Studies of time dependent CP violation in charm decays of B_s^0 mesons

Agnieszka Dziurda

The Henryk Niewodniczański Institute of Nuclear Physics
Polish Academy of Sciences
Krakow, Poland

Thesis submitted for the Degree of Doctor of Philosophy in Physics
Prepared under the supervision of
thesis supervisor: prof. dr hab. Tadeusz Lesiak
auxiliary supervisor: dr Vladimir Vava Gligorov

Krakow, April 2015 r.

"*Two things fill the mind with ever new and increasing admiration and awe,*
the oftener and the more steadily we reflect on them:
the starry heavens above and the moral law within."
– Immanuel Kant

Abstract

The thesis describes the world-first, time-dependent measurement of charge-parity (CP) violation in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. The study is performed at the LHCb experiment using data sample corresponding to an integrated luminosity of $1.0\,\text{fb}^{-1}$ of proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV, recorded in 2011. The CP violating observables are found to be: $C = 0.52 \pm 0.25 \pm 0.04$, $D_f = 0.29 \pm 0.42 \pm 0.17$, $D_{\bar{f}} = 0.14 \pm 0.41 \pm 0.18$, $S_f = -0.90 \pm 0.31 \pm 0.06$, $S_{\bar{f}} = -0.36 \pm 0.34 \pm 0.06$, where the first (second) uncertainty is statistical (systematic). These observables are used to perform the first measurement of the Cabibbo-Kobayashi-Maskawa angle γ in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays, finding $\gamma = (113^{+30}_{-44})^{\circ}$ modulo 180° at 68% CL, where the error contains both statistical and systematic uncertainties.

Streszczenie

Praca doktorska opisuje pierwszy na świecie, zależny od czasu pomiar łamania symetrii przestrzenno-ładunkowej (CP) w rozpadzie $B_s^0 \to D_s^{\mp} K^{\pm}$. Badania zostały przeprowadzone w eksperymencie LHCb używając próbki danych odpowiadającej $1.0\,\text{fb}^{-1}$ zderzeń proton-proton w układzie środka masy $\sqrt{s} = 7$ TeV, zarejestrowanych w 2011 roku. Parametry łamania CP zostały wyznaczone: $C = 0.52 \pm 0.25 \pm 0.04$, $D_f = 0.29 \pm 0.42 \pm 0.17$, $D_{\bar{f}} = 0.14 \pm 0.41 \pm 0.18$, $S_f = -0.90 \pm 0.31 \pm 0.06$, $S_{\bar{f}} = -0.36 \pm 0.34 \pm 0.06$, gdzie pierwsza (druga) niepewność oznacza przyczynek statystyczny (systematyczny). Parametry te zostały wykorzystane do wykonania pierwszego pomiaru kąta γ macierzy Cabibbo-Kobayashi-Maskawa, wyznaczając $\gamma = (113^{+30}_{-44})^{\circ}$ modulo 180° w przedziale 68% CL, gdzie błąd zawiera niepewność statystyczną jak i systematyczną.

Acknowledgements

I would like to express my gratitude to Tadeusz Lesiak, who many years ago agreed to be my master's and later doctoral thesis supervisor. Thank you for introducing me to the field of particle physics, showing me the exciting side of science and for the support over all of these years.

None of this work would have been possible without the support of my auxiliary supervisor Vladimir Vava Gligorov, who I owe my knowledge about the LHCb experiment. I am deeply grateful for always believing in me, his tireless patience and assistance in all moments when I needed it.

I would like to thank the Henryk Niewodniczański Institute of Nuclear Physics for a healthy work environment and KNOW for PhD scholarship. In addition, I am grateful to Mariusz Witek for the opportunity to work in the LHCb collaboration. I would like to also thank all institute members of the LHCb and ATLAS experiments for many inspired discussions.

My stays at CERN were always fruitful experiences. I would like to say thank you to all people who made them even more scientifically valuable as well as enjoyable. Special thanks go to the conveners of the B2OC working group: Till Moritz Karbach, Anton Poluektov, Marie-Helene Schune, the conveners of the time-dependent B2OC subgroup: Conor Fitzpatrick, Eduardo Rodrigues, Manuel Schiller and working group members: Suvayu Ali, Rose Koopman, Maximillian Schlupp, Giulia Tellarini and Stefania Vecchi. I am indebted to Manuel for his assistance over my stays at CERN, for inviting me to his office and, especially, his tireless support close to the publication deadlines.

This thesis couldn't be done without financial support. I would like to express my gratitude to National Science Center for partially funding my research in grant (DEC-2012/07/N/ST2/02890) and scholarship (DEC-2013/08/T/ST2/00035).

Last but not least, I would like to thank my family for their love and believing in me in any situation. In particular, I am indebted to my parents and brothers, who built my moral system. I owe You who I am. Thank You. Secondly, I need to thank my parents in-law for encouraging me to be simultaneously a perfect wife and scientist. Finally, my husband Grzegorz deserves for special thanks, without his never ending love, patience and understanding, I wouldn't be here.

Moritz, I know what the measurement of the CKM γ angle meant for you. It is too late for giving you my thesis, but I hope you would like it. Thank you for doing this research with me. Rest in peace.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Standard Model of particle physics describes the fundamental particles and their interactions. For the past half century this theory has been remarkably successful at predicting the behaviour of elementary particles. However there are several experimental observations that cannot be explained in the Standard Model. In particular, the theory does not explain the amount of visible asymmetry between matter and antimatter in the universe. One of the three Sakharov conditions [1] required for this asymmetry is charge-parity (*CP*) violation, which can be efficiently studied in the decays of beauty mesons.

This dissertation presents a measurement of *CP* violation in the process $B_s^0 \to D_s^{\mp} K^{\pm}$ using the LHCb detector at the Large Hadron Collider (LHC) at CERN, Geneva. The B_s^0 mesons[1] oscillate into each other during their propagation through space-time by the exchange of virtual particles. These are $W^{\pm}$ bosons in the Standard Model, but can be different in theories beyond the Standard Model. Because both decays are accessible for B_s^0 and $\overline{B}_s^0$ mesons, a measurement of *CP* violation is possible in the interference of mixing and decay. This *CP* violation is in turn sensitive to the angle γ of the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix. The analysis presented in this dissertation is performed for the first time in the world.

This thesis is organised in the following way: Chapter 2 presents the theoretical formalism which contains an introduction to the Standard Model and *CP* violation, a phenomenological model of beauty mesons decays and the experimental summary of the determination of the CKM angle γ using beauty mesons. Chapter 3 provides information about the LHCb experiment, in particular the parts of detector and software which are crucial in the measurement. This is followed by Chapter 4 where a general overview of the analysis is given. Chapter 5 discusses the selection necessary to obtain a pure $B_s^0 \to D_s^{\mp} h^{\pm}$ sample. Chapter 6 shows the signal and background discrimination achieved by the multidimensional fit performed to the $m(D_s^{\mp} h^{\pm})$ invariant mass, the $m(h^- h^+ h^{\pm})$ invariant mass, and the identification variable PIDK for the bachelor. This type of fit is used for the first time in the LHCb experiment. Next, ingredients crucial to perform the time-dependent measurement are presented: the reconstruction of initial flavour in Chapter 7 as well as decay-time resolution and acceptance in Chapter 8. The decay-time fit to $B_s^0 \to D_s^{\mp} h^{\pm}$ is discussed in Chapter 9, which is followed by a description of systematics effects in Chapter 10. Finally, the determination of the CKM angle γ is shown in Chapter 11 and the conclusion is given in Chapter 12.

[1] Unless otherwise stated, charge conjugate states are also considered.

The author of the presented dissertation also served as a contact author of the analysis in question inside the LHCb collaboration. Therefore, the author not only contributed to almost every step of the presented measurement but also checked and coordinated the work of other analysts. However, it is worth noting the author's personal contributions. The first of them is the Boosted Decision Trees selection trained specifically for this measurement and used by other analyses with similar decay chains Ref [2–6]. The next tasks were the optimisation of the offline selection and obtaining the data and simulation samples after all requirements. The author of this thesis was responsible for two major steps of the analysis: the multidimensional fit and the *sFit*. All work related to the multidimensional fit is also her direct personal contribution. In case of the flavour tagging a dedicated working group trains neural networks, which have to be calibrated for each measurement. Although, the calibration as well as the determination of the CKM angle γ were the tasks of another analysts, the author provided the *sFit* which was used in these studies. Furthermore, the signal and background PDFs created for the combined mistag and time uncertainty are a personal contribution of the author. In addition, the author adapted the spline technique for describing the decay-time acceptance and performed all the fits on simulated samples. The author wrote the *sFit* code for both the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ decay modes. In addition, the author is the main person who contributed to the standalone generator of pseudo experiments, and performed both the multidimensional fit and *sFit* validation studies. Finally, the author performed all pseudo experiment studies which were used in obtaining the systematics. More information about the author's activities can be found in the recommendation letters attached to this thesis.

Chapter 2

Theoretical formalism

This chapter provides a theoretical introduction to the topics studied in the dissertation. First a brief summary of the Standard Model is given. Next, the different types of symmetries in physics are discussed. An introduction to *CP* violation is followed by a phenomenological model of beauty mesons decays. In addition, different types of *CP* violation in this system are presented. Finally, the experimental summary of the determination of the CKM angle γ using beauty mesons is described together with the measurement strategy for $B_s^0 \to D_s^{\mp} K^{\pm}$. The following brief descriptions are based on [7–9].

2.1 The Standard Model

The Standard Model (SM) of particle physics is a renormalisable quantum field theory that describes the propagation and interactions of elementary particles. This theory combines the theory of strong interactions known as Quantum ChromoDynamics (QCD) [10,11], with the theory of electroweak interactions commonly labelled as the Glashow-Salam-Weinberg (GSW) model [12–14].

In the framework of the Standard Model matter is built from 12 particles called fermions which possess half-integer spin. Each fermion has a partner with opposite quantum numbers, referred to as its antiparticle. Fermions are split in two categories according to their interactions. The first group encompasses quarks, which carry colour charge and therefore interact via the strong interaction. Quarks come in six flavours: up u, charm c, top t (down d, strange s, beauty b) with electric charge of $\frac{2}{3}e$ $(-\frac{1}{3}e)$[2]. Besides the t quark[3], other quarks are bound inside hadrons. Hadrons comprise two groups: mesons which contain a quark-antiquark ($q\overline{q}$) pair and baryons made of three quarks (qqq). The second group of fermions contains leptons: three charged particles: electron e, muon μ, tau τ, and their neutral neutrino partners ν_e, ν_μ, ν_τ. The basic properties of the elementary fermions are shown in Tab. 2.1.

The general properties of the Standard Model originate from the symmetries of the direct product of the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge groups. Interactions between fermions are mediated by bosons which carry integer spin. The Standard Model depicts three of the four fundamental forces of nature, the electromagnetic, weak and strong interactions, but does not include gravity.

[2]The elementary charge e is the electric charge carried by a single electron.

[3]The lifetime of the top quark is so small that it decays before combining with other quarks to form hadrons.

Table 2.1: Basic properties of elementary fermions in the Standard Model.

Name	Spin [$\hbar$]	Generation			Electric Charge	Colour Charge
		1^{st}	2^{nd}	3^{rd}		
Quarks	$\pm\frac{1}{2}$	u	c	t	$\frac{2}{3}e$	{red,blue,green}
	$\pm\frac{1}{2}$	d	s	b	$-\frac{1}{3}e$	{red,blue,green}
Leptons	$\pm\frac{1}{2}$	e	μ	τ	$-e$	-
	$\pm\frac{1}{2}$	ν_e	ν_μ	ν_τ	0	-

Strong interactions are described in terms of the $SU(3)_C$ gauge symmetry group where the index C refers to the colour charge which takes three values labelled as red, blue and green, together with the respective anti-colours attributed to antiquarks. The gluons, massless particles which carry colour charge, are mediators of the strong interaction. Due to the fact that gluons themselves have colour charge they self-interact, which leads to the confinement phenomenon: the lack of observation of colourful states at distances exceeding the size of a nucleon.

The unification of weak and electromagnetic forces is described by the direct product of the $SU(2)_L \otimes U(1)_Y$ gauge groups. The electroweak bosons are massless W_1^+, W_2^-, W_3^0 $[SU(2)_L]$ and B^0 $[U(1)_Y]$ states. The $SU(2)_L \otimes U(1)_Y$ symmetry is spontaneously broken (SSB) to the $U(1)_Q$ group associated with electric charge via the so-called Higgs mechanism. The latter requires the presence of a doublet, scalar, complex Higgs field carrying four degrees of freedom. Three of them are absorbed by the electroweak bosons which become massive in this way. The remaining fourth field, the Higgs one, permeates all space. As particles pass through the Higgs field they acquire mass. The field is associated with an elementary particle called the Higgs boson which was observed in 2012 by the ATLAS [15] and CMS [16] collaborations.

In the presence of spontaneous symmetry breaking the electromagnetic force is mediated by the photon, γ, which is a linear combination of the third $SU(2)_L$ boson W_3^0 and the $U(1)_Y$ boson B^0. Due to the fact that the photon remains a massless particle after SSB, the range of the electromagnetic interaction is infinite.

The weak interaction is mediated by three observable bosons: two charged states $W^\pm$ and one neutral Z^0. The first two are linear combinations of $SU(2)_L$ vector bosons W_1^+ and W_2^-, whereas the last is, as the photon, a linear combination of W_3^0 and B^0. All three mediators are massive, therefore the range of the weak interaction is short. Weak interactions provide the only mechanism in the SM by which quarks can change their flavour.

The fundamental bosons of the Standard Model are listed in Tab. 2.2.

Table 2.2: Basic properties of elementary bosons in the Standard Model.

Name	Spin [$\hbar$]	Mass [GeV/c^2]	Force mediated
Photon γ	1	0	electromagnetic
Gluon g	1	0	strong
W^+, W^-	1	80.4	weak (charged current)
Z^0	1	91.2	weak (neutral current)
H^0	0	125.9 $\pm$ 0.4	-

2.2 Symmetries in physics

A physics system exhibits a definite symmetry if it remains the same after a certain transformation: e.g. translation, rotation etc. Each individual symmetry coincides with a well defined conserved quantity which is called a constant of motion. According to Noether's theorem [17] this parameter stays the same before and after the translation in question. In classical mechanics the indication of a symmetry is the invariance of the Lagrangian (Hamiltonian) describing the system with respect to the given transformation. For symmetries in quantum mechanics the Hamiltonian must commute with the respective operator which parametrises the symmetry operation. Group theory provides a mathematical description of symmetries in physics, in particular in the SM. Transformations of the system and the associated symmetries may be divided into continuous and discrete.

The most prominent examples of continuous symmetries are the conservation of energy, which follows from the invariance of physical laws under translations in time; the conservation of momentum, which appears because of invariance under translations in space, and angular momentum conservation due to invariance under spatial rotations.

The most relevant discrete symmetries are: parity P, charge conjugation C, and time reversal T. Each of them leads to the conservation of a multiplicative quantum number.

The first discrete symmetry, parity, changes the sign of the space coordinates $\vec{r} = (x, y, z)$ and, as a result, the handedness of the system of axes. The eigenfunction of the parity P operator satisfies the condition:

$$P\Psi(\overrightarrow{r}) = \Psi(-\overrightarrow{r}) = p\Psi(\overrightarrow{r}), \tag{2.1}$$

where p is an eigenvalue and Ψ denotes an eigenstate. A second application of this operator leads to the initial state, and so $p = \pm 1$. By convention the value $p = 1$ is assigned to elementary fermions and $p = -1$ to their antiparticles. The wave function Ψ can be separated into radial and angular parts:

$$\Psi(\overrightarrow{r}) = R(r)Y_l^m(\theta, \phi). \tag{2.2}$$

The parity operator does not change the spatial distribution R, but transforms a spherical harmonic into:

$$Y_l^m(\pi - \theta, \pi + \phi) = (-1)^L Y_l^m(\theta, \phi). \tag{2.3}$$

where L denotes the angular momentum, $l = 0, 1, 2, ...$ and $m = -l, -l + 1, ..., l - 1, l$. Parity is a multiplicative quantum number, so for mesons which are quark-antiquark bound states with an angular momentum L can be expressed as:

$$p = p_q p_{\overline{q}}(-1)^L = (-1)^{L+1}, \tag{2.4}$$

where $p_q(p_{\overline{q}})$ is the quark (antiquark) parity, and $p_q = -p_{\overline{q}}$.

