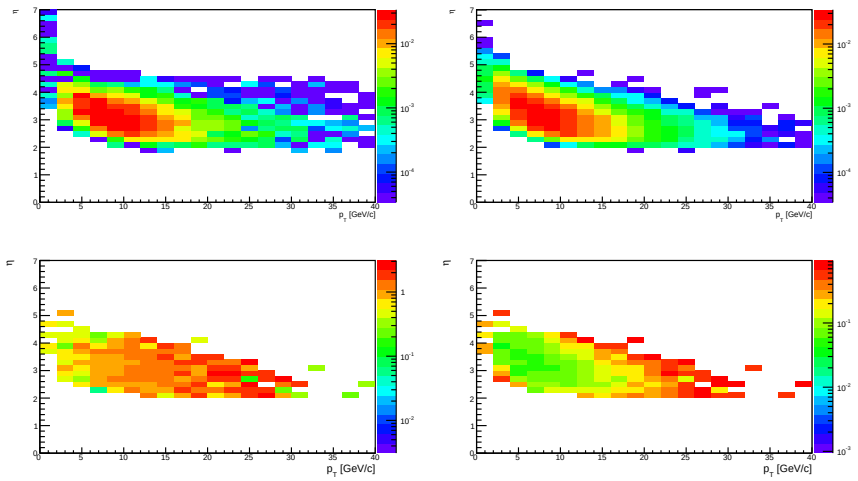


Fig. C.16 Two-dimensional distribution (p_{T} vs. η) of the sWeigted (top left) $B_s^0 \to J/\psi\phi$ and (top right) $B_s^0 \to D_s\pi$ data. The (bottom left) two-dimensional weight and (bottom right) error obtained from the ratio of the $B_s^0 \to J/\psi\phi$ to $B_s^0 \to D_s\pi$ data samples.

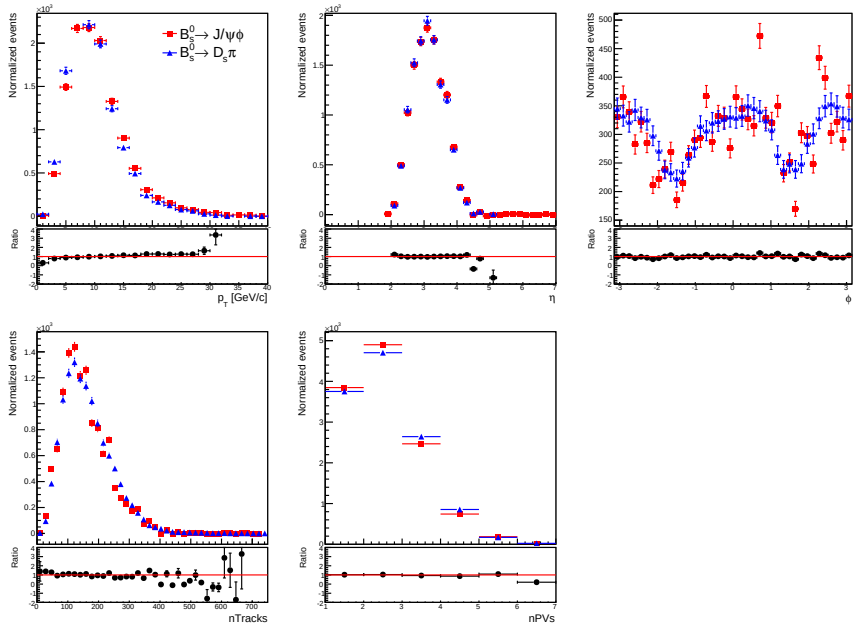


Fig. C.17 Comparison of kinematic distributions for sWeighted $B_s^0 \to J/\psi\phi$ and $B_s^0 \to D_s\pi$ data after reweighting.

C.10 Profile likelihood scans

Fig. C.19 shows scans of the negative log-likelihood for each fit parameter from the nominal fit. The scans were performed using fixed decay time resolution and average flavour tagging parameters. At each point the fit is repeated with all other parameters floating. The scans are generally parabolic in nature but larger asymmetries are seen for some parameters, notably λ, F_S and δ_S. As such, asymmetric statistical uncertainties are quoted for the final results.

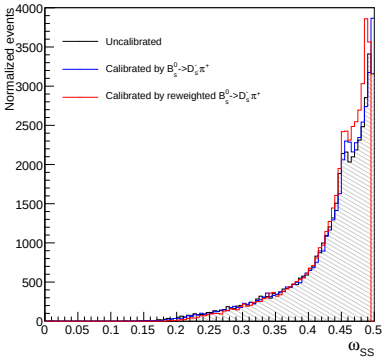


Fig. C.18 SSKNNet mistag distribution for the $B_s^0 \to J/\psi\phi$ signal events before (black) and after calibration (blue) and after calibration by reweighted data (red).

Table C.7 Two set of the SSKNNet parameters used in Eq. 4.17 to calibrate the per-event mistag probabilities. If two uncertainties are quoted, the first one is statistical and the second one is systematic.

Parameter	SSKNNet (old)	SSKNNet (new)
$\langle\eta\rangle$	0.4342	0.4349
$p_0 - \langle\eta\rangle$	0.005±0.004±0.003	0.0018±0.007
$\Delta p_0/2$	-0.0079±0.0007	-0.0047±0.0033
p_1	0.976±0.071±0.057	0.886±0.1036
$\Delta p_1/2$	0.0035±0.0111	0.0744±0.0518