The next symmetry is charge conjugation which changes the sign of all quantum charges. The eigenfunction of the charge conjugation C operator is:

$$C|\text{particle} >= |\text{antiparticle} >= \text{c}|\text{particle} >, \tag{2.5}$$

where c is again the respective eigenvalue with possible values ± 1. Thus only particles which are their own antiparticles can be eigenstates of C. Charge conjugation exchanges quarks with antiquarks. In consequence it inverts the relative position vector, which has

the same effect as the parity operation. In addition the spin wave function is also flipped. For a fermion-antifermion system this leads to the expression:

$$c = (-1)^{L+S}, \tag{2.6}$$

where L is the angular momentum and S is the spin of such a system.

Finally, the operator T reverses the direction of motion by reflection in the time axis. This anti-unitary operator has the form:

$$T\Psi(\overrightarrow{r}, t) = UK\Psi(\overrightarrow{r}, t) = \Psi^{\dagger}(\overrightarrow{r}, -t) \tag{2.7}$$

where U is a unitary operator and K denotes a complex conjugation.

The strong and electromagnetic interactions are invariant with respect to each of the C, P and T transformations. The weak interaction is not invariant under charge conjugation. The C and T operators change the left-handed neutrino into a left-handed antineutrino, which is not observed in nature, thus leading to the corollary that C symmetry is maximally violated in this interaction. P violation in the weak interaction was first observed as the spatial asymmetry of the decay products of $^{60}\mathrm{Co} \rightarrow ^{60}\mathrm{Ni}^{*} + \mathrm{e} + \overline{\nu_{\mathrm{e}}}$ [18], where the electrons favoured a direction of decay opposite to spin of the nuclei. Moreover in 1964 it was found that the CP symmetry (with respect to the product of C and P) is broken by weak interactions in the decays of neutral kaons [19]. In view of the CPT theorem which states that the product of C, P and T transformations should be an exact symmetry of an interaction that is described by a Lorentz invariant quantum field theory with a hermitian Hamiltonian, the violation of CP symmetry points unequivocally to the non-conservation of time reversal in weak interactions.

2.3 CP violation in the Standard Model

2.3.1 The Cabibbo-Kobayashi-Maskawa matrix

In the Standard Model CP violation is embedded in a single place, in processes mediated by $W^{\pm}$ exchange and described in the Lagrangian of the so-called weak charged currents. The quarks can be split into two groups according to their electric charges:

$$U = \begin{pmatrix} u \\ c \\ t \end{pmatrix}, \qquad D = \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.8}$$

The charged current is given by:

$$\frac{g}{\sqrt{2}}(\overline{U}\gamma^{\mu}V_{CKM}D_{L}W_{\mu}^{-} + \overline{D}_{L}\gamma^{\mu}V_{CKM}^{\dagger}U_{L}W_{\mu}^{+}), \tag{2.9}$$

where g is a coupling constant of weak interactions, W_{μ} corresponds to the field of the intermediate boson $W^{\pm}$, γ^{μ} are Dirac's matrices ($\mu = 0, 1, 2, 3$) and V_{CKM} is the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [20, 21]. The CKM mixing matrix reflects the fact that the flavour eigenstates (d', s', b') do not coincide with the mass eigenstates (d, s, b):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.10}$$

Here each matrix element V_{ij} corresponds to the flavour-changing weak coupling between the respective quarks q_i and q_j. The overall 3x3 matrix of complex elements is dependent on 18 real parameters. The fact that the CKM matrix is unitary reduces the number of its parameters by half. Moreover, five of these parameters are relative phases between the quark fields, which can be factored out. Therefore, in the case of three generations, the CKM matrix is parametrised by three real angles and one irreducible imaginary phase that is called the weak phase. The latter is the unique source of *CP* violation in the framework of the SM.

A commonly-used parametrisation for the CKM matrix elements is:

$$V_{CKM} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}c_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.11}$$

where $c_{ij} = cos(\theta_{ij})$, $s_{ij} = sin(\theta_{ij})$, θ_{ij} parametrises the size of mixing between quark generations i and j and δ refers to a weak phase.

The hierarchy of the moduli $|V_{ij}|$ which is observed experimentally can be presented in terms of powers of a small parameter $\lambda = s_{12} = sin(\theta_{12}) = 0.22$, which describes the mixing between the first and second generation of quarks (also known as the Cabibbo angle [20]):

$$|V_{CKM}| \sim \begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix}. \tag{2.12}$$

This leads directly to the so-called Wolfenstein parametrisation [22], which is shown here up to the third order in the λ parameter:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}, \tag{2.13}$$

where the parameters A, ρ, λ and η are related to Eq. 2.11 by the following expressions: $\lambda = s_{12}$, $A\lambda^2 = s_{23}$ and $A\lambda^3(\rho + i\eta) = s_{13}e^{i\delta}$. The imaginary parameter η quantifies the size of *CP* violating effects in the Standard Model.

2.3.2 The Unitarity Triangles

The unitarity of the CKM matrix ($V_{CKM}^{\dagger}V_{CKM} = 1$) implies six orthogonality conditions and requires that the products of different rows and columns of the V_{CKM} should sum up to zero. These orthogonality conditions can be represented as triangles in the complex plane, known as Unitarity Triangles (UT). The six relations together with the order of

magnitude of the length of each side of the triangle read:

$$V_{ud}^* V_{us} + V_{cd}^* V_{cs} + V_{td}^* V_{ts} = 0, \qquad (\lambda\ ,\quad \lambda\ ,\quad \lambda^5), \tag{2.14}$$
$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0, \qquad (\lambda^3,\quad \lambda^3,\quad \lambda^3), \tag{2.15}$$
$$V_{us}^* V_{ub} + V_{cs}^* V_{cb} + V_{ts}^* V_{tb} = 0, \qquad (\lambda^4,\quad \lambda^2,\quad \lambda^2), \tag{2.16}$$
$$V_{ud}^* V_{td} + V_{us}^* V_{ts} + V_{ub}^* V_{tb} = 0, \qquad (\lambda^3,\quad \lambda^3,\quad \lambda^3), \tag{2.17}$$
$$V_{td}^* V_{cd} + V_{ts}^* V_{cs} + V_{tb}^* V_{cb} = 0, \qquad (\lambda^4,\quad \lambda^2,\quad \lambda^2), \tag{2.18}$$
$$V_{ud}^* V_{cd} + V_{ds}^* V_{cs} + V_{ub}^* V_{cb} = 0, \qquad (\lambda\ ,\quad \lambda\ ,\quad \lambda^5). \tag{2.19}$$

Only the UTs given by Eq. 2.15 and Eq. 2.17 have all sides of the same order of magnitude. They both describe the triangle which satisfies the equation:

$$[(\rho + i\eta) + (-1) + (1 - \rho - i\eta)]\, A\lambda^3 = 0. \tag{2.20}$$

The UT described by Eq. 2.17 is experimentally less accessible due to transitions involving the top quark. Therefore, the first of them (Eq. 2.15, shown in Fig. 2.1) is commonly used and labelled as the Unitarity Triangle. Taking into account also terms proportional to λ^5, the sides of the UT are:

$$V_{ub}^* V_{ud} = A\lambda^3(\rho + i\eta)(1 - \frac{\lambda^2}{2}), \tag{2.21}$$
$$V_{cb}^* V_{cd} = -A\lambda^3, \tag{2.22}$$
$$V_{tb}^* V_{td} = A\lambda^3((1 - \rho - i\eta) + \lambda^2(\rho + i\eta)). \tag{2.23}$$

It is common to normalise the above equations by dividing them by the factor $A\lambda^3$. This leads to a useful geometrical interpretation of the UT in the complex plane in terms of the variables $(\overline{\rho}, \overline{\eta})$, where $\overline{\rho} = \rho(1 - \frac{\lambda^2}{2})$ and $\overline{\eta} = \eta(1 - \frac{\lambda^2}{2})$, with the angles of the triangle defined as:

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) = \arg\left(\frac{\left(1 - \frac{\lambda^2}{2}\right)(i\eta - \rho)}{1 - \rho - i\eta}\right), \tag{2.24}$$
$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) = \arg\left(\frac{1}{1 - \rho - i\eta}\right), \tag{2.25}$$
$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) = \arg\left(\left(1 - \frac{\lambda^2}{2}\right)(\rho - i\eta)\right). \tag{2.26}$$

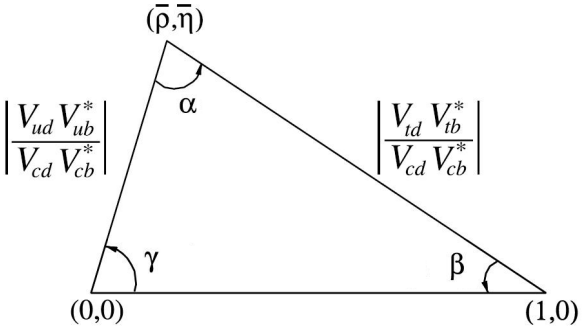

Figure 2.1: The Unitary Triangle as given by Eq. 2.15.

2.3.3 Current constraints on the CKM matrix

The values of the parameters of the CKM matrix are not predicted in the Standard Model, however, they are accessible experimentally. The precise determination of a set of observables which depend on V_{CKM} parameters would over-constrain the UT, thus providing one of the most important tests of the SM. In particular, any potential discrepancies in relation 2.15 would indicate contributions from physics beyond the SM. Figure 2.2 shows the results of the most recent measurements of the angles and sides of the UT. At the current level of precision there is no significant discrepancy, although the measurements of the CKM angle β and the matrix element V_{ub} are in some tension with each other. This can be seen from Fig. 2.3 where the two dimensional constraint between the measurements of $\sin(2\beta)$ and the branching fraction of the decay $B \to \tau\nu$, proportional to V_{ub}, is shown. Nevertheless, there is still no significant evidence for a departure from the Standard Model picture of *CP* violation in the UT measurements. The averages of the experimental results for the CKM angles from the CKMFitter group [23] are collected in Tab. 2.3. As it can be seen, the CKM angle γ angle, whose determination will be presented in this dissertation, is the least precisely measured of the CKM angles.

The CKM angle γ can be measured in many different decays of beauty hadrons as will be described in Sec. 2.4.4. In recent years measurements of this quantity have been performed by three experiments: Belle [24] at KEK [25](Japan), BaBar [26] at SLAC [27] (USA) and LHCb [28] at CERN [29] (Switzerland). Each of them has also provided its individual collaboration-wide average of the CKM angle γ measurement, whose results are listed in Tab. 2.4. It is worthwhile to underline that the average presented by the LHCb experiment [30] included the result of studies presented in this thesis.

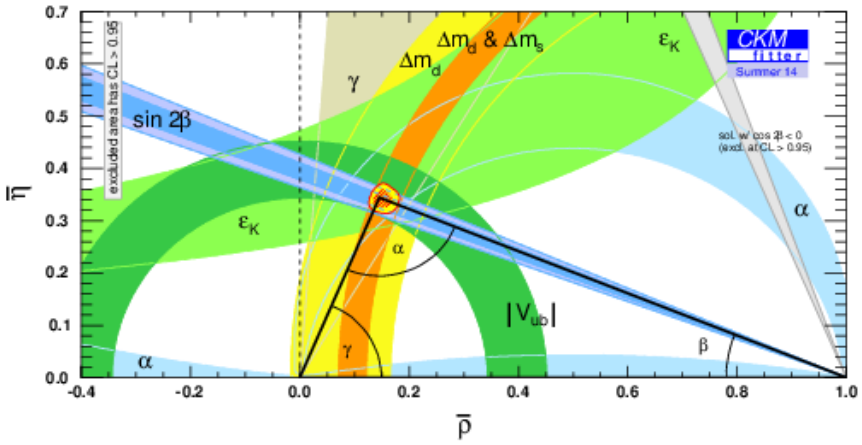

Figure 2.2: Status of measurements of sides and angles of the Unitary Triangle given by Eq. 2.15. Ref. [23].

Table 2.3: Estimates of the CKM angles α, β and γ as given by the CKMFitter group [23].

The CKM angle	Estimate from direct measurements [°]
α	$87.7^{+3.5}_{-3.3}$
β	$21.50^{+0.75}_{-0.74}$
γ	$73.2^{+6.3}_{-7.0}$

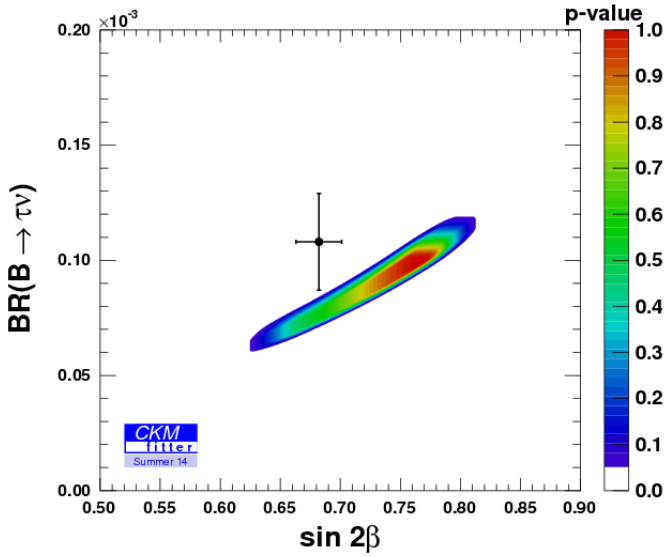

Figure 2.3: The experimental estimates of $\sin(2\beta)$ vs $\mathcal{B}(B \to \tau\nu)$ (data points) and the respective SM predictions (the area with colours corresponding to the p-values as given by the legend on the right side of the plot) Ref. [23].

Table 2.4: Recent combinations of the CKM angles γ made by the LHCb, Belle and BaBar collaborations Ref. [23].

Experiment	Combination of the CKM angle γ [°]
the LHCb Collaboration	$74.6^{+8.4}_{-9.2}$
the Belle Collaboration	73^{+13}_{-15}
the BaBar Collaboration	70 ± 18

2.4 Heavy flavour physics

Heavy flavour physics studies processes containing heavy elementary fermions such as the b, c quarks or the τ lepton.

2.4.1 Mixing of neutral mesons

In a system composed of a pair of neutral mesons, that differ only by a quantum number which is conserved in weak interactions, mixing arises as a result of a mismatch between their mass and flavour eigenstates. The following description focuses on the beauty meson mixing for which the flavour eigenstates are defined as:

$$B_q^0 = |\bar{b}q\rangle, \qquad \overline{B}_q^0 = |b\bar{q}\rangle, \tag{2.27}$$

where $q = \{d, s\}$. The convention $B_d^0 \equiv B^0$ is used. The mass eigenstates for the lighter (L) and heavier (H) components are a linear combination of the flavour eigenstates:

$$|B_L\rangle = p|B_q^0\rangle + q|\overline{B}_q^0\rangle, \qquad |B_H\rangle = p|B_q^0\rangle - q|\overline{B}_q^0\rangle, \tag{2.28}$$

with $|p|^2 + |q|^2 = 1$. Their inversions are given by:

$$|B_q^0\rangle = \frac{1}{2p}\left(|B_L\rangle + |B_H\rangle\right), \qquad |\overline{B}_q^0\rangle = \frac{1}{2q}\left(|B_L\rangle - |B_H\rangle\right). \tag{2.29}$$

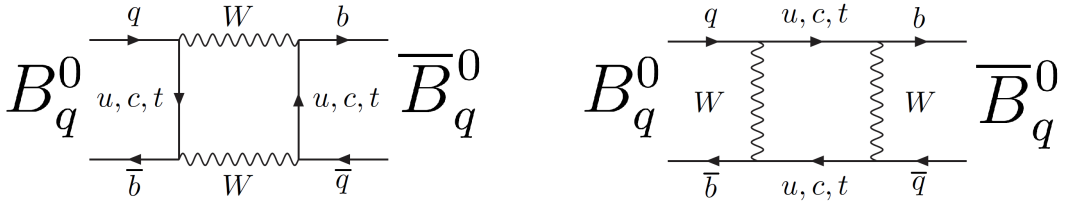

Figure 2.4: Feynman diagrams responsible for beauty meson mixing.