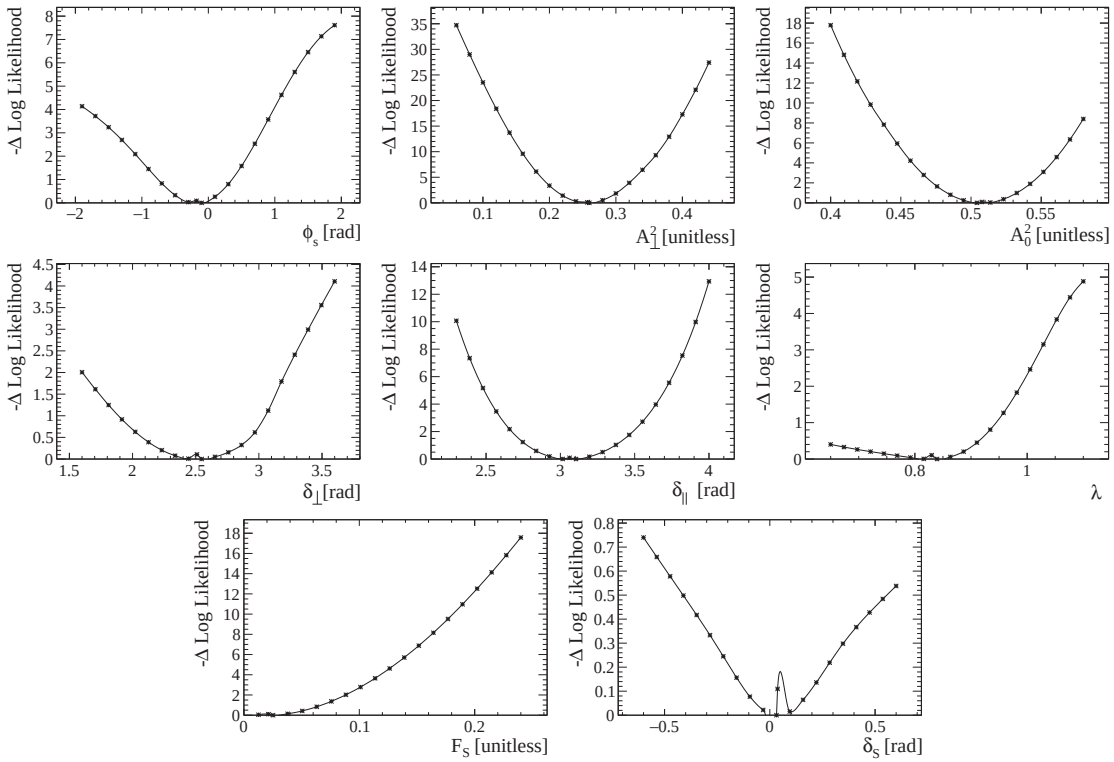


Fig. C.19 Log-likelihood scans of the fit parameters.

C.11 Factorization of $m(e^+e^-K^+K^-)$ distribution

The sWeighting procedure relies on the fact that the control variable $m(e^+e^-K^+K^-)$ used to determine the sWeights is uncorrelated with the variables used in the maximum likelihood fit (decay time, angles). Tables C.8-C.11 show the variation in the parameters of double CB mass model [101] when fitting $m(e^+e^-K^+K^-)$ dividing into Bremsstrahlung categories (0γ, 1γ and 2γ) in different bins of the helicity angles and decay time. The width and mean of the mass model is $\sigma_1 = \sigma_2$ and μ, respectively. The slope parameters $\alpha_{1,2}$, fraction f and number of degrees of freedom $n_1 = n_2 = n$ are fixed from a fit to the signal MC sample. The bins of the decay time and angles are chosen to have a similar number of events in each bin. Some dependence is observed as a function of all dimension, the strongest being as a function of $\cos\theta_e$.

Using the mass model in each bin of $\cos\theta_e$, the sWeights are recomputed for the full data sample and the time dependent angular fit repeated using the newly sWeighted dataset. The results of this fit are compared to the nominal physics parameter values (Table C.12) and any bias is assigned as a systematic uncertainty.

Table C.8 Mass model parameters from fits to the $m(e^+e^-K^+K^-)$ distribution in bins of $\cos\theta_e$ distributions.

Parameter	$\lvert\cos\theta_e\rvert \in[0,0.31]$	$\lvert\cos\theta_e\rvert \in(0.31,0.65]$	$\lvert\cos\theta_e\rvert \in(0.65,1.0]$
		0γ	
α_1	0.155±0.009	0.129±0.007	0.132±0.006
α_2	-0.65±0.5	-1.12±0.2	-0.56±0.4
f	0.998±0.002	0.983±0.009	0.998±0.003
n	2.9±0.2	4.9±0.7	17.0±4.0
σ[MeV/c^2]	30.6±4.8	32.5±5.8	23.2±6.7
μ[MeV/c^2]	5333.2±6.0	5340.8±5.8	5345.2±6.8
		1γ	
α_1	0.30±0.02	0.239±0.006	0.291±0.007
α_2	-0.639±0.03	-0.526±0.02	-0.465±0.02
f	0.75±0.01	0.750±0.008	0.797±0.009
n	4.9±0.7	22.0±5.0	29.0±5.0
σ[MeV/c^2]	61.3±5.5	55.6±6.1	43.4±6.5
μ[MeV/c^2]	5324.1±4.7	5334.4±4.6	5334.4±5.3
		2γ	
α_1	0.35±0.02	0.36±0.01	0.42±0.02
α_2	-0.524±0.03	-0.588±0.03	-0.564±0.03
f	0.57±0.02	0.58±0.02	0.63±0.02
n	11.0±4.0	20.0±6.0	20.0±6.0
σ[MeV/c^2]	64.0±12.0	56.1±9.4	82.0±19.0
μ[MeV/c^2]	5351.8±7.1	5360.3±6.0	5346.0±15.0

Table C.9 Mass model parameters from fits to the $m(e^+e^-K^+K^-)$ distribution in bins of $\cos\theta_K$ distributions.

Parameter	$\lvert\cos\theta_K\rvert \in$[0,0.4]	$\lvert\cos\theta_K\rvert \in$(0.4,0.73]	$\lvert\cos\theta_K\rvert \in$(0.73,1.0]
	0γ		
α_1	0.134±0.007	0.128±0.007	0.149±0.009
α_2	-1.28±0.3	-0.65±0.4	-0.67±0.3
f	0.985±0.009	0.997±0.004	0.996±0.003
n	6.0±1.0	6.3±1.0	3.6±0.3
σ[MeV/c^2]	19.1±4.2	25.6±4.4	41.9±8.3
μ[MeV/c^2]	5343.9±4.9	5345.3±5.0	5327.8±9.1
	1γ		
α_1	0.270±0.010	0.259±0.008	0.26±0.01
α_2	-0.554±0.02	-0.507±0.02	-0.549±0.02
f	0.76±0.01	0.768±0.009	0.754±0.010
n	15.0±4.0	19.0±5.0	10.0±2.0
σ[MeV/c^2]	48.7±6.7	51.3±7.2	57.1±5.9
μ[MeV/c^2]	5334.7±5.2	5332.4±5.8	5329.±4.7
	2γ		
α_1	0.37±0.02	0.39±0.02	0.36±0.01
α_2	-0.549±0.03	-0.594±0.03	-0.533±0.03
f	0.60±0.02	0.59±0.02	0.58±0.02
n	16.0±5.0	17.0±5.0	19.0±6.0
σ[MeV/c^2]	85.0±13.0	43.6±9.0	64.0±14.0
μ[MeV/c^2]	5348.6±9.9	5348.7±6.7	5368.1±7.7

Table C.10 Mass model parameters from fits to the $m(e^+e^-K^+K^-)$ distribution in bins of ϕ distribution.

Parameter	$\lvert\phi\rvert \in$[0,1.05]	$\lvert\phi\rvert \in$(1.05,2.10]	$\lvert\phi\rvert \in$(2.10,3.14]
	0γ		
α_1	0.129±0.007	0.147±0.008	0.135±0.008
α_2	-1.05±0.3	-0.65±0.3	-0.67±0.5
f	0.987±0.008	0.996±0.003	0.997±0.004
n	7.0±1.0	4.0±0.4	5.0±0.6
σ[MeV/c^2]	28.8±9.0	29.4±6.4	27.7±5.1
μ[MeV/c^2]	5338.9±6.2	5335.3±7.6	5345.5±5.7
	1γ		
α_1	0.263±0.009	0.25±0.01	0.27±0.01
α_2	-0.527±0.02	-0.551±0.02	-0.530±0.02
f	0.763±0.009	0.75±0.01	0.759±0.010
n	16.0±4.0	15.0±5.0	13.0±3.0
σ[MeV/c^2]	49.2±5.9	62.4±7.3	49.9±5.5
μ[MeV/c^2]	5332.4±5.	5331.3±5.3	5331.9±4.7
	2γ		
α_1	0.37±0.01	0.38±0.02	0.37±0.02
α_2	-0.511±0.03	-0.568±0.03	-0.601±0.03
f	0.60±0.02	0.60±0.02	0.57±0.02
n	20.0±6.0	61.0±1.0	15.0±5.0
σ[MeV/c^2]	66.0±14.0	67.0±15.0	70.0±15.0
μ[MeV/c^2]	5355.1±8.8	5359.8±7.4	5351.0±9.1

Table C.11 Mass model parameters from fits to the $m(e^+e^-K^+K^-)$ distribution in bins of decay time distribution.

Parameter	t ∈(0.3,0.95]	t ∈(0.95,2.0]	t ∈(2.0,14.0)
		0γ	
α_1	0.140±0.008	0.14±0.01	0.137±0.008
α_2	-0.78±0.3	-6.0±7.	-0.88±0.4
f	0.993±0.005	0.98±0.02	0.994±0.007
n	4.4±0.5	6.0±1.0	5.3±0.7
σ[MeV/c^2]	81.0±22.0	25.3±4.3	24.9±3.1
μ[MeV/c^2]	5309.0±15.0	5344.1±4.2	5340.±4.
		1γ	
α_1	0.26±0.01	0.256±0.008	0.27±0.01
α_2	-0.57±0.02	-0.533±0.02	-0.505±0.02
f	0.76±0.01	0.75±0.009	0.77±0.01
n	14.0±4.0	18.0±5.0	12.0±3.0
σ[MeV/c^2]	63.0±13.0	48.6±4.4	52.1±4.1
μ[MeV/c^2]	5325.3±9.1	5336.8±3.7	5329.2±3.5
		2γ	
α_1	0.37±0.02	0.36±0.02	0.39±0.02
α_2	-0.607±0.03	-0.562±0.03	-0.508±0.03
f	0.58±0.02	0.57±0.02	0.61±0.02
n	18.0±6.0	18.0±6.0	15.0±5.0
σ[MeV/c^2]	89.0±53.0	55.5±9.7	65.7±8.0
μ[MeV/c^2]	5373.0±19.0	5350.7±5.4	5348.4±5.3

Table C.12 Results of the unbinned maximum likelihood fit to the selected $B_s^0 \to J/\psi(e^+e^-)\phi$ candidates when sWeights are generated using a mass shape from a fit to $m(e^+e^-K^+K^-)$ in a single bin of $\cos\theta_e$. The last column shows the maximal difference with respect to the nominal fit result.

Parameter	Nominal	$\lvert\cos\theta_e\rvert$ ∈[0,0.31]	$\lvert\cos\theta_e\rvert$ ∈(0.31,0.65]	$\lvert\cos\theta_e\rvert$ ∈(0.65,1.0]	Max diff
$A_\perp^2$	0.258 ± 0.022	0.257 ± 0.022	0.257 ± 0.022	0.257 ± 0.023	0.001
A_0^2	0.508 ± 0.018	0.504 ± 0.017	0.508 ± 0.017	0.512 ± 0.018	0.004
$\delta_\parallel$ [rad]	3.06 ± 0.22	3.05 ± 0.22	3.06 ± 0.22	3.07 ± 0.21	0.01
$\delta_\perp$ [rad]	2.52 ± 0.45	2.6 ± 0.47	2.51 ± 0.46	2.45 ± 0.41	0.08
F_S	0.02 ± 0.029	0.011 ± 0.019	0.018 ± 0.027	0.035 ± 0.034	0.015
δ_S [rad]	0.05 ± 0.32	0.04 ± 0.38	0.04 ± 0.32	0.05 ± 0.25	0.01
ϕ_s [rad]	-0.19 ± 0.37	-0.18 ± 0.35	-0.18 ± 0.36	-0.19 ± 0.38	0.01
$\lvert\lambda\rvert$	0.83 ± 0.112	0.801 ± 0.149	0.825 ± 0.116	0.852 ± 0.081	0.029

C.12 $e^+e^-K^+K^-$ mass model

The fit result of the reconstructed $m(e^+e^-K^+K^-)$ distribution in selected data sample using a double Ipatia function fo the signal candidates is shown in Fig. C.20 and Table C.13. The time dependent angular fit is repeated using the newly sWeighted data sample. The result of this fit is compared to the nominal physics parameter values and any bias is assigned as a systematic uncertainty (Table C.14).

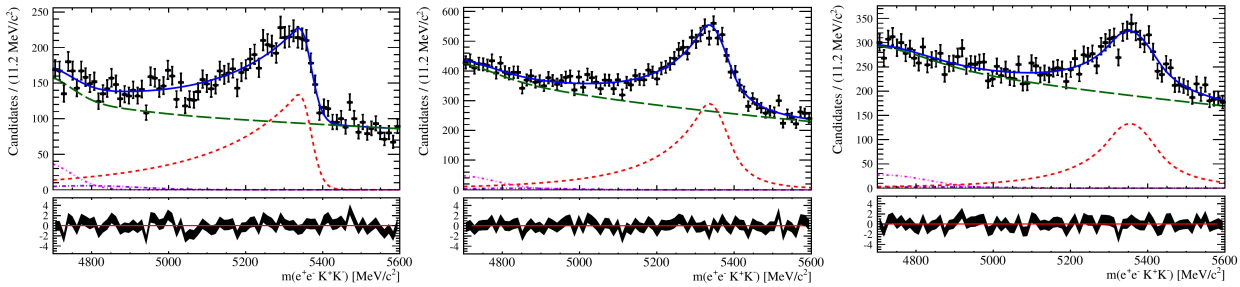


Fig. C.20 Fit to the reconstructed $e^+e^-K^+K^-$ mass distribution in data after selections using a double Ipatia function for the signal candidates, an exponential for the combinatorial background and a double and single Gaussian for partially reconstructed background. The dataset is divided into three Bremsstrahlung categories: (left) without, (middle) one and (right) more reconstructed photons coming from the electron radiation.

Table C.13 Results of the fit to the $m(e^+e^-K^+K^-)$ distribution in data sample using a double Ipatia function for the signal candidates with splitting into three Bremsstrahlung categories. The mass model parameters are defined as $n_1 = n_2 = n$ and $\beta = \zeta = 0$ and $\alpha_{1,2}$, λ and n are fixed to MC fit.