In the SM the beauty meson mixing process arises from so-called box-diagram transitions whose corresponding Feynman diagrams are shown in Fig. 2.4. The mesons B_q^0, $\overline{B}_q^0$ transform under CP symmetry as:

$$CP|B_q^0\rangle = -|\overline{B}_q^0\rangle, \qquad CP|\overline{B}_q^0\rangle = -|B_q^0\rangle. \tag{2.30}$$

The Schrödinger equation for the time evolution is given by:

$$i\frac{d}{dt}\begin{pmatrix}|B_q^0\rangle\\|\overline{B}_q^0\rangle\end{pmatrix} = \mathbf{H}\begin{pmatrix}|B_q^0\rangle\\|\overline{B}_q^0\rangle\end{pmatrix} = \left(\mathbf{M} - \frac{i}{2}\Gamma\right)\begin{pmatrix}|B_q^0\rangle\\|\overline{B}_q^0\rangle\end{pmatrix}, \tag{2.31}$$

where $\mathbf{H}$ denotes the Hamiltonian, while $\mathbf{M}$ and Γ are two-dimensional hermitian matrices describing the masses and widths of the system:

$$\mathbf{M} = \begin{pmatrix}m_{11} & m_{12}\\ m_{21} & m_{22}\end{pmatrix}, \qquad \Gamma = \begin{pmatrix}\Gamma_{11} & \Gamma_{12}\\ \Gamma_{21} & \Gamma_{22}\end{pmatrix}, \tag{2.32}$$

satisfying the relations $m_{21} = m_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$. Furthermore, CPT invariance implies that:

$$\langle B_q^0|\mathbf{H}|B_q^0\rangle = \langle\overline{B}_q^0|\mathbf{H}|\overline{B}_q^0\rangle, \tag{2.33}$$

which leads to the following additional constraints: $m = m_{11} = m_{12}$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$. As a result Eq. 2.31 reads:

$$i\frac{d}{dt}\begin{pmatrix}|B_q^0\rangle\\|\overline{B}_q^0\rangle\end{pmatrix} = \begin{pmatrix}m - \frac{i}{2}\Gamma & m_{12} - \frac{i}{2}\Gamma_{12}\\ m_{12}^* - \frac{i}{2}\Gamma_{12}^* & m - \frac{i}{2}\Gamma\end{pmatrix}\begin{pmatrix}|B_q^0\rangle\\|\overline{B}_q^0\rangle\end{pmatrix}. \tag{2.34}$$

The eigenvalues of $\mathbf{H}$ in the mass eigenstate basis can be expressed as:

$$\omega_{L,H} = m_{L,H} - \frac{i}{2}\Gamma_{L,H}, \tag{2.35}$$

where $m_{L,H}$ are the masses of lighter, heavier component, respectively and $\Gamma_{L,H}$ are their decay widths. In terms of the elements of matrices $\mathbf{M}$ and Γ, the eigenvalues are given by:

$$\omega_{L,H} = m - \frac{i}{2}\Gamma \mp \sqrt{\left(m_{12} - \frac{i}{2}\Gamma_{12}\right)\left(m_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}, \tag{2.36}$$

which can be rewritten as:

$$\omega_{L,H} = \left(m \mp \frac{\Delta m}{2}\right) - \frac{i}{2}\left(\Gamma \pm \frac{\Delta\Gamma}{2}\right) \tag{2.37}$$

with the following relations:

$$m = \frac{m_H + m_L}{2}, \qquad \Gamma = \frac{\Gamma_H + \Gamma_L}{2},$$
$$\Delta m = m_H - m_L, \qquad \Delta\Gamma = \Gamma_L - \Gamma_H. \tag{2.38}$$

Finally, the matrix from Eq. 2.34 can be expressed in terms of $\omega_{L,H}$ as:

$$\begin{pmatrix} m - \frac{i}{2}\Gamma & m_{12} - \frac{i}{2}\Gamma_{12} \\ m^*_{12} - \frac{i}{2}\Gamma^*_{12} & m - \frac{i}{2}\Gamma \end{pmatrix} = \begin{pmatrix} \frac{1}{2|q|^2} & 0 \\ 0 & \frac{1}{2|p|^2} \end{pmatrix} \begin{pmatrix} q^* & q^* \\ p^* & -p^* \end{pmatrix} \begin{pmatrix} \omega_L & 0 \\ 0 & \omega_H \end{pmatrix} \begin{pmatrix} q & p \\ q & -p \end{pmatrix}$$
$$= \begin{pmatrix} \omega & -\frac{p}{2q}\Delta\omega \\ -\frac{p}{2q}\Delta\omega & \omega \end{pmatrix}, \tag{2.39}$$

where $\omega = \frac{1}{2}(\omega_H + \omega_L)$ and $\Delta\omega = \omega_H - \omega_L$. Moreover, since the off-diagonal elements of the matrix from Eq. 2.39 coincide with the parameters p and q (Eq. 2.28) it gives the relation:

$$\frac{p}{q} = \sqrt{\frac{m_{12} - \frac{i}{2}\Gamma_{12}}{m^*_{12} - \frac{i}{2}\Gamma^*_{12}}}. \tag{2.40}$$

2.4.2 Time evolution of neutral meson states

The time evolution of mass eigenstates is given by the relations:

$$|B_L(t)\rangle = e^{-\left(im_L + \frac{1}{2}\Gamma_L\right)t}|B_L\rangle, \qquad |B_H(t)\rangle = e^{-\left(im_H + \frac{1}{2}\Gamma_H\right)t}|B_H\rangle. \tag{2.41}$$

Using Eq. 2.28 and Eq. 2.41, the latter can be transformed to:

$$|B^0_q(t)\rangle = g_+(t)|B^0_q(0)\rangle + \frac{q}{p}g_-(t)|B^0_q(0)\rangle,$$
$$|\overline{B}^0_q(t)\rangle = \frac{p}{q}g_-(t)|B^0_q(0)\rangle + g_+(t)|B^0_q(0)\rangle, \tag{2.42}$$

where the functions $g_\pm(t)$ read:

$$g_\pm(t) = \frac{1}{2}\left(e^{-im_L t}e^{-\frac{1}{2}\Gamma_L t} \pm e^{-im_H t}e^{-\frac{1}{2}\Gamma_H t}\right). \tag{2.43}$$

Due to mixing, both B^0_q and $\overline{B}^0_q$ mesons can decay to the same final state which will be denoted as f. In the following, its charge conjugation will be marked as $\overline{f}$. There are four possible decays $B^0_q \to f$, $B^0_q \to \overline{f}$, $\overline{B}^0_q \to f$ and $\overline{B}^0_q \to \overline{f}$ and their decay amplitudes are given by:

$$A_f = \langle f|\mathbf{H}|B^0_q\rangle, \quad A_{\overline{f}} = \langle \overline{f}|\mathbf{H}|B^0_q\rangle, \quad \overline{A}_f = \langle f|\mathbf{H}|\overline{B}^0_q\rangle, \quad \overline{A}_{\overline{f}} = \langle \overline{f}|\mathbf{H}|\overline{B}^0_q\rangle, \tag{2.44}$$

respectively. In the presence of CPT invariance the following conditions are satisfied:

$$|\overline{A}_{\overline{f}}| = |A_f|, \qquad |\overline{A}_f| = |A_{\overline{f}}|. \tag{2.45}$$

The decay rate of a $|B^0_q\rangle$ meson produced at time $t = 0$ to a final state f at time t is given by the formula:

$$\frac{d\Gamma_{B^0_q \to f}(t)}{dt} = |\langle f|\mathbf{T}|B^0_q(t)\rangle|^2, \tag{2.46}$$

where $\mathbf{T}$ corresponds to the respective transition matrix. Similar expressions can be defined for the other three decay amplitudes described in Eq. 2.44. Using Eq. 2.42, the decay rate of the $B_q^0 \to f$ decay is expressed by the formula:

$$\begin{aligned}\frac{d\Gamma_{B_q^0\to f}(t)}{dt} &= |\langle f|\mathbf{T}|B_q^0(t)\rangle|^2 = |g_+(t)A_f + \frac{q}{p}g_-\overline{A}_f|^2 \\ &= |A_f|^2\left(|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 + \lambda_f^* g_+(t)g_-^*(t) + \lambda_f g_+^*(t)g_-(t)\right) \\ &= \frac{1}{2}|A_f|^2 e^{-\Gamma t}\Bigg[\left(1+|\lambda_f|^2\right)\cosh\left(\frac{\Delta\Gamma}{2}t\right) + \left(1-|\lambda_f|^2\right)\cos\left(\Delta m t\right) \\ &\qquad - 2\Re\left(\lambda_f\right)\sinh\left(\frac{\Delta\Gamma}{2}t\right) - 2\Im(\lambda_f)\sin\left(\Delta m t\right)\Bigg]. \end{aligned} \tag{2.47}$$

The respective term for the $\overline{B}_q^0 \to f$ decay is:

$$\begin{aligned}\frac{d\Gamma_{\overline{B}_q^0\to f}(t)}{dt} &= |\langle f|\mathbf{T}|\overline{B}_q^0(t)\rangle|^2 \\ &= \frac{1}{2}\left|\frac{p}{q}\right|^2 |A_f|^2 e^{-\Gamma t}\Bigg[\left(1+|\lambda_f|^2\right)\cosh\left(\frac{\Delta\Gamma}{2}t\right) - \left(1-|\lambda_f|^2\right)\cos\left(\Delta m t\right) \\ &\qquad - 2\Re\left(\lambda_f\right)\sinh\left(\frac{\Delta\Gamma}{2}t\right) + 2\Im(\lambda_f)\sin\left(\Delta m t\right)\Bigg]. \end{aligned} \tag{2.48}$$

In the above the parameter λ_f is defined by:

$$\lambda_f = \frac{1}{\overline{\lambda}_f} = \frac{q}{p}\frac{\overline{A}_f}{A_f}. \tag{2.49}$$

The rates of the decays to the charge conjugate final state $\overline{f}$ are given by the same equations but after substituting the index f by $\overline{f}$ and with the parameter $\lambda_{\overline{f}}$ given by:

$$\lambda_{\overline{f}} = \frac{1}{\overline{\lambda}_{\overline{f}}} = \frac{q}{p}\frac{\overline{A}_{\overline{f}}}{A_{\overline{f}}}. \tag{2.50}$$

Based on the above discussion, the following CP asymmetry observables C_f, S_f, D_f ($C_{\bar{f}}$, $S_{\bar{f}}$, $D_{\bar{f}}$) for f ($\overline{f}$) final states are defined:

$$\begin{aligned} C_f &= \frac{1-|\lambda_f|^2}{1-|\lambda_f|^2}, \qquad S_f = \frac{2\Im(\lambda_f)}{1-|\lambda_f|^2}, \qquad D_f = \frac{-2\Re(\lambda_f)}{1-|\lambda_f|^2}, \\ C_{\bar{f}} &= \frac{1-|\lambda_{\overline{f}}|^2}{1-|\lambda_{\overline{f}}|^2}, \qquad S_{\bar{f}} = \frac{2\Im(\lambda_{\overline{f}})}{1-|\lambda_{\overline{f}}|^2}, \qquad D_{\bar{f}} = \frac{-2\Re(\lambda_{\overline{f}})}{1-|\lambda_{\overline{f}}|^2}. \end{aligned} \tag{2.51}$$

Here CPT invariance yields $C = C_f = -C_{\bar{f}}$. The relevant SM parameters such as the weak phase $\gamma - 2\beta_q$ and the strong phases δ, can be extracted using the relations:

$$\gamma - 2\beta_q = \frac{1}{2}[\arg(\overline{\lambda}_{\overline{f}}) - \arg(\lambda_f)], \qquad \delta = \frac{1}{2}[\arg(\overline{\lambda}_{\overline{f}}) + \arg(\lambda_f)], \tag{2.52}$$

where β_q denotes the phase associated with mixing. Finally, the *CP* parameters defined in Eq. 2.51 exhibit the following dependence on the SM parameters:

$$\begin{aligned} C &= \frac{1-r^2}{1-r^2}, \\ S_f = \frac{2r\sin(\delta-(\gamma-2\beta_s))}{1-r^2}, &\qquad D_f = \frac{-2r\cos(\delta-(\gamma-2\beta_s))}{1-r^2}, \\ S_{\bar{f}} = \frac{-2r\sin(\delta+(\gamma-2\beta_s))}{1-r^2}, &\qquad D_{\bar{f}} = \frac{-2r\cos(\delta+(\gamma-2\beta_s))}{1-r^2}, \end{aligned} \tag{2.53}$$

where r is the ratio of the magnitudes of the decay amplitudes of the contributing diagrams. As can be seen in the above relations, a small value of r limits the sensitivity of the measurement, whereas maximal *CP* violation sensitivity is expected when the contributing diagrams have the same order of magnitude.

2.4.3 Types of *CP* violation

CP violation phenomena can be categorised into three types:

- in decay,
- in mixing,
- in the interference between decay and mixing.

CP violation in decay occurs for charged and neutral mesons. This type of *CP* violation happens when the decay probability of $B_q^0 \to f$ is not equal to that of $\overline{B}_q^0 \to \overline{f}$. It implies that $\left|\frac{\overline{A}_{\overline{f}}}{A_f}\right| \neq 1$. The ratio $\left|\frac{\overline{A}_{\overline{f}}}{A_f}\right|$ can be measured through the following asymmetry:

$$A_{CP} = \frac{\Gamma(B_q^0 \to f) - \Gamma(\overline{B}_q^0 \to \overline{f})}{\Gamma(B_q^0 \to f) + \Gamma(\overline{B}_q^0 \to \overline{f})} = \frac{|A(B_q^0 \to f)|^2 - |A(\overline{B}_q^0 \to \overline{f})|^2}{|A(B_q^0 \to f)|^2 + |A(\overline{B}_q^0 \to \overline{f})|^2} = \frac{1-|\overline{A}_{\overline{f}}/A_f|^2}{1+|\overline{A}_{\overline{f}}/A_f|^2} \neq 0. \tag{2.54}$$

Direct *CP* violation is expected and has been observed, for example, in the decays: $B^0_{(s)} \to K^+\pi^-$, $\overline{B}^0_{(s)} \to K^-\pi^+$ [31], but also $B^\pm \to D^0 K^\pm$ [32].

The second type of *CP* violation (in mixing) occurs when the oscillation probabilities $B_q^0 \to \overline{B}_q^0$ and $\overline{B}_q^0 \to B_q^0$ differ, which implies that $|q/p| \neq 1$ and $|\lambda_f| \neq 1$. The ratio $|q/p|$ can be measured from semileptonic decays of neutral beauty mesons using the asymmetry defined by:

$$A_{CP} = \frac{\Gamma(B_q^0 \to l^-\overline{\nu}X) - \Gamma(\overline{B}_q^0 \to l^+\overline{\nu}X)}{\Gamma(B_q^0 \to l^-\overline{\nu}X) + \Gamma(\overline{B}_q^0 \to l^+\overline{\nu}X)} = \frac{1-|q/p|^4}{1+|q/p|^4}. \tag{2.55}$$

For B_q^0 mesons *CP* violation in mixing is expected to be small, at the level $O(10^{-4})$ [33], and has not been observed yet.

Finally, *CP* violation in the interference between mixing and decay is possible when $\Im(\lambda_f) \neq 0$. That type of *CP* violation may happen when the final states f and $\overline{f}$ are accessible for both B_q^0 and $\overline{B}_q^0$ mesons. As shown in Fig. 2.5, the B_q^0 meson can decay directly to a final state f, or alternatively, firstly mix to $\overline{B}_q^0$ and then decay to f. A similar process is also possible for $\overline{B}_q^0 \to \overline{f}$. This kind of *CP* violation is relevant for the decay $B_s^0 \to D_s^\mp K^\pm$ which is studied in this thesis.

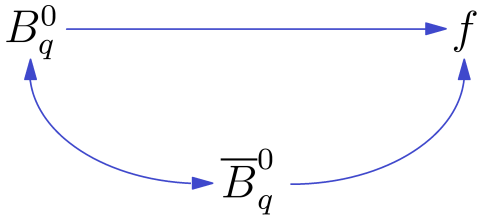

Figure 2.5: Illustration of CP violation in interference between mixing and decay.

2.4.4 Measurement of the CKM angle γ using beauty meson decays.

Taking into account the decays of the b quark which are relevant for this thesis, the Standard Model permits two generic types of decays: tree-level and loop-level, as shown in Fig. 2.6. Tree-level decays occur due to the exchange of $W^{\pm}$ bosons and thus they are not sensitive to new phenomena originating from physics beyond the SM. In the loop-level processes the $W^{\pm}$ boson is a part of a loop, from which a gluon, photon or Z boson is emitted. The latter then creates a quark-antiquark pair. In such loops massive, virtual and so far undetected particles can be exchanged. Thus this kind of diagram provides potential sensitivity to physics beyond the SM. The CKM angle γ can be determined using both types of decays as shown in Fig. 2.7. In view of the above discussion, a comparison between results obtained from both methods is a valuable test of the SM.

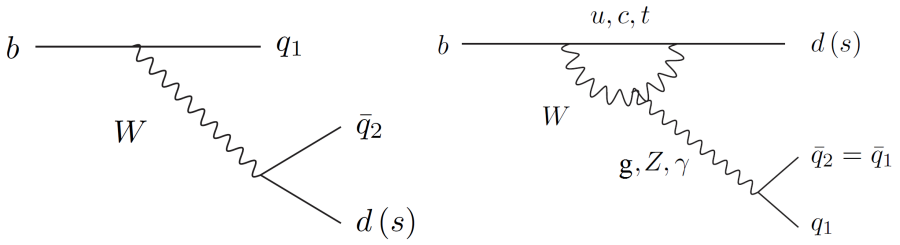

Figure 2.6: Tree-level (left) and loop-level (right) processes for the decays relevant for the CKM angle γ determination Ref. [34]

Loop-level measurements are accessible from decays to charmless final states and they further split into time-integrated and time-dependent methods. Time-integrated measurements use three-body decays [35, 36]. The determination of the CKM angle γ is also possible using the Dalitz plot technique [37], which separates $CP = +1$ and $CP = -1$ final states. On the other hand, time-dependent methods are based on two-body decays [38, 39]. Measurement of the CKM angle γ from such decays requires knowledge of hadronic factors, which are not theoretically accessible, but can be extracted from U-spin symmetry.

Tree-level measurements are also categorised into time-integrated and time-dependent ones. For the former, the CKM angle γ can be obtained by exploiting the interference between $b \to u$ and $b \to c$ transitions in the decays of beauty mesons with a charm meson in the final state, such as $B^- \to D^{(*)0}K^{(*)-}$. In $B^- \to D^{(*)0}K^{(*)-}$ decays the colour-favoured $B^- \to D^{(*)0}K^{(*)-}$ and the colour-suppressed $B^- \to \bar{D}^{(*)0}K^{(*)-}$ transitions interfere when the $D^{(*)0}$ and $\bar{D}^{(*)0}$ decay to a common final state. The relationship between γ and the physical observables depends on the $D^{(*)0}$ final state. Based on that, three separate methods of extracting the CKM angle γ can be considered:

- the Gronau-London-Wyler (GLW) method [40] [41] for *CP* eigenstates ($D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$),
- the Atwood-Dunietz-Soni (ADS) method [42] for flavour-specific eigenstates ($D^0 \to K^+\pi^-$ and $D^0 \to K^+\pi^-\pi^+\pi^-$)
- the Giri-Grossman-Soffer-Zupan (GGSZ) method [43] for self-conjugate three-body final states ($D^0 \to K^0_S\pi^+\pi^-$ and $D^0 \to K^0_S K^+K^-$).

The time dependent method with tree-level transitions will be discussed in details in the next section.

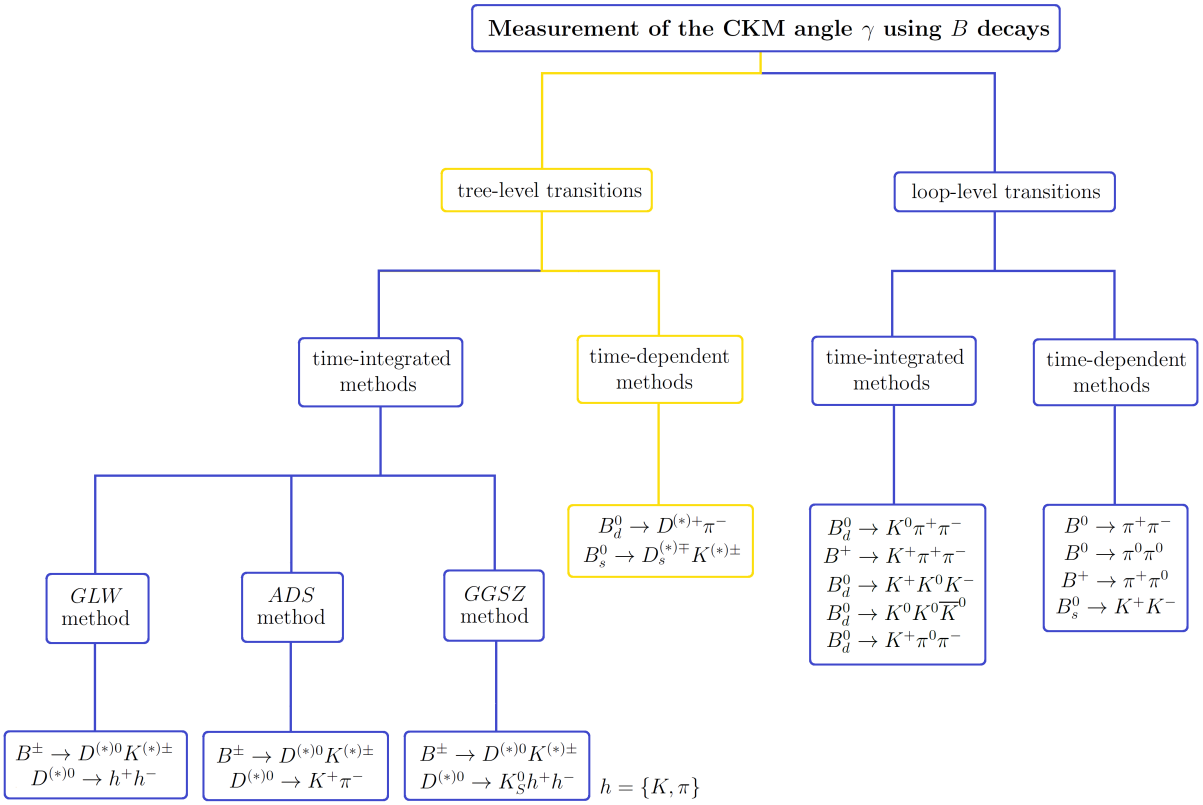

Figure 2.7: Compilation of methods of the CKM angle γ measurement. The approach which is the topic of this thesis is highlighted as yellow boxes.

2.5 Measurement of the CKM angle γ from $B^0_s \to D^\mp_s K^\pm$ decays

Time-dependent measurements of the CKM angle γ with tree-level processes exploit, just as time-integrated methods, the interference between $b \to u$ and $b \to c$ transitions [44–46].

Time-dependent measurements using $B^0 \to D^{(*)\mp}\pi^\pm$ decays were performed by both the BaBar [47, 48] and the Belle [49, 50] collaborations. For these decays, however the amplitude ratios:

$$r_{D^{(*)}\pi} = \frac{A(B^0 \to D^{(*)-}\pi^+)}{A(B^0 \to D^{(*)+}\pi^-)}, \tag{2.56}$$

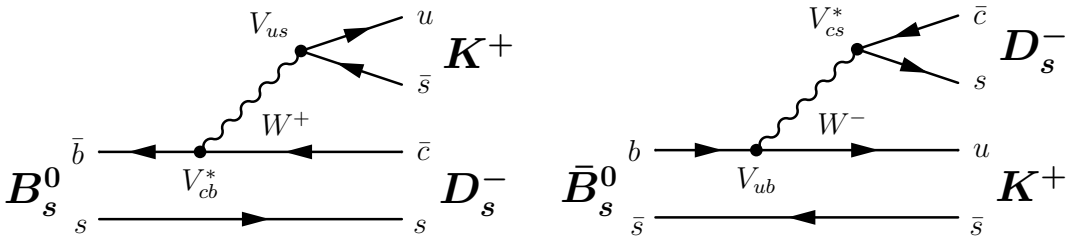

Figure 2.8: Tree diagrams for the decay $B_s^0 \to D_s^{\mp} K^{\pm}$.

between the interfering decay modes are expected to be small, $r_{D^{(*)}\pi} \approx 0.02$ [51], which limits the sensitivity.

In case of the decay $B_s^0 \to D_s^{\mp} K^{\pm}$, there exist two tree diagrams shown in Fig. 2.8. Here, both B_s^0 and $\overline{B}_s^0$ mesons can decay directly to either $D_s^- K^+$ or $D_s^+ K^-$, which leads to the possibility of *CP* violation in the interference between mixing and decay. The diagrams describing $B_s^0 \to D_s^- K^+$ and $\overline{B}_s^0 \to D_s^- K^+$ decays (Fig. 2.8) are of the order $A\lambda^3$ in the Wolfenstein parametrisation. However, the process $\overline{B}_s^0 \to D_s^- K^+$ is suppressed by a factor $|(\rho - i\eta)| \approx 0.36$ due to the value of the CKM element V_{ub}. The suppression is relatively small, so a large interference between the two amplitudes is expected.

The measurement of time-dependent *CP* violation in the decay mode $B_s^0 \to D_s^{\mp} K^{\pm}$ allows for an efficient and unambiguous measurement of the CKM angle γ. The importance of this measurement lies in the fact that $B_s^0 \to D_s^{\mp} K^{\pm}$ is a pure tree-level decay. Because of this, the theoretical uncertainty on the interpretation of the *CP* observables in terms of the CKM angle γ is tiny, $\frac{\delta\gamma}{\gamma} = 10^{-7}$ [52].

Two of the four decay rates for $B_s^0 \to D_s^{\mp} K^{\pm}$ are given by Eq. 2.47 and Eq. 2.48. The next two, for the charge conjugate final state $\overline{f}$, can be obtained by substituting f by $\overline{f}$ and λ_f by $\lambda_{\overline{f}}$ as discussed in Sec. 2.4.2. The *CP* violation parameters are defined by Eq. 2.51 and their dependence on the CKM angle γ by Eq. 2.53. The resulting *CP* asymmetries for f ($\overline{f}$) final states are:

$$A_{CP}^{f(\overline{f})}(t) = \frac{d\Gamma_{\overline{B}_s^0 \to f(\overline{f})}(t)/dt - d\Gamma_{B_s^0 \to f(\overline{f})}(t)/dt}{d\Gamma_{\overline{B}_s^0 \to f(\overline{f})}(t)/dt + d\Gamma_{B_s^0 \to f(\overline{f})}(t)/dt} = \frac{-C_{f(\overline{f})}\cos(\Delta m_s t) + S_{f(\overline{f})}\sin(\Delta m_s t)}{\cosh(\frac{\Delta\Gamma_s}{2}t) + D_{f(\overline{f})}\sinh(\frac{\Delta\Gamma_s}{2}t)}. \tag{2.57}$$

All measurements of the CKM angle γ are characterised by at least twofold, but often four or eight fold, ambiguities because of the mathematical relationship between the decay rates and the CKM angle γ. Figure 2.9 shows a simulation for the $B_s^0 \to D_s^{\mp} K^{\pm}$ *CP* observables: S_f, D_f, $S_{\bar{f}}$, $D_{\bar{f}}$ in the (γ,δ) plane. The combined likelihood together with the projection onto the CKM angle γ are presented in Fig. 2.10. It can be seen that thanks to the extra hyperbolic terms, D_f and $D_{\bar{f}}$, this ambiguity is reduced to only two solutions. The $B_s^0 \to D_s^{\mp} K^{\pm}$ decay is one of the few ways of measuring γ with this property.

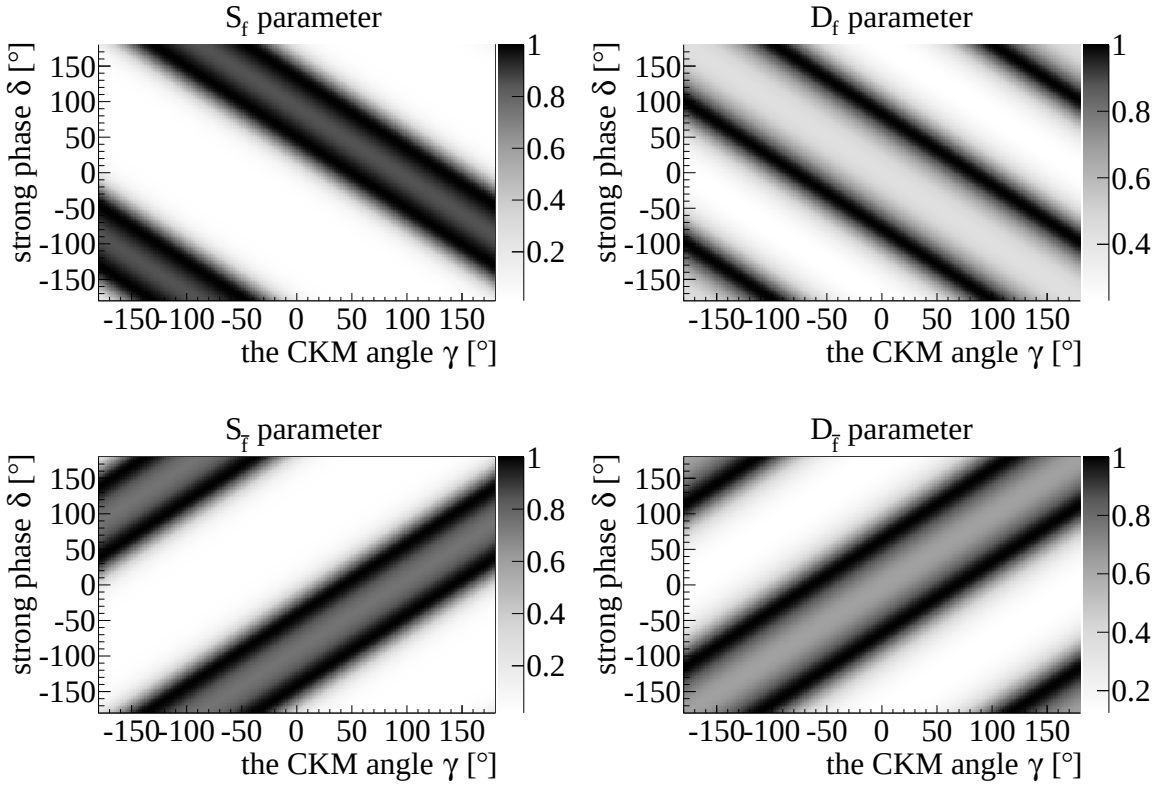

Figure 2.9: Likelihood for *CP* parameters, from top left to bottom right: S_f, D_f, $S_{\bar{f}}$, $D_{\bar{f}}$.