Parameter	0γ	1γ	2γ
α_1	0.027±0.003	0.12±0.02	0.21±0.01
α_2	0.62±0.03	0.36±0.06	0.32±0.02
λ	-10.377±0.06	-6.0±2.0	-5.72±0.2
n	4.7±0.3	3.29±0.08	3.3±0.1
σ[MeV/c^2]	123.0±15.0	180.0±13.0	244.0±30.0
μ[MeV/c^2]	5342.5±3.3	5335.4±2.4	5354.3±4.9

Table C.14 Results of the unbinned maximum likelihood fit to the selected $B_s^0 \to J/\psi(e^+e^-)\phi$ candidates when sWeights are generated using an alternative mass model function to $m(e^+e^-K^+K^-)$ distribution. The last column shows the difference with respect to the nominal fit result.

Parameter	Nominal	double Ipatia	Diff
$A_\perp^2$	0.257 ± 0.022	0.257 ± 0.022	0.000
A_0^2	0.508 ± 0.018	0.508 ± 0.018	0.000
$\delta_\parallel$ [rad]	3.06 ± 0.21	3.07 ± 0.21	0.01
$\delta_\perp$ [rad]	2.51 ± 0.44	2.48 ± 0.43	0.04
F_S	0.021 ± 0.029	0.024 ± 0.029	0.002
δ_S [rad]	0.04 ± 0.3	0.04 ± 0.3	0.00
ϕ_s [rad]	-0.18 ± 0.37	-0.17 ± 0.36	0.01
$\lvert\lambda\rvert$	0.83 ± 0.107	0.841 ± 0.096	0.011

C.13 Peaking background from $\Lambda_b \to J/\psi pK(\pi)$

The $\Lambda_b \to J/\psi(ee)pK(\pi)$ simulated data sample is used to determine the Λ_b background in the $m(e^+e^-K^+K^-(\pi^-))$ distribution in data sample. The shape of the Λ_b events is modelled by a double CB function. The MC sample of peaking background is divided into three Bremsstrahlung categories in the same way as the $B_s^0(B^0)$ data sample. The mass fit is

shown in Fig. C.21. The obtained fit parameters are reported in Table C.15 where the mean μ, width σ and fraction f are used in the mass fit performed on data sample. As estimated in Sec. 4.1.5 and 4.7.3 1.03%(0.85%) of Λ_b events is considered under the $B_s^0(B^0)$ mass peak. Using the mass model with Λ_b contribution, the sWeights are recomputed for the full data sample (Fig. C.22) and new decay time efficiency is calculated. Then, the time dependent angular fit is repeated using the newly sWeighted dataset and decay time acceptance. The results of this fit are compared to the nominal physics parameter values (Table C.16) and the shift is assigned as a systematic uncertainty.

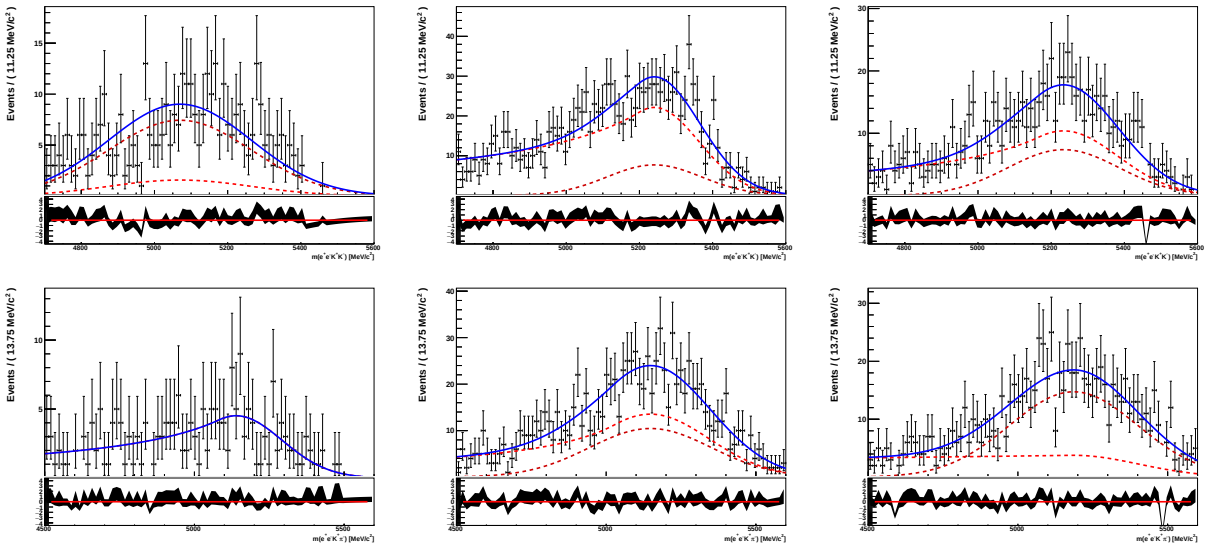


Fig. C.21 Distribution of (top) the $m(e^+e^-K^+K^-)$ for the $\Lambda_b \to J/\psi pK$ simulated data sample and (bottom) $m(e^+e^-K^+\pi^-)$ for the $\Lambda_b \to J/\psi p\pi$ simulated data sample divided into three Bremsstrahlung categories: (left) without, (middle) one and (right) more photons coming from the electron radiation.

C.14 Decay time resolution

The decay time resolution of the analysed decay has been determined from the maximum likelihood fit where the decay time resolution parameters are constrained to double difference of fit parameters of the $B_s^0 \to J/\psi(e^+e^-)\phi$ and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ simulated samples (Sec. 4.3.2). This constraint is increased 3 and 4 times to estimate the systematic contribution. The largest bias between the physics parameters determined using extended constraints of the decay time parameters and the nominal values is considered as a systematic error (Table C.17).

The result of the time dependent angular fit with and without $\Delta\Gamma_s$ and Γ_s constraint is reported in Table C.18.

Table C.15 Results of the fit to the $m(e^+e^-K^+K^-(\pi^-))$ distribution in the $\Lambda_b \to J/\psi pK(\pi)$ simulated data sample divided into three Bremsstrahlung categories. The shape is modelled by a double CB function.