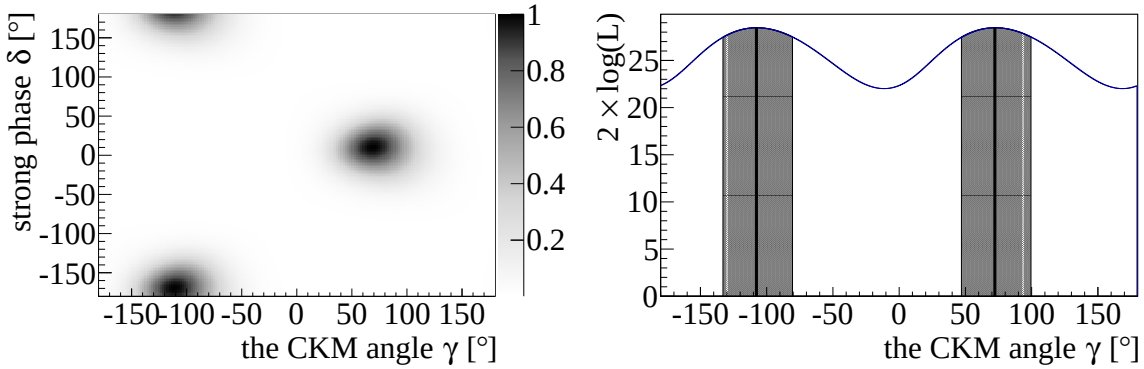

Figure 2.10: Combined likelihood for the *CP* parameters: S_f, D_f, $S_{\bar{f}}$, $D_{\bar{f}}$ (left). Projection onto the CKM angle γ (right), where the dashed region corresponds to $\pm 1\sigma$ and the black vertical line denotes the central value.

Chapter 3

The LHCb experiment

In this chapter the experimental apparatus relevant for the studies presented in this thesis is outlined. First the main features of the Large Hadron Collider are briefly presented. Next the LHCb spectrometer is discussed with special attention paid to the subdetectors which are crucial for this dissertation.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [53] is a proton-proton accelerator built by the European Laboratory for Particle Physics (CERN) at the French-Swiss border near Geneva. The LHC is located 100 m underground in a 27 km long tunnel.

A schematic view of the LHC is shown in Fig. 3.1. Protons from the source are accelerated in several stages. Firstly, they go through the linear collider (LINAC2 [54]) where they reach an energy of 50 MeV. The next step is an acceleration up to an energy of 1.4 GeV in the Booster [55]. Finally protons pass through the Proton Synchrotron (PS) [56] and the Super Proton Synchrotron (SPS) [57] where they reach an energy of 25 GeV and 450 GeV, respectively. The LHC was designed to accelerate proton beams up to an energy of 7 TeV, leading to a center-of-mass energy of $\sqrt{s} = 14$ TeV with a peak luminosity[4] of $\mathcal{L} = 1 \times 10^{34}$ cm^{-2}s^{-1}. However, in the first years after commissioning, due to issues with the superconducting dipole magnets, the LHC operated at reduced energies of $\sqrt{s} = 7$ TeV in 2010 and 2011 and $\sqrt{s} = 8$ TeV during 2012.

Four major experiments are situated in the LHC ring: ALICE (A Lead Ion Collision Experiment, [58]), ATLAS (A Toroidal LHC ApparatuS, [15]), CMS (Compact Muon Spectrometer, [16]) and LHCb (Large Hadron Collider beauty, see Sec. 3.2). ATLAS and CMS are general purpose detectors. The ALICE experiment is dedicated to the study of ultra relativistic heavy ions, which can also be collided in the LHC. The LHCb experiment will be described in detail in the next section.

3.2 The LHCb experiment

The LHCb experiment is dedicated to the study of beauty and charm hadron decays with special attention paid to *CP* violating phenomena as well as searches for physics beyond the SM through rare decays. The beauty production cross-section (the left part of

[4] $\mathcal{L} = \frac{dR}{dt}\frac{1}{\sigma_R}$, where $\frac{dR}{dt}$ is the number of candidates for a given process R detected in time t and σ_R is the interaction cross section for the process in question.

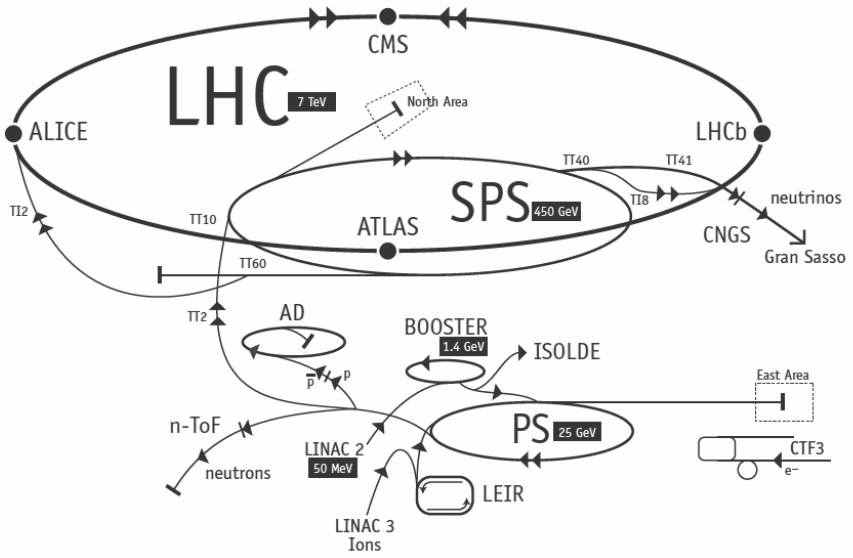

Figure 3.1: A schematic view of the LHC accelerator [59].

Fig. 3.2) is dominated by gluon fusion, and thus it predominantly occurs in the forward region. Therefore, in contrast to other detectors, the LHCb detector is a single-arm forward spectrometer with a pseudorapidity[5] η spanning the range $1.8 < \eta < 4.9$.

The LHCb detector was designed to operate at lower instantaneous luminosities than those used by ATLAS and CMS. At the LHCb crossing point one beam is moved relative to the other, so that the beams do not collide head-on and the interaction area is bigger. This procedure is known as luminosity levelling. The integrated luminosity collected by the LHCb experiment in 2010-2012 is shown on the right part of Fig. 3.2. Due to luminosity levelling the LHCb experiment collected around $1\,\text{fb}^{-1}$ in 2011, compared to $5\,\text{fb}^{-1}$ for ATLAS and CMS.

Thanks to the excellent performance of the LHCb detector, the LHCb experiment is uniquely able to study *CP* violation in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. Even though the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay was seen earlier by the Belle [24] and CDF [60] experiments, neither of them (which are no longer taking data) collected as many candidates as LHCb has collected so far. In addition, Belle does not have sufficient decay time resolution to perform time dependent *CP* violation measurements in B_s^0 decays, while CDF has much worse signal purity than LHCb because of their inferior identification of charged hadrons.

The reconstruction of beauty and charm mesons requires effective tracking and particle identification systems specialised in collecting specific information about particles produced in proton-proton collisions. The LHCb spectrometer, shown in Fig. 3.3, is composed of several sub-detectors. The LHCb co-ordinate system is a right handed Cartesian system with the origin at the interaction point. The x-axis is oriented horizontally towards the outside of the LHC ring, the y-axis is pointing upwards with respect to the beamline and the z-axis is aligned with the beam direction.

The following subsections consist of brief descriptions of these sub-systems, focusing on those which are the most relevant for the analysis presented in this thesis. More detailed information can be found in Ref. [28].

[5] $\eta = -\ln[\tan\left(\frac{\theta}{2}\right)]$ where θ is the angle between the beam axis (for the LHCb experiment z axis) and the particle's momentum.

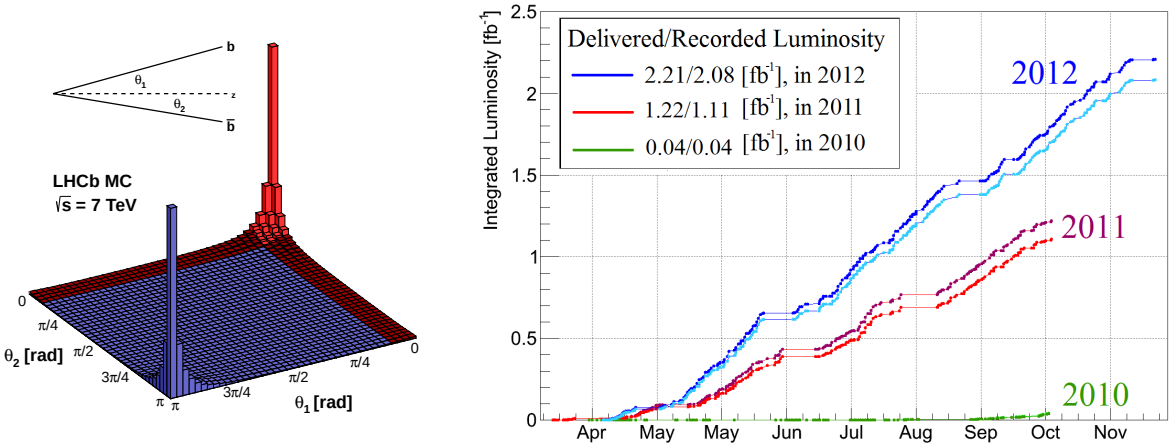

Figure 3.2: Left: the relative cross section for $b\bar{b}$ production as a function of polar angles at $\sqrt{s} = 7$ TeV Ref. [61]. Right: the integrated luminosity collected by the LHCb detector in 2010-2012 Ref. [62].

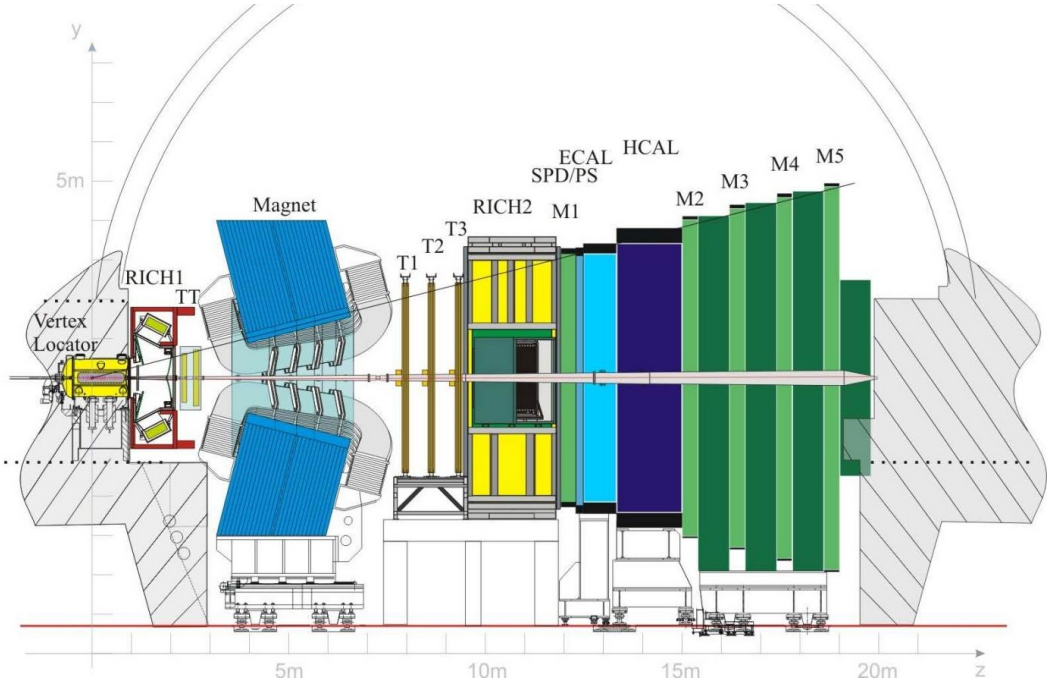

Figure 3.3: A schematic view of the LHCb detector Ref. [28]. The labels RICH1 and RICH2 denote the Ring Imaging Cherenkov detectors, TT and T1-T3 are the tracking stations, M1-M5 show the muon detectors, SPD(PS) is the scintillator pad(preshower) detector. Finally ECAL and HCAL mark the electromagnetic and hadronic calorimeters, respectively.

3.2.1 The tracking system

The main goal of the tracking system is to find and measure the signatures of charged particles passing through the LHCb detector. These signatures are labelled by tracks and they constitute a necessary input to the reconstruction algorithms.

The VErtex LOcator

The Vertex Locator (VELO) subdetector is responsible for a precise determination of charged track positions close to the interaction point. As a result it is crucial in the reconstruction of vertices. These can be divided in two categories: primary, which correspond to the position of a pp collision, and secondary, formed by the decay of long-lived particles. The relative position between primary and secondary vertices, together with momentum information from other subdetectors, allows the decay-time of b hadrons to be reconstructed. This is a necessary observable in many analyses, in particular time-dependent CP violation measurements.

The VErtex LOcator (VELO) is composed of 21 semicircular stations and two additional VETO stations, which can be used to reject candidates with too many pp collisions per bunch crossing (so-called pile-up). Each station has left and right parts, which can be moved apart during the beam injection period and moved back together once the beams are stable. In nominal working conditions the active parts of the VELO are about 8 mm away from the beam. Figure 3.4 shows the layout of the VELO sensor stations in the x-z plane.

Each half of the VELO is made of a silicon wafer covered by aluminium strips. Its thickness (diameter) amounts to 300 μm (84mm). To provide two-dimensional information about particle positions the VELO sensors are split into two categories: R-type and ϕ-type. R-type sensors have semicircular strips and measure the radial coordinate, whereas ϕ-type sensors provide information about the azimuthal direction (ϕ angle) and their strips are placed along the radius. Figure 3.5 shows an overview of a single sensor, which is composed of one R-type and one ϕ-type sensor. Each VELO half-module has 2000 strips, which gives 180 000 read-out channels for the entire detector.

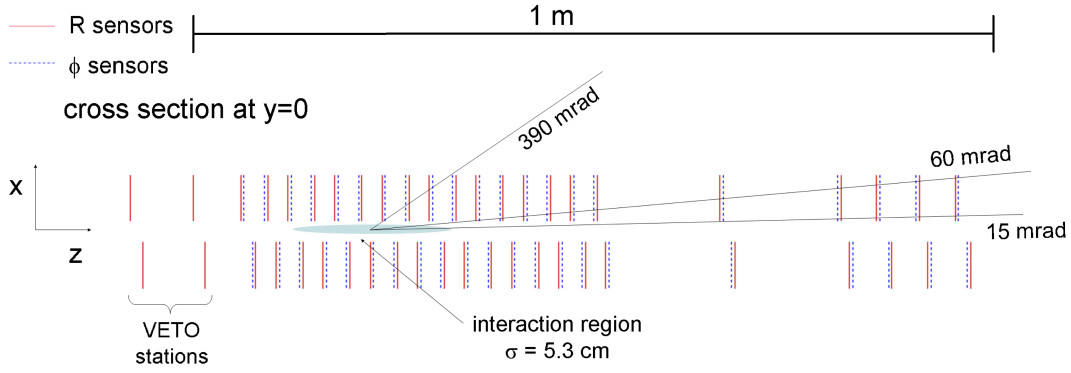

Figure 3.4: Layout of the Vertex Locator Ref. [28].

The Tracker Turicensis

The Tracker Turicensis (TT) is located between RICH1 and the LHCb magnet as shown in Fig. 3.3. Its main purpose is to improve the track momentum resolution as well as the reconstruction of long-lived particles that decay outside the VELO. The TT is composed of two stations (TTa and TTb), each of them made of two layers of silicon micro-strip sensors. The layers are divided into sections depending on the beam occupancy. Determination of the transverse component of the momentum is possible due to the presence of two type of layers. The TTa (TTb) stations consist of the so-called X-U (X-V) layers, respectively. The X layers have vertical readout strips, while U and V stereo layers are rotated along

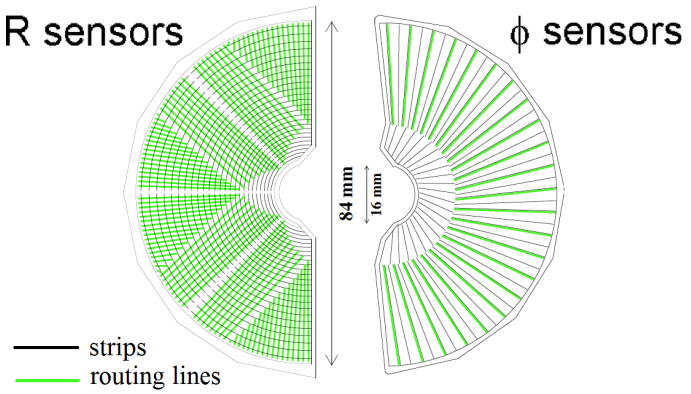

Figure 3.5: Layout of the R-type and ϕ-type sensors of the Vertex Locator Ref. [28].