Parameter	0γ	1γ	2γ
	$\Lambda_b \to J/\psi pK$		
α_1	3.0±4.0	0.33±0.05	0.44±0.09
α_2	-7.7±6.0	-8.3±2.0	-8.1±1.0
σ[MeV/c^2]	196.0±9.0	123.0±4.0	149.0±7.0
scale	1.0±0.05	1.05±0.07	1.02±0.07
μ[MeV/c^2]	5070.0±11.0	5241.0±6.0	5234.0±8.0
$f_{\rm CB1}$	0.2±0.3	0.81±0.04	0.66±0.07
n	1.0±97.0	1.0±23.0	1.0±23.0
$N_{\rm evt}^{\rm MC}$	380±19	1194±35	725±27
$N_{\rm evt}^{\rm data}$	32±3	63±4	30±3
	$\Lambda_b \to J/\psi p\pi$		
α_1	0.4±0.1	0.61±0.1	0.0±7.0
α_2	-8.4±5.0	-7.6±2.0	-3.2±8.0
σ[MeV/c^2]	153.0±13.0	197.0±8.0	215.0±57.0
scale	1.0±0.1	1.04±0.06	1.0±0.1
μ[MeV/c^2]	5139.0±19.0	5148.0±9.0	5186.0±19.0
$f_{\rm CB1}$	1.0±0.9	0.61±0.07	0.3±0.3
n	1.0±20.0	1.0±23.0	1.0±62.0
$N_{\rm evt}^{\rm MC}$	193±14	989±31	807±27
$N_{\rm evt}^{\rm data}$	71±1	249±3	150±3

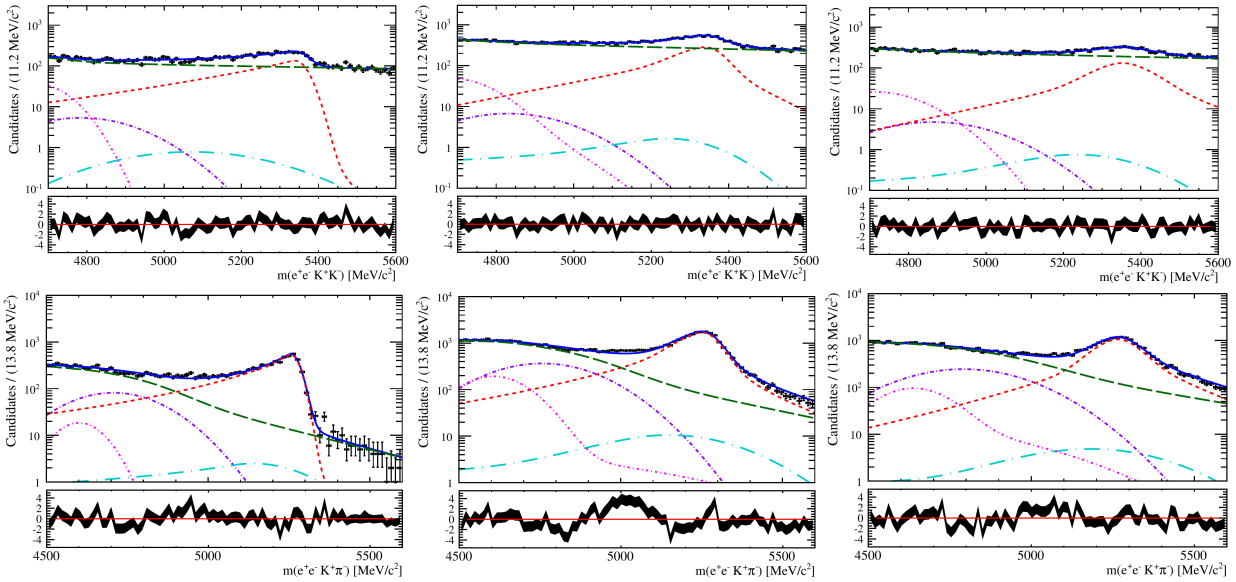


Fig. C.22 Distribution of (top) the $m(e^+e^-K^+K^-)$ for the $B_s^0 \to J/\psi\phi$ data sample with the contribution from $\Lambda_b \to J/\psi pK$ decay and (bottom) $m(e^+e^-K^+\pi^-)$ for the $B^0 \to J/\psi K^*$ data sample with the contribution from $\Lambda_b \to J/\psi p\pi$ decay (light blue dash-dotted line). The sample is divided into three Bremsstrahlung categories: (left) without, (middle) one and (right) more photons coming from the electron radiation. The blue line shows the total fit which is included the signal (red line) and combinatorial background (green line) contributions. A partially reconstructed background from the $B_s^0(B^0) \to \psi(2S)\phi(K^*)$ and $B_s^0(B^0) \to \chi_{c1}(1P)\phi(K^*)$ decays is indicated by pink and purple lines, respectively. The distributions are in a logarithmic scale.

Table C.16 Results of the unbinned maximum likelihood fit to the selected $B_s^0 \to J/\psi(e^+e^-)\phi$ candidates when sWeights and decay time acceptance are generated taken into account the Λ_b decay contribution. The last column shows the difference with respect to the nominal fit result.

Parameter	Nominal	Λ_b background	Diff
$A_\perp^2$	0.257 ± 0.022	0.254 ± 0.021	0.003
A_0^2	0.508 ± 0.018	0.478 ± 0.015	0.03
$\delta_\parallel$ [rad]	3.06 ± 0.21	2.99 ± 0.24	0.07
$\delta_\perp$ [rad]	2.51 ± 0.44	2.54 ± 0.48	0.03
F_S	0.021 ± 0.029	0.002 ± 0.003	0.019
δ_S [rad]	0.04 ± 0.3	0.05 ± 0.3	0.01
ϕ_s [rad]	-0.18 ± 0.37	-0.03 ± 0.29	0.15
$\lvert\lambda\rvert$	0.83 ± 0.107	0.786 ± 0.145	0.044

Table C.17 Results of the unbinned maximum likelihood fit to the selected $B_s^0 \to J/\psi(e^+e^-)\phi$ candidates when the constraint on the decay time resolution is changed. The last column shows the maximal difference with respect to the nominal fit result.

Parameter	Nominal	Constraint×3	Constraint×4	Max diff
$A_\perp^2$	0.257 ± 0.022	0.256 ± 0.023	0.258 ± 0.022	0.001
A_0^2	0.508 ± 0.018	0.508 ± 0.018	0.507 ± 0.018	0.001
$\delta_\parallel$ [rad]	3.06 ± 0.214	3.07 ± 0.22	3.06 ± 0.21	0.01
$\delta_\perp$ [rad]	2.51 ± 0.44	2.42 ± 0.53	2.63 ± 0.5	0.12
F_S	0.021 ± 0.029	0.029 ± 0.034	0.015 ± 0.033	0.008
δ_S [rad]	0.04 ± 0.3	0.07 ± 0.3	0.01 ± 0.39	0.03
ϕ_s [rad]	-0.18 ± 0.37	-0.22 ± 0.42	-0.14 ± 0.38	0.04
$\lvert\lambda\rvert$	0.83 ± 0.107	0.849 ± 0.088	0.799 ± 0.18	0.031

Table C.18 Results of the unbinned maximum likelihood fit to the selected $B_s^0 \to J/\psi(e^+e^-)\phi$ candidates with and without constraint on $\Delta\Gamma_s$ and Γ_s. The last column shows the difference with respect to the nominal fit result.

Parameter	Nominal	w/o constraint	Diff
Γ_s [ps^{-1}]	0.659 ± 0.003	0.615 ± 0.016	0.044
$\Delta\Gamma_s$ [ps^{-1}]	0.08 ± 0.009	0.117 ± 0.049	0.037
$A_\perp^2$	0.257 ± 0.022	0.236 ± 0.029	0.021
A_0^2	0.508 ± 0.018	0.519 ± 0.022	0.011
$\delta_\parallel$ [rad]	3.06 ± 0.21	3.05 ± 0.22	0.01
$\delta_\perp$ [rad]	2.51 ± 0.44	2.53 ± 0.44	0.02
F_S	0.021 ± 0.029	0.016 ± 0.024	0.006
δ_S [rad]	0.04 ± 0.3	0.04 ± 0.33	0.00
ϕ_s [rad]	-0.18 ± 0.37	-0.13 ± 0.28	0.05
$\lvert\lambda\rvert$	0.83 ± 0.107	0.821 ± 0.114	0.01

C.15 Sensitivity studies

In order to determine the expected parameter sensitivity and any potential biases in the maximum likelihood fitting procedure, a toy study is performed where the PDF is used to generate 550 datasets. The same number of signal events as in the data is used with the

same decay time and angular acceptance and physics parameter values set to those obtained from the nominal fit to the data. The flavour tagging parameters are fixed to their average values in data. Then, these datasets are fit using the same PDF. The parameter values and uncertainties are recorded from each fit and results are summarized in Table C.19 and Figs. C.23-C.24.

Table C.19 Results of a toy study that is used to evaluate the expected sensitivity for the physics parameters.

Parameter	Sensitivity	Pull mean	Pull width
$A_\perp^2$	0.009	-0.083±0.041	0.916±0.034
A_0^2	0.007	-0.010±0.045	0.971±0.035
$\delta_\parallel$ [rad]	0.126	-0.072±0.053	1.128±0.043
$\delta_\perp$ [rad]	0.109	0.048±0.041	0.908±0.033
F_S	0.005	-0.032±0.042	0.93±0.033
δ_S [rad]	0.104	-0.002±0.042	0.954±0.032
ϕ_s [rad]	0.049	0.039±0.045	0.985±0.037
$\lvert\lambda\rvert$	0.020	-0.022±0.046	1.03±0.03

C.16 Analysis of the subsets of the data sample

Tables C.20 to C.23 report the fit results split into different subcategories based on: run period, B_s^0 $p_{\rm T}$, magnet polarity and number of PVs. Fig. C.25 shows the decay time acceptance histograms evaluated in bins of number of PVs. All efficiency curves are consistent with the average reported in Fig. 4.21. Figs. C.26 to C.28 summarize each table by plotting the differences between the fit parameters in each bin with respect to the first bin, divided by the quadrature sum of the parameter uncertainties. The plots are distributed roughly as a Gaussian function, as expected for random variations within each data subset. The fit parameter values obtained in all subsets of the data are basically consistent.

Table C.20 Fit results in the run periods of 2011 and 2012.

Parameter	Nominal	2011	2012
$A_\perp^2$	0.257 ± 0.022	0.252 ± 0.048	0.261 ± 0.027
A_0^2	0.508 ± 0.018	0.487 ± 0.036	0.512 ± 0.021
$\delta_\parallel$ [rad]	3.06 ± 0.21	3.26 ± 0.22	3.03 ± 0.3
$\delta_\perp$ [rad]	2.51 ± 0.44	3.11 ± 0.22	2.43 ± 0.48
F_S	0.021 ± 0.029	0.025 ± 0.026	0.044 ± 0.039
δ_S [rad]	0.06 ± 0.3	-1.2± 0.37	0.35 ± 0.28
ϕ_s [rad]	-0.18 ± 0.37	1.07 ± 1.31	-0.21± 0.41
$\lvert\lambda\rvert$	0.83 ± 0.107	0.364 ± 0.326	0.899 ± 0.079

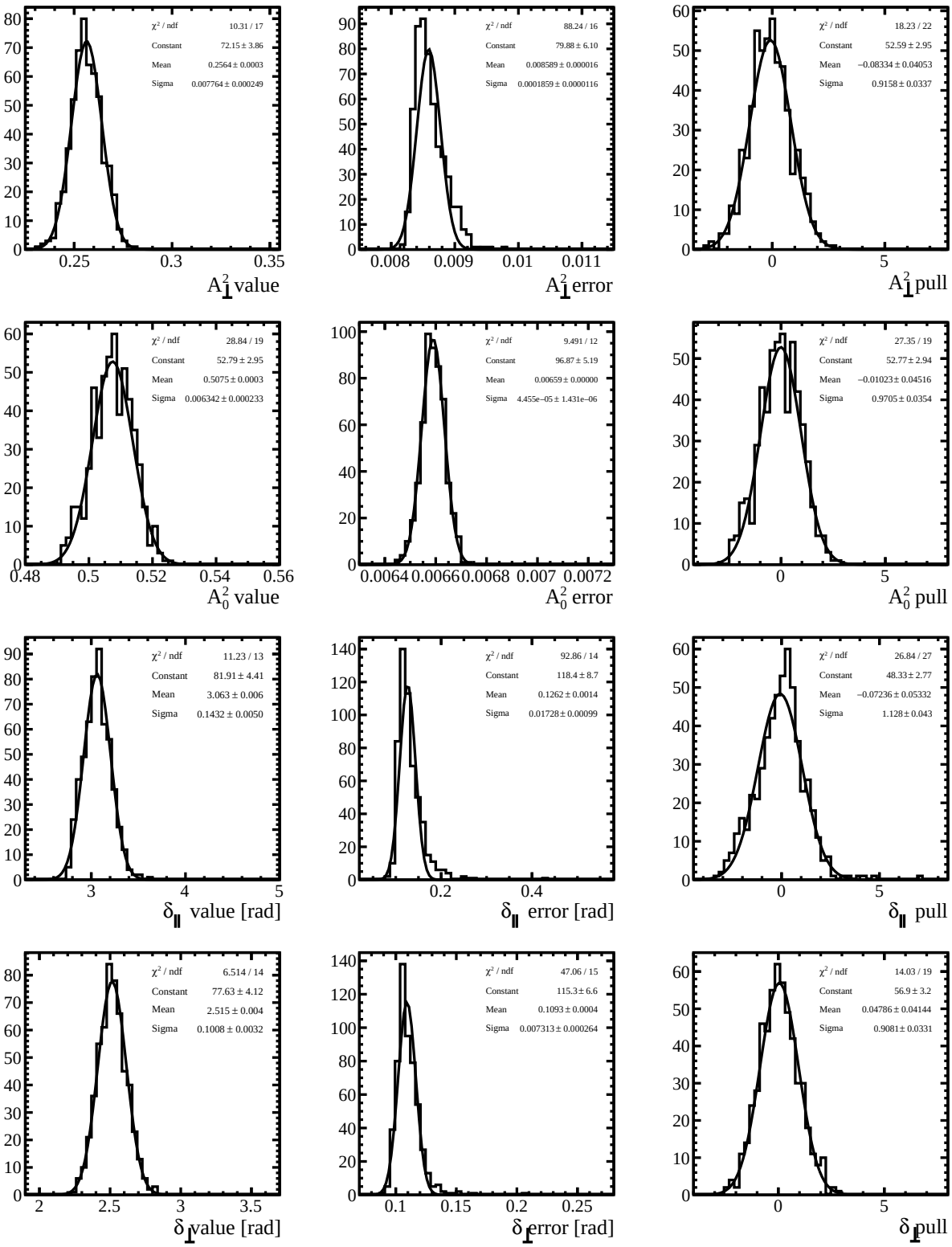


Fig. C.23 Distributions of (left) central value, (middle) error and (right) pull from 550 toy of potenial fit bias.