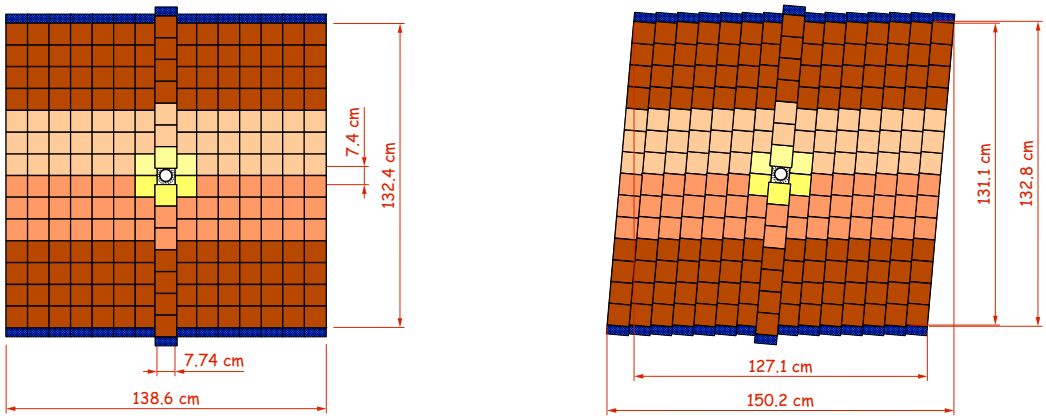

Figure 3.6: Schematic views of a single layer of the Tracker Turicensis. The left (right) plots show X (U) view, respectively. Yellow/brown colours denote different readout sections, while a navy blue corresponds to readout electronics. Ref. [28].

the y-axis by $+5°$ or $-5°$, respectively. The design of the two different layers is shown in Fig. 3.6. In total the TT stations contain around 140 000 readout strips covering an active area of 8.4 m^2. The hit resolution of the TT is about 50 μm.

The Inner Tracker

The Inner Tracker (IT) covers the central part of the T1-T3 tracking stations, located behind the LHCb magnet as shown in Fig. 3.3. The detector covers the 1.3% of the LHCb acceptance closest to the beamline. In this area the particle flux is much greater compared to the remaining regions of the T1-T3 acceptance. For this reason, the IT is composed of radiation-hard silicon detectors. It is built from four sections placed above, below, and on each side of the beamline as presented on the right part of Fig. 3.7. Each IT station comprises four layers arranged in the same manner (X-U-V-X) as the layers in the TT stations. The IT has a hit resolution of about 50 μm and contains around 129 000 readout channels.

The outer tracker

The outermost layers (w.r.t the beam pipe) of the T1-T3 stations are filled with the Outer Tracker (OT) as shown in Fig. 3.7. Each OT station covers a large area of 340 m^2.

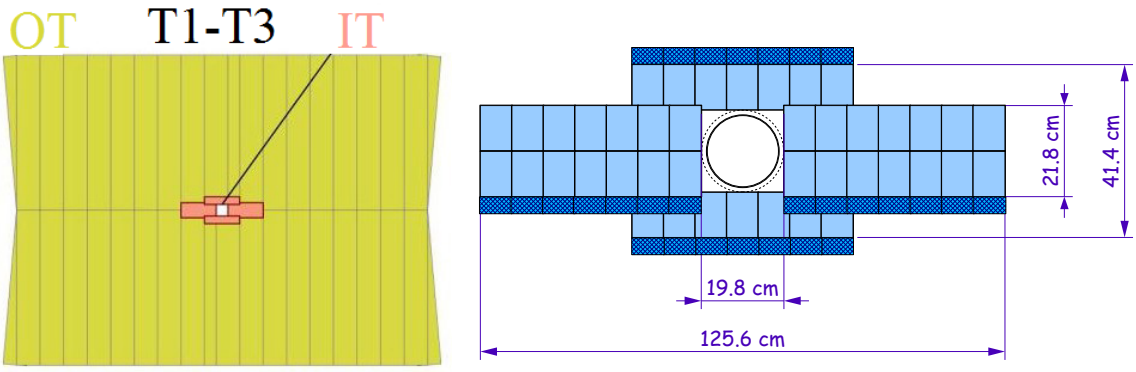

Figure 3.7: Overall layout of the tracking stations T1-T3 (left) and of the single (X) layer of the IT (right) Ref. [28].

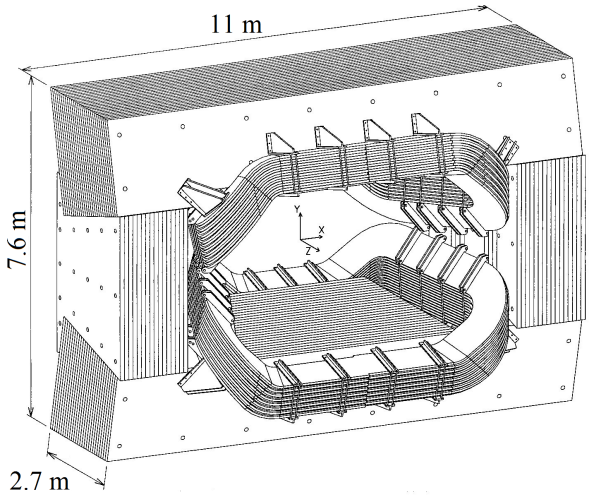

Figure 3.8: A view of the LHCb magnet Ref. [28].

The OT is a gaseous straw-tube detector filled with an $Ar/CO_2/O_2$ (70%/28.5%/1.5%) gas mixture. Each of the three stations has four layers in a (X-U-V-X) configuration. In addition, each layer is divided into 20 modules with two layers of straws. The straws have an inner diameter of 4.9 mm. For optimal spatial resolution, the straw layers are shifted by half a straw diameter inside one module. The OT is composed of around 54 000 straw tubes and provides a hit resolution of about 200 μm.

The LHCb magnet

The dipole magnet curves the trajectories of charged particles enabling the measurement of their momenta. It is situated between the TT and T1-T3 stations as presented in Fig. 3.3. The magnet consists of a steel cover surrounding two identical aluminium non-superconductive coils placed with mirror symmetry to the beam axis as shown in Fig. 3.8. The integrated magnetic field for tracks with length of 10 m is 4 Tm. To reduce detection asymmetries the polarity of the magnet is switched repeatedly during data taking.

3.2.2 Particle Identification

In a single proton-proton collision a lot of different charged particles are produced, primarily pions, kaons, electrons, protons or muons. Their correct identification is crucial for the reconstruction of beauty meson decays. To perform this task, the difference in logarithmic likelihood is defined as:

$$\mathrm{DLL(A-B)} = \ln \mathcal{L}_{\mathrm{A}} - \ln \mathcal{L}_{\mathrm{B}} \tag{3.1}$$

where $\mathcal{L}_A$ and $\mathcal{L}_B$ denote the likelihood that a particle is identified under the A and B hypotheses, respectively. For example, in case of $\mathrm{DLL(A-B)} < 0$ the particle is more likely B, whereas for $\mathrm{DLL(A-B)} > 0$ it is more likely A.

The final state particles of the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay are pions and kaons. Moreover, the most important peaking background is due to the $B_s^0 \to D_s^- \pi^+$ decay, with a 15 times larger branching fraction than $B_s^0 \to D_s^{\mp} K^{\pm}$ [63]. For this background a pion is misidentified as a kaon. Hence the distinction between kaons and pions is crucial for the analysis, and therefore the important observable is $\mathrm{PIDK} = \mathrm{DLL(K}-\pi)$. Other PID variables used in the presented analysis are $\mathrm{PID}\mu = \mathrm{DLL}(\mu - \pi)$ and $\mathrm{PID}p = \mathrm{DLL(p}-\pi)$, which provides discrimination against muons and protons, respectively.

Several subdetectors provide the necessary input to the identification procedure, based on which the combined likelihood for a specific particle hypothesis is obtained. These subdetectors are: the Ring Imaging Cherenkov detectors, the calorimeters and the muon chambers. Their brief description is given in the following subsections.

The Ring Imaging Cherenkov detectors

The Ring Imaging CHerenkov (RICH) detectors exploit the so-called Cherenkov effect [64] to measure the velocity of charged particles. This effect occurs when a charged particle travels faster than the speed of light in a medium, which is given by:

$$c' = \frac{c}{n}, \tag{3.2}$$

where n is the refractive index of the medium and c denotes the speed of light in vacuum. The LHC produces highly relativistic particles, whose velocity can be greater than c' for some materials. In this case the particle emits photons in a cone w.r.t. the direction of propagation, characterised by the so-called Cherenkov angle θ_C, which satisfies the formula:

$$\cos(\theta_C) = \frac{c'}{v} = \frac{1}{\beta n}, \tag{3.3}$$

where β is the velocity of the particle. Thus, the determination of the Cherenkov angle provides information about particle's velocity which, together with the momentum measurement from the tracking system, allows the particle's mass to be determined and therefore gives particle identification.

Cherenkov angles for different particles as a function of track momentum are shown on the left part of Fig. 3.9. As can be seen, in order to assure an efficient particle identification over a broad range of momenta (1-100 GeV/c in case of the LHCb experiment) it is necessary to use several radiators. The aerogel radiator is well suited for the identification of low-momentum particles while the CF_4 and C_4F_{10} media provide good discriminating power at high-momenta (above 20 GeV/c).

The RICH1 detector is located in front of the LHCb magnet as presented in Fig. 3.3. To provide an efficient identification in the momentum range from 2 GeV/c to 60 GeV/c,

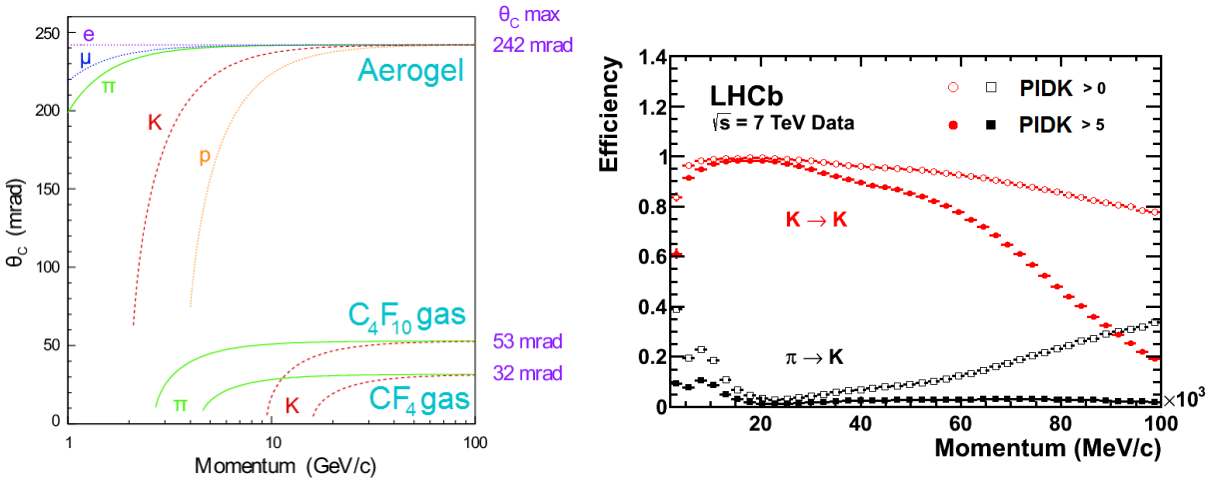

Figure 3.9: Left: Cherenkov angle θ_C versus track momentum for different particles and radiators Ref. [28]. Right: the kaon identification efficiency (red points) and a pion misidentification rate (black points) versus momentum for data collected in 2011 by the LHCb experiment Ref. [65]. The closed (open) points correspond to PIDK > 5 (> 0).

it uses two radiators: aerogel with refractive index $n = 1.03$ and C_4F_{10} with $n = 1.0014$. RICH1 covers an acceptance of 25 mrad to 250 (300) mrad in the vertical (horizontal) direction. RICH2 is placed behind the LHCb magnet and the T1-T3 tracking stations (see Fig. 3.3). It uses a CF_4 gas radiator with a refractive index $n = 1.0005$, which allows for an efficient identification of particles with momenta between 15 GeV/c and 100 GeV/c. The acceptance of RICH2 is limited to 15-120 mrad (15-100 mrad) in x (y) directions. Since track momentum is strongly correlated with polar angle in the LHCb detector, the reduced RICH2 acceptance covers the region where the majority of high momentum particles are produced. The design of both detectors is shown in Fig. 3.10. Particles which go through the detector emit Cherenkov light which is collected by flat and spherical mirrors and then reflected onto the Hybrid Photo Detectors (HPDs) and photomultiplier tubes. The HPDs are sensitive in the 200-600 nm wavelength range.

In the case of the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay, the distinction between kaons and pions is crucial. The kaon identification efficiency and pion misidentification rate are shown as a function of momentum, for two different values of the PIDK requirement: $PIDK > 0$ (open shapes) and $PIDK > 5$ (solid shapes) using data collected in 2011 on the right part of Fig. 3.9.

The calorimeter system

The calorimeter measures the energy of charged and neutral particles using the following principle: a particle travels through it and interacts with an absorber material. This interaction creates a secondary electromagnetic and/or hadronic shower, which is measured in an active material of the calorimeter (e.g. scintillators or silicon detectors). Thus, the traversing particles deposit almost all their energy in the calorimeter.

Electron and photon identification is possible thanks to the combined information from the Scintillator Pad (SPD), the Preshower Detector (PS), and the Electromagnetic Calorimeter (ECAL). They are located behind the first muon station as shown in Fig. 3.3. The SPD and PS use scintillating pads, whose granularity depends on the distance from the beam direction as presented in Fig. 3.11. Neutral photons do not introduce a signal in the pads before triggering an electromagnetic shower, allowing the SPD to distinguish

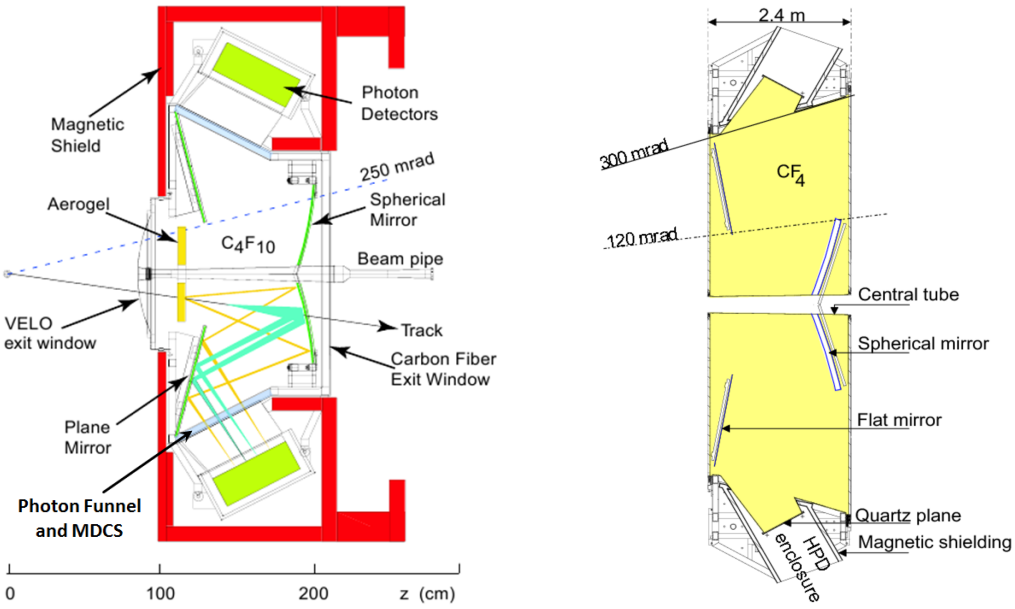

Figure 3.10: Layout of the LHCb Ring Imaging Cherenkov detectors: RICH1 (left) and RICH2 (right) Ref. [28].

between electrons and photons. The SPD and PS are separated by a 15 mm thick lead plate, which initiates a shower of photons and electrons. The PS is used to distinguish between electrons and charged particles which do not introduce a shower, for example pions. The particles then pass through the ECAL, where they deposit their energy and where photons and electrons are absorbed. The ECAL is made of 2 mm lead plates followed by 4 mm scintillator pads. Its granularity depends on the distance from the beam direction in the same way as for the SPD and PS. The energy resolution of the ECAL is:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E/\text{GeV}}} \oplus 1\%, \tag{3.4}$$

where $\oplus$ denotes that the two terms are added in quadrature. In total, the ECAL has about 6000 readout channels.