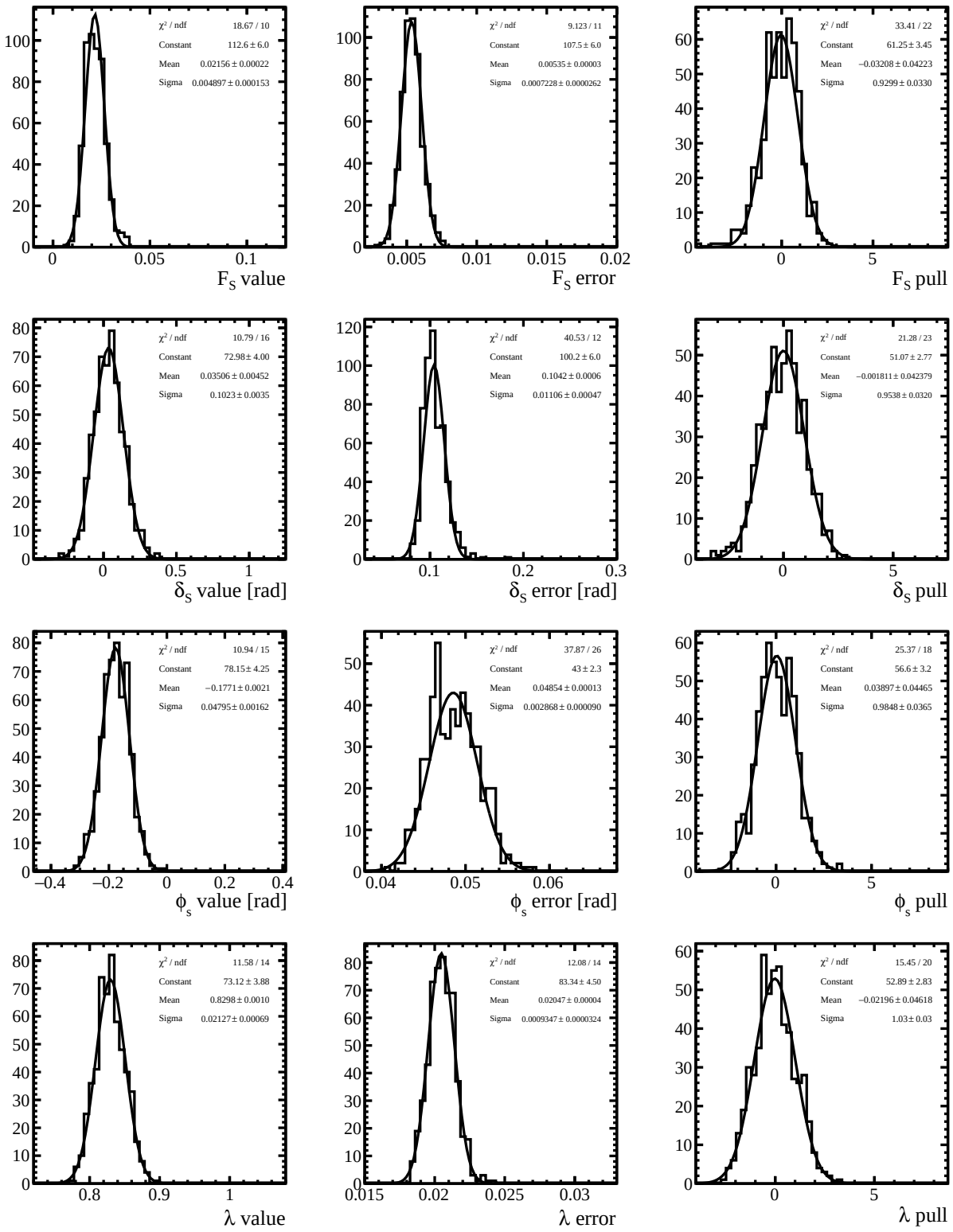


Fig. C.24 Distributions of (left) central value, (middle) error and (right) pull from 550 toy of potenial fit bias.

Table C.21 Fit results in the bins of magnet polarity.

Parameter	Nominal	Mag Up	Mag Down
$A_\perp^2$	0.257 ± 0.022	0.207 ± 0.034	0.286 ± 0.031
A_0^2	0.508 ± 0.018	0.541 ± 0.026	0.482 ± 0.023
$\delta_\parallel$ [rad]	3.06 ± 0.21	3.12 ± 0.23	2.88 ± 0.29
$\delta_\perp$ [rad]	2.51 ± 0.44	2.21 ± 0.44	2.92 ± 0.29
F_S	0.021 ± 0.029	0.058 ± 0.047	0.002 ± 0.005
δ_S [rad]	0.04 ± 0.3	-0.15 ± 0.42	0.82 ± 1.41
ϕ_s [rad]	-0.18 ± 0.37	0.13 ± 0.47	-0.22 ± 0.55
$\lvert\lambda\rvert$	0.83 ± 0.107	0.777 ± 0.099	0.642 ± 0.243

Table C.22 Fit results in the bins of the number of PVs.

Parameter	Nominal	1	2	3 or more
$A_\perp^2$	0.257 ± 0.022	0.255 ± 0.041	0.265 ± 0.033	0.259 ± 0.056
A_0^2	0.508 ± 0.018	0.535 ± 0.033	0.511 ± 0.026	0.463 ± 0.04
$\delta_\parallel$ [rad]	3.06 ± 0.21	3.17 ± 0.31	2.82 ± 0.28	3.16 ± 0.31
$\delta_\perp$ [rad]	2.51 ± 0.44	2.04 ± 0.5	2.42 ± 0.4	3.12 ± 0.26
F_S	0.021 ± 0.029	0.127 ± 0.061	10^{-8}±10^{-5}	0.009 ± 0.012
δ_S [rad]	0.04 ± 0.3	0.1 ± 0.21	-0.15 ± 0.8	-0.05 ± 0.23
ϕ_s [rad]	-0.18 ± 0.37	0.26 ± 0.5	-0.42 ± 0.42	-1.79 ± 1.57
$\lvert\lambda\rvert$	0.83 ± 0.107	0.894 ± 0.096	0.707 ± 0.17	0.345 ± 0.326

Table C.23 Fit results in the bins of the B_s^0 $p_{\rm T}$ where 1) $p_{\rm T}$<7. GeV/c; 2) $p_{\rm T}$ ∈[7.,10.5) GeV/c; 3) $p_{\rm T}$ ≥10.5 GeV/c.

Parameter	Nominal	1	2	3
$A_\perp^2$	0.257 ± 0.022	0.175 ± 0.062	0.267 ± 0.044	0.291 ± 0.025
A_0^2	0.508 ± 0.018	0.527 ± 0.046	0.49 ± 0.032	0.502 ± 0.019
$\delta_\parallel$ [rad]	3.06 ± 0.21	3.16 ± 0.28	2.93 ± 0.45	2.76 ± 0.24
$\delta_\perp$ [rad]	2.51 ± 0.44	2.19 ± 0.81	1.24 ± 1.08	2.75 ± 0.21
F_S	0.021 ± 0.029	0.042 ± 0.072	0.013 ± 0.023	0.004 ± 0.005
δ_S [rad]	0.04 ± 0.3	-0.27 ± 0.56	1.15 ± 1.85	0.68 ± 1.2
ϕ_s [rad]	-0.18 ± 0.37	-0.03 ± 0.9	-0.68 ± 0.79	-0.16 ± 0.36
$\lvert\lambda\rvert$	0.83 ± 0.107	0.766 ± 0.172	1.107 ± 0.151	0.537 ± 0.14

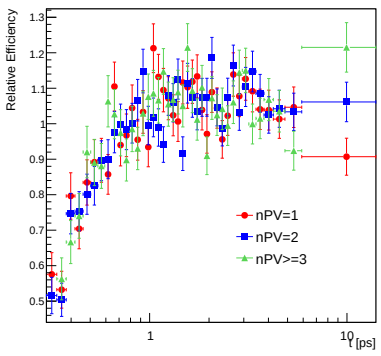


Fig. C.25 Decay time efficiency in bins of number of PVs defined in Eq. 4.7. The x-axis is in a logarithmic scale.

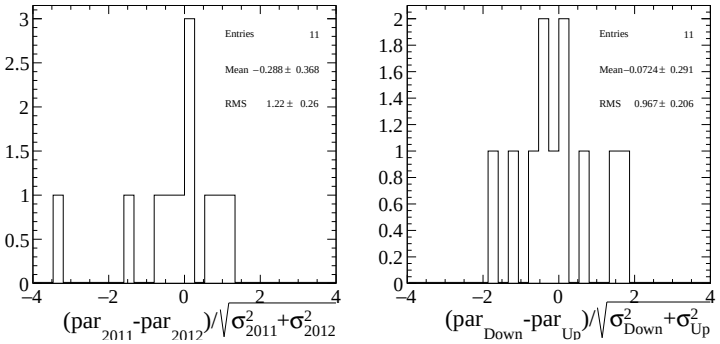


Fig. C.26 Distributions of the difference between fit results (left) in the 2011 and 2012 subsets and (right) in bins of magnet polarity. The summary of fit results in Tables C.20 and C.21.

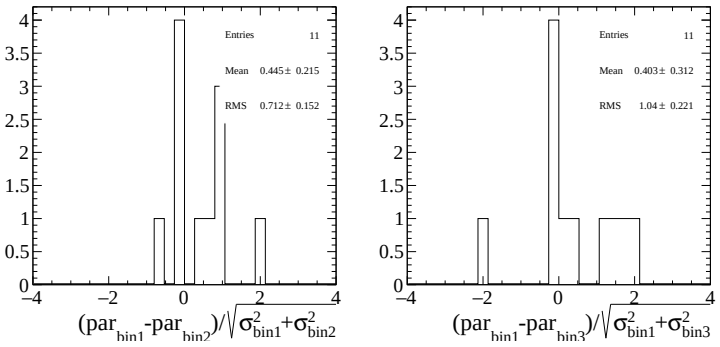


Fig. C.27 Distributions of the difference between fit results in bins of number of PVs, relative to the total uncertainty. The summary of fit results in Table C.22.

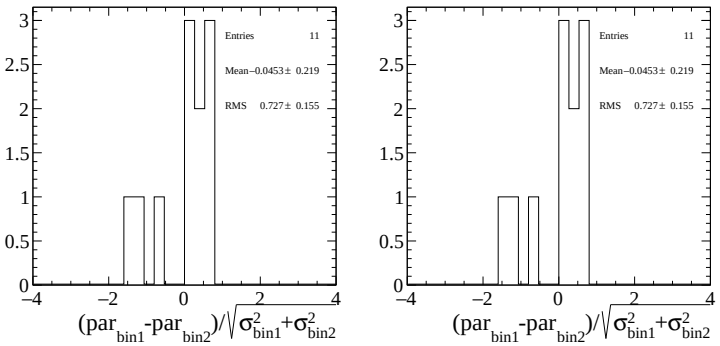


Fig. C.28 Distributions of the difference between fit results in bins of B_s^0 p_{T}, relative to the total uncertainty. The summary of fit results in Table C.23.

C.17 Closure test using simulated $B_s^0 \to J/\psi\phi$ sample

There are 298 804 signal candidates in the simulated sample after the full selection and truth matching. The unbinned maximum likelihood fit is performed with the angular (Table 4.17) and decay time acceptances (Fig. 4.20 middle) determined from this MC sample. The decay time resolution parameters are fixed to the MC fit result reported in Table 4.13. The flavour tagging parameters are Gaussian constrained to their average values in data. The fit result to the simulated $B_s^0 \to J/\psi\phi$ sample and difference in σ from nominal fit are shown in Table C.24. The measured value of F_S is negligible caused to the definition of the polarization amplitudes in the generation of the simulated signal events (Sec. 4.1). No bias is observed within the statistical uncertainties of the studied sample.

Table C.24 Result of the unbinned maximum likelihood fit to the simulated $B_s^0 \to J/\psi\phi$ signal candidates.

Parameter	Fit result	Diff(σ)
$A_\perp^2$	0.2523 ± 0.0022	-1.5
A_0^2	0.5154 ± 0.0016	+0.6
$\delta_\parallel$ [rad]	3.197 ± 0.066	+0.7
$\delta_\perp$ [rad]	2.976 ± 0.065	+1.3
F_S	$4.468{\cdot}10^{-6}$ ± $2.772{\cdot}10^{-5}$	-1.0
δ_S [rad]	1.02 ± 1.86	+0.5
ϕ_s [rad]	0.056 ± 0.019	+0.8
$\lvert\lambda\rvert$	1.01 ± 0.012	+2.1