The last part of the LHCb calorimeter system is the Hadron Calorimeter (HCAL), where the energy deposit left by charged and neutral hadrons which penetrated through the ECAL can be measured. The HCAL is made of iron blocks and scintillating tiles, and its layers have a parallel orientation to the beam direction. The granularity of this detector is presented on the right part of Fig. 3.11. The energy resolution of the HCAL is:

$$\frac{\sigma_E}{E} = \frac{69\%}{\sqrt{E/\text{GeV}}} \oplus 9\%. \tag{3.5}$$

It has about 1500 readout channels.

The muon chambers

Muon detection is performed in the five stations of muon chambers placed in the outermost layers of the LHCb spectrometer. The first station (M1) is located before the calorimeter system, while the remaining four stations (M2-M5) are placed behind the system, as shown in Fig. 3.12. The muon chambers are gaseous detectors. Due to greater occupancy

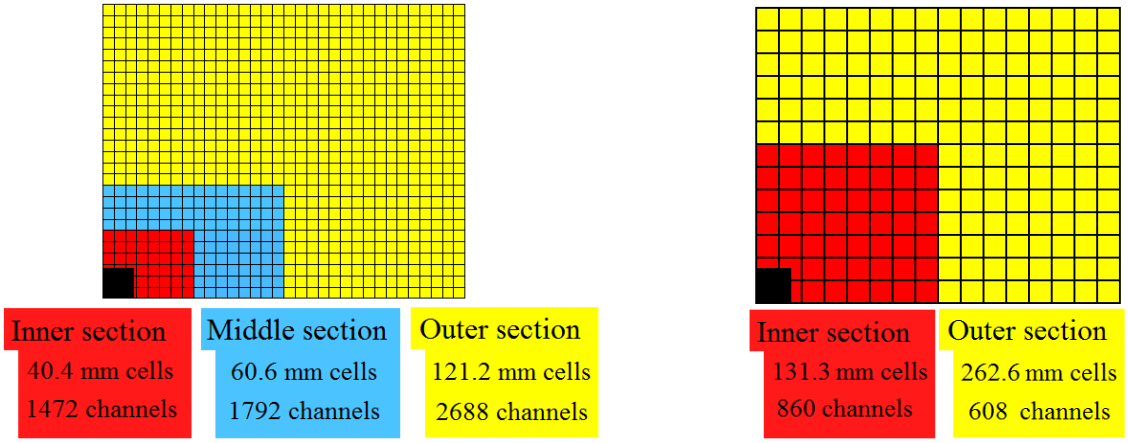

Figure 3.11: Granularity for the different detector regions of the SPD, PS, and ECAL (left) and of the HCAL (right) Ref. [28].

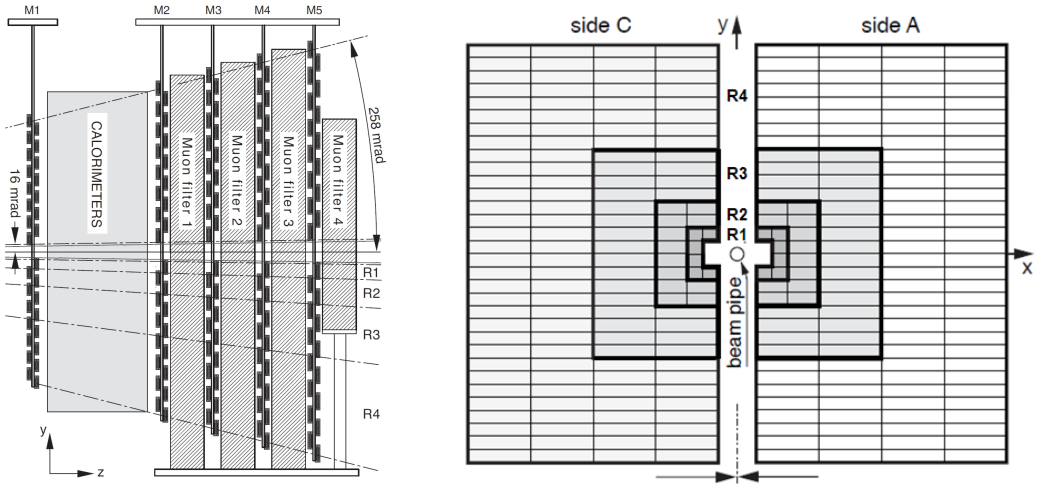

Figure 3.12: The LHCb muon chambers Ref. [28].

M1 uses gas electron multipliers (GEM), while the remaining stations (M2-M5) are built from multiwire proportional chambers (MWPC). To absorb hadrons from the calorimeter system, 80 cm thick iron walls are placed between the stations. The presence of these iron walls means that muons with momenta below 6 GeV/c cannot reach the M5 station. The granularity pattern of muon stations follows the one of the overall calorimeter system, as shown in Fig. 3.12.

3.3 The LHCb trigger

In the nominal conditions of the LHC operation, the proton beams would intersect (so-called bunch crossing) with a rate of 40 Hz. However, in 2010-2012 the LHC operated at a reduced energy with a lower bunch crossing rate of 15 MHz. This value is still several orders of magnitude above the current abilities of data recording and storage. Moreover, candidates which are potentially interesting for further studies comprise only a very small fraction of the overall rate of proton-proton collisions. Therefore, a trigger system is required to reduce this rate to an acceptable value which for LHCb was set at about 5 kHz. Figure 3.13 shows a schematic view of the LHCb trigger. It is composed of three

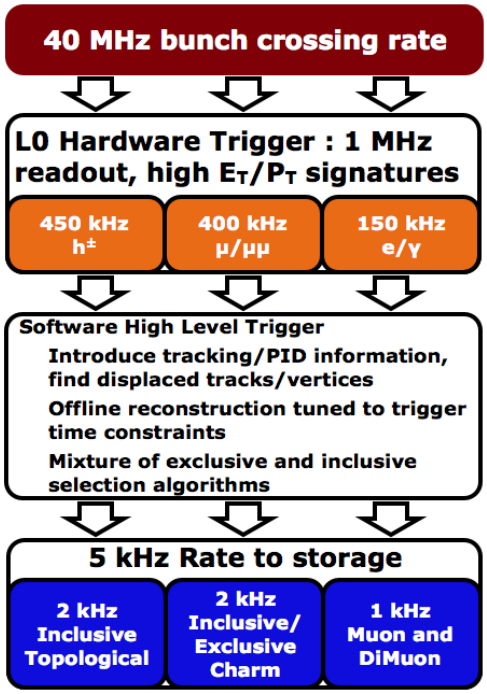

Figure 3.13: The layout of the LHCb trigger Ref. [66].

stages: one hardware known as Level0 (L0) and two software called High-Level-Trigger (HLT1 and HLT2). The conditions which describe each trigger decision are collected in so-called trigger lines.

3.3.1 Hardware trigger

The main purpose of the L0 trigger stage is to quickly find general signatures of interesting decays and reduce the rate of accepted candidates to about 1 MHz, at which the full detector can be read out. Different L0 decisions are assigned to the candidate if the following conditions are satisfied [67]:

- `L0-Muon`: occurrence of one stub in the muon chambers with transverse momentum above 1.48 GeV/c,
- `L0-DiMuon`: presence of a pair of stubs in the muon chambers with a product of the transverse momentum above 1.296 $(\text{GeV/c})^2$,
- `L0-Hadron`: occurrence of one cluster in the HCAL with transverse energy above 3.5 GeV/c,
- `L0-Photon`: presence of one cluster in the ECAL with transverse energy above 2.5 GeV/c, hits in the PS, and no hits in the SPD,
- `L0-Electron`: occurrence of one cluster in the ECAL with transverse energy above 2.5 GeV/c, hits in the PS, and at least one hit in the SPD.

Moreover, the estimation of the number of tracks in the event is done based on the number of hits in the SPD. The threshold is set to be fewer than 900 hits for `L0-DiMuon` and fewer than 600 hits for other trigger decisions. It allows the L0 trigger to reject candidates with too many particles.

3.3.2 Software trigger

All candidates accepted by the L0 trigger are processed by the software trigger, which is composed of two stages: HLT1 and HLT2. It uses a large computing cluster with about 29,000 logical CPUs known as the Event Filter Farm (EFF). For 2011 data taking, the first stage reduced the rate of candidates to about 50 kHz, while the corresponding value for the second one amounted to 3 kHz (5 kHz in 2012).

The HLT1 trigger performs a partial reconstruction of the candidates. In this stage all requirements are inclusive, which means that the selection is applied only to subset of the final state particles. Information from VELO and T1-T3 stations is used to search for tracks with high transverse momentum and large displacement from the primary vertex. Specific requirements are imposed depending on the decision from the L0 trigger. Candidates which successfully pass these selections receive a positive `1Track` decision [68].

The HLT2 trigger contains both inclusive and exclusive algorithms which are more time-consuming but provide a more precise reconstruction. Exclusive algorithms are used to select a specific decay at the trigger stage. A large part of HLT2 is dedicated to inclusive topological trigger lines, which, based on multivariate algorithms (see Sec. 5.3), reconstruct all b-hadron decays with charged particles in the final state. These topological trigger lines select a displaced two-, three- or four-body vertex and are labeled as 2-, 3- or 4-body `TopoBBDT` ones [69, 70], respectively. The other trigger lines are inclusive or exclusive selections of c-hadron decays (exploiting information about reconstructed secondary vertices) and single or di-muon trigger lines. One of the inclusive trigger lines is the `IncPhi` decision which selects $\phi^0(1020) \to K^+K^-$ candidates. The word "inclusive" means here that the $\phi^0(1020)$ meson can come from any decay for example $D_s^0 \to \phi^0(1020)\pi^0$, $B_s^0 \to J\psi\phi^0(1020)$, $B_s^0 \to \phi^0(1020)\gamma$ etc.

The candidates accepted by the trigger are split into three categories:

- **Triggered On Signal** (TOS): a candidate for which a positive trigger decision is generated due to sufficient presence of the signal i.e. the trigger tracks have overlap with the signal.
- **Triggered Independently of Signal** (TIS): a candidate for which a positive trigger decision is generated independently of the signal presence.
- **Triggered On Both** (TOB): a candidate which is neither TIS nor TOS.

Details on the `1Track`, topological triggers and TISTOS categories are given in dedicated public note [68].

3.4 Reconstruction algorithms

Reconstruction algorithms provide information about tracks and primary vertices.

The tracking reconstruction consists of three main parts. Firstly, the signatures produced in the detector by charged particles are found and the tracks are formed from hits. Then, all found tracks are fitted using a Kalman fit [71] which obtains the best possible estimate of the true trajectory. This procedure includes corrections due to energy loss and multiple scattering. Finally, any duplicated tracks are removed and a container with the best unique tracks is created.

For each track a state vector $\vec{s} = (x, y, t_x, t_y, q/p)^T$ is set, where x and y are coordinates, t_x (t_y) are slopes in x-z and y-z projections and q/p denotes charge divided by track

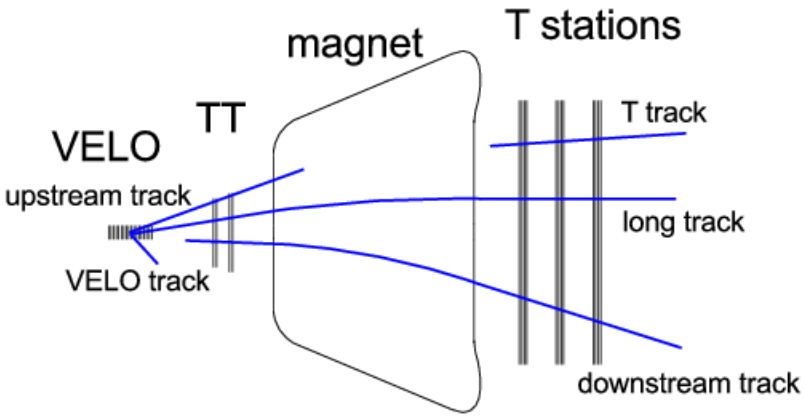

Figure 3.14: Different type of tracks in the LHCb experiment Ref. [72].

momentum. The track trajectory and its uncertainty are represented by a state vector and its covariance martix.

In the LHCb experiment, as presented in Fig. 3.14, tracks are grouped in five categories:

- **VELO track**,
  has a signature only in the VELO detector and is used in finding primary vertices. Due to the lack of magnetic field, VELO tracks are always reconstructed as straight lines.

- **Upstream track**,
  is reconstructed from VELO tracks and TT hits. Usually it is a low momentum track which is bent out of the LHCb acceptance by the magnet.

- **T track**,
  with hits in the main tracking stations T1-T3.

- **Downstream track**,
  has a signature in both the TT and T1-T3 tracking stations, but doesn't in the VELO detector. This type of track is important in the reconstruction of long-lived particles such as K_s^0 or Λ, which often decay after passing the VELO detector.

- **Long track**,
  has signatures in all subdetectors, from the VELO detector to the T1-T3 tracking stations. Since these tracks pass through the magnetic field, they have the most accurate measured momentum and thus are the most useful for analysis.

The properties of selected tracks vary depending on their transverse momentum. In 2011 data taking the low momentum tracks had momentum resolution $\Delta p/p = 0.4\%$, while for high momentum the value $\Delta p/p = 0.6\%$ was found [73].

In addition to track reconstruction, the primary vertex finding plays a crucial role in *CP* measurements. The position of the *pp* collision is determined using a set of input tracks. The first step of the primary vertex algorithm is the so-called seeding. A search is performed in a loop over tracks and for each of them a number of *close* tracks is determined. A track is defined as *close* when the distance of the closest approach to the reference track is less then 1 mm. A selected group of tracks is called a *seed* and it creates a candidate for a primary vertex. For each pair of tracks in the *seed*, the distance of the closest approach

is calculated and their weighted average determines the final coordinates of candidate. In the second step the position of the primary vertex is found by minimising the the primary vertex χ^2 defined as:

$$\text{Vertex}\chi^2 = \sum_{i=1}^{\text{nTracks}} \text{IP}\chi_i^2 \cdot w_i, \tag{3.6}$$

where nTracks denotes the number of tracks associated to the vertex, $\text{IP}\chi_i^2$ is a single track IP χ^2 and w_i is its weight, which parametrises the significance of the given track in the vertex fit. The weights w_i are obtained for each track i based on the adaptive weighted least square method with Tukey biweights [74] and are given by:

$$w = \begin{cases} (1 - \chi_{IP}^2/C_T^2)^2, & \text{if } \chi_{IP} < C_T, \\ 0, & \text{if } \chi_{IP} \geq C_T, \end{cases} \tag{3.7}$$

where C_T is Tukey's constant [74].

The author of this thesis has been working on developments and the optimization of primary vertex reconstruction for data taking which starts in 2015. More detailed information about these studies can be found in App. A.

3.5 The LHCb software

The LHCb experiment provides several software packages which allow to either generate simulation or analyse collected data. The structure of LHCb software is shown in Fig. 3.15. The packages are based on ROOT [75] and Gaudi [76] frameworks. The main of them are:

- **GAUSS** [77],
  allows to generate the simulated candidates. Other software packages are used inside Gauss. In particular, the simulation of the proton-proton interaction is done by PYTHIA [78], while the decays of hadronic particles are realised by EvtGen [79]. Moreover, final state radiation is simulated by PHOTOS [80]. Finally the recent detector description is based on informations from GEANT4 [81, 82].

- **BOOLE** [83],
  the output of GAUSS is digitised by BOOLE and the detector response in the individual subdetectors is added.

- **MOORE** [84],
  provides information about trigger decisions for both simulation and data.

- **BRUNEL** [85],
  is used by both simulation and data. The main purpose of the package is to perform the full event reconstruction from raw candidates.

- **DaVinci** [86],
  allows particles to be combined into a decay chain and the signal decays to be reconstructed for analyses. DaVinci uses the decay tree fitter (DTF) [87], which refits simultaneously all vertices specified in the chain. Also at this stage a very loose preselection is applied. DaVinci framework can be used on both simulation and real data.

- **URANIA** [88],
  is a group of packages dedicated to different physics analyses. The analysis presented in this thesis was done using a part of Urania, namely the `B2DXFitters` package [89]. The author of this thesis is one of the main authors of that package. More information about this package can be found in App. B.

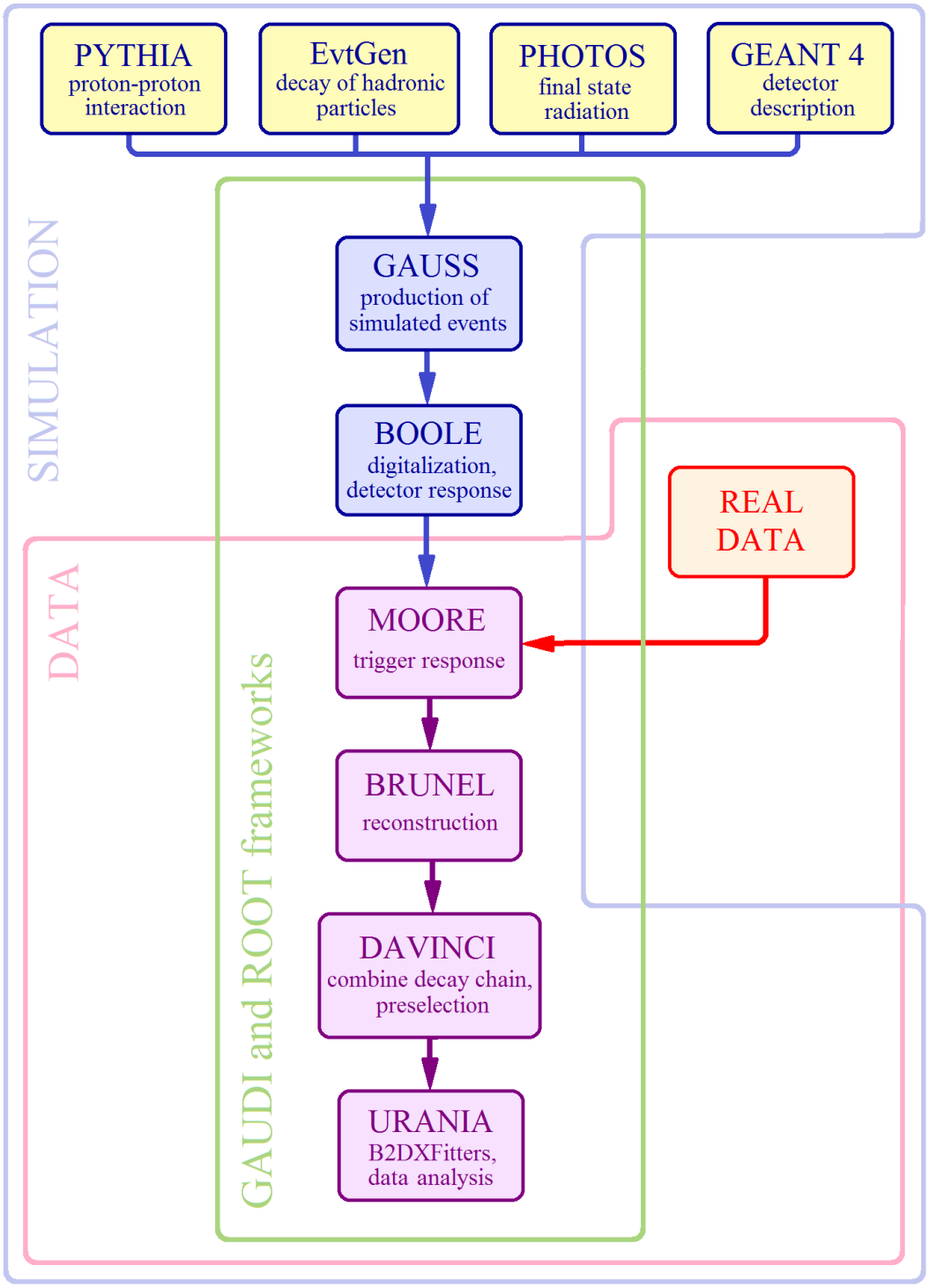

Figure 3.15: A flowchart of the LHCb software.

Chapter 4

Analysis strategy

This chapter outlines the major steps of the analysis which will be discussed in detail in the next sections. Also, it provides the information about the data sample as well as the signal and control decay channels.

4.1 Data sample

This analysis uses an integrated luminosity $\int \mathcal{L}dt = 1.03\,\text{fb}^{-1}$ of data recorded in proton-proton collisions with the LHCb detector at a center-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ in 2011 [62]. All detector components were fully operational and in stable conditions during data taking.

The dataset comprises two distinct sub-samples recorded with opposite directions of the magnetic field: 0.44 fb^{-1} with the magnet up and 0.59 fb^{-1} with the magnet down polarities.

4.2 Signal and normalisation decay channels

In addition to the signal channel $B_s^0 \to D_s^{\mp} K^{\pm}$, the decays $B_s^0 \to D_s^- \pi^+$ and $B^0 \to D^- \pi^+$ are used as control modes. The latter are characterised by a similar content of final state particles and overall topology, to compare with the signal mode, and their branching fractions are measured with reasonable precision.

The first control channel, $B_s^0 \to D_s^- \pi^+$, is a flavour specific mode, thus only the decays $B_s^0 \to D_s^- \pi^+$ and $\overline{B}{}_s^0 \to D_s^+ \pi^-$ are allowed. In addition, the $B_s^0 \to D_s^- \pi^+$ mode is a purely tree-level process as shown in Fig. 4.1. The decay rates given by Eq. 2.47 and Eq. 2.48 are simplified to:

$$\frac{d\Gamma_{B_s^0 \to D_s^- \pi^+ (\overline{B}{}_s^0 \to D_s^- \pi^+)}(t)}{dt} = \frac{1}{2}|A_f|^2 e^{-\Gamma t}\left[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) \pm \cos\left(\Delta m_s t\right)\right], \qquad (4.1)$$

from which a measurement of Δm_s is possible. Moreover, the $B_s^0 \to D_s^- \pi^+$ decay is topologically and geometrically identical to the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay. This allows the use of information obtained from the $B_s^0 \to D_s^- \pi^+$ mode in many stages of the $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis. For this reason, a full analysis of $B_s^0 \to D_s^- \pi^+$ decays is performed along with the signal mode.

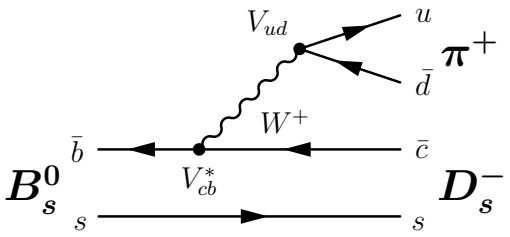

Figure 4.1: Feynman tree diagram for the $B_s^0 \to D_s^- \pi^+$ decay.

The second control mode is $B^0 \to D^- \pi^+$, $D^- \to K^+ \pi^- \pi^-$ decay. Its yield is used to compute the number of misidentified D-like backgrounds that lie under the $B_s^0 \to D_s^\mp h^\pm$ signal peaks, as well as to reduce the difference between simulation and data.

The signal and normalisation B_s^0 decays are distinguished based on the mass hypothesis of the bachelor particle accompanying the D_s^- mesons. The latter is reconstructed in the following three final states $D_s^- \to K^- K^+ \pi^-$, $D_s^- \to K^- \pi^+ \pi^-$, and $D_s^- \to \pi^- \pi^+ \pi^-$. In case of the decay $D_s^- \to K^- K^+ \pi^-$, two resonant states are selected: $D_s^- \to \phi \pi^-$ and $D_s^- \to K^{*0} K^-$. The remaining $D_s^- \to K^- K^+ \pi^-$ candidates are called non-resonant and labelled as $D_s^- \to (KK\pi)_{\text{nonres}}$.

In all cases the $D_{(s)}^-$ meson is reconstructed from three particles with the appropriate mass hypotheses; a "bachelor" $h^+ = \{K^+, \pi^+\}$ is then added to form the $B_{(s)}^0$ candidates. Table 4.1 collects the most relevant properties of the studied particles. The production rates of b hadrons are determined by the fragmentation functions f_u, f_d, f_s, f_Λ, which describe the probability that a b quark will fragment into a B_q^0 meson (where $q = \{u, d, s\}$) or Λ_b^0 baryon, respectively. An overview of the relevant branching fractions as well as the fragmentation fraction f_s/f_d are collected in Tab. 4.2.

Table 4.1: Major properties of the particles relevant for the analysis, taken from [63].

Particle	Quark composition	Mass [MeV/c^2]	Lifetime [ps]	J^P
$B_s^0/\overline{B}_s^0$	$s\bar{b}/\bar{s}b$	5366.77 ± 0.24	1.512 ± 0.007	0^-
D_s^+/D_s^-	$c\bar{s}/\bar{c}s$	1968.30 ± 0.11	0.500 ± 0.007	0^-
K^+/K^-	$u\bar{s}/\bar{u}s$	493.677 ± 0.016	12380 ± 21	0^-
π^+/π^-	$u\bar{d}/\bar{d}s$	139.57018 ± 0.00035	26033 ± 5	0^-

4.3 Major steps of the analysis

To measure the *CP* violating observables defined in Sec. 2.4.2, it is necessary to perform a fit to the decay-time distribution of the selected $B_s^0 \to D_s^\mp K^\pm$ candidates. This fit has to be preceded by several preparatory stages. The general strategy of the analysis is presented in Fig. 4.2. All fits in this thesis are based on the unbinned maximum likelihood technique[6].

[6] A set of unknown parameters $\overrightarrow{\alpha} = \{\alpha_1, \alpha_2, ...\}$ is estimated from a set of measured observables $\overrightarrow{X} = \{X_1, X_2, ...\}$. The unbinned fit takes information about every single i candidate from the data set, for which the probability density function (PDF) $f(\overrightarrow{\alpha}; \overrightarrow{X})$ is normalized over the range of measured observables: $PDF(\overrightarrow{\alpha}; \overrightarrow{X}_i) = \frac{f(\overrightarrow{\alpha}; \overrightarrow{X}_i)}{\int f(\overrightarrow{\alpha}; \overrightarrow{X'}_i) d\overrightarrow{X'}_i}$ The likelihood function $\mathcal{L}$ is the product of the PDFs over the

Table 4.2: Branching fractions of B_s^0 decays as given by Ref. [90]; those of D_s^- decays are taken from the PDG [63], while the ratio f_s/f_d is the LHCb combination from hadronic/semi-leptonic decays [91].

Decay	Branching fraction
$B_s^0 \to D_s^{\mp} K^{\pm}$	$(1.90 \pm 0.12(\text{stat}) \pm 0.13(\text{syst}) \pm 0.14(f_s/f_d)) \times 10^{-4}$
$B_s^0 \to D_s^- \pi^+$	$(2.95 \pm 0.05(\text{stat}) \pm 0.17(\text{syst}) \pm 0.22(f_s/f_d)) \times 10^{-3}$
$B^0 \to D^- \pi^+$	$(2.68 \pm 0.13) \times 10^{-3}$
$D_s^- \to K^- K^+ \pi^-$	$(5.49 \pm 0.27) \times 10^{-2}$
$D_s^- \to K^- \pi^+ \pi^-$	$(6.9 \pm 0.5) \times 10^{-3}$
$D_s^- \to \pi^- \pi^+ \pi^-$	$(1.10 \pm 0.06) \times 10^{-2}$
$D^- \to K^+ \pi^- \pi^-$	$(9.54 \pm 0.26) \times 10^{-3}$
f_s/f_d	$0.259 \pm 0.015((\text{stat})\&(\text{syst}))$

The first step is improving the signal purity. Candidates have to pass several stages of selection, where each of them is tighter than the previous one. Both B_s^0 decay modes are selected using the same criteria, except for the requirement on the PIDK of the bachelor track. The optimisation of the last step, the offline selection, is performed using the control mode $B_s^0 \to D_s^- \pi^+$. A detailed description of the selection will be discussed in Chapter 5.

Then, it is necessary to distinguish the signal and background candidates in the selected sample. This analysis uses three variables to maximise sensitivity when discriminating between signal and background: the $m(D_s^{\mp} h^{\pm})$ invariant mass; the $m(h^- h^+ h^{\pm})$ mass; and the log-likelihood difference PIDK between the pion and kaon hypotheses for the bachelor particle. Here the $h^- h^+ h^{\pm}$ denotes possible combinations of the final state particles from the $D_s^{\pm}$ decay: $h^- h^+ h^{\pm} = \{K^+ K^- \pi^{\pm}, \pi^- \pi^+ K^{\pm}, \pi^- \pi^+ \pi^{\pm}\}$. Chapter 6 describes how a simultaneous extended maximum likelihood fit (in the following referred to as multidimensional fit) to these three variables is used to determine the yields of signal and background components in the samples of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates.

When the signal and background components are distinguished, the remaining inputs to the decay-time fits are information about the flavour of the B_s^0 meson production, the decay-time acceptance and resolution. Chapter 7 describes how to reconstruct the flavour at production of the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates using a combination of flavour-tagging algorithms, whose performance is calibrated with data using flavour-specific control modes. The decay-time resolution and acceptance are determined using a mixture of data control modes and simulated signal candidates, described in Chapter 8.

Finally, Chapter 9 presents the fit to the decay-time distribution of the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates which extracts the *CP* violating observables. The decay-time fit, henceforth referred to as the *sFit* 9, uses the results of the multidimensional fit to obtain the so-called *sWeights* [92] which allow the background components to be statistically subtracted [93]. The *sWeights* are defined by:

$$sWeights_n(y_e) = \frac{\sum_{j=1}^{N_s} V_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)}, \tag{4.2}$$

where V_{nj} means the covariance matrix resulting from the likelihood maximisation, N_k is

entire data sample: $\mathcal{L}(\overrightarrow{\alpha}; \overrightarrow{X}) = \prod_i^N PDF(\overrightarrow{\alpha}; \overrightarrow{X}_i)$ The fit algorithm maximises the likelihood by varying the parameters $\overrightarrow{\alpha}$. The solution is found for a set $\overrightarrow{\alpha}_{max}$, for which the probability is maximal.

the number of candidates expected on average for the k^{th} component, $f_i(y_e)$ is the value of the PDFs of the discriminating variables y for the i^{th} component and for candidate e, and N_s denotes the number of components. The *sPlot* technique [92] is implemented using the RooFit package [75]. The decay-time distribution is then subjected to the *sFit* procedure using only the PDF of the signal component.

The second decay-time fit, henceforth referred to as the *cFit*, performs a fit with full signal and background components description. However, the *cFit* is not a topic of this dissertation and is used only as a cross-check of the *sFit*.

Finally, the determination of the CKM angle γ is performed based on the result from the decay-time fit.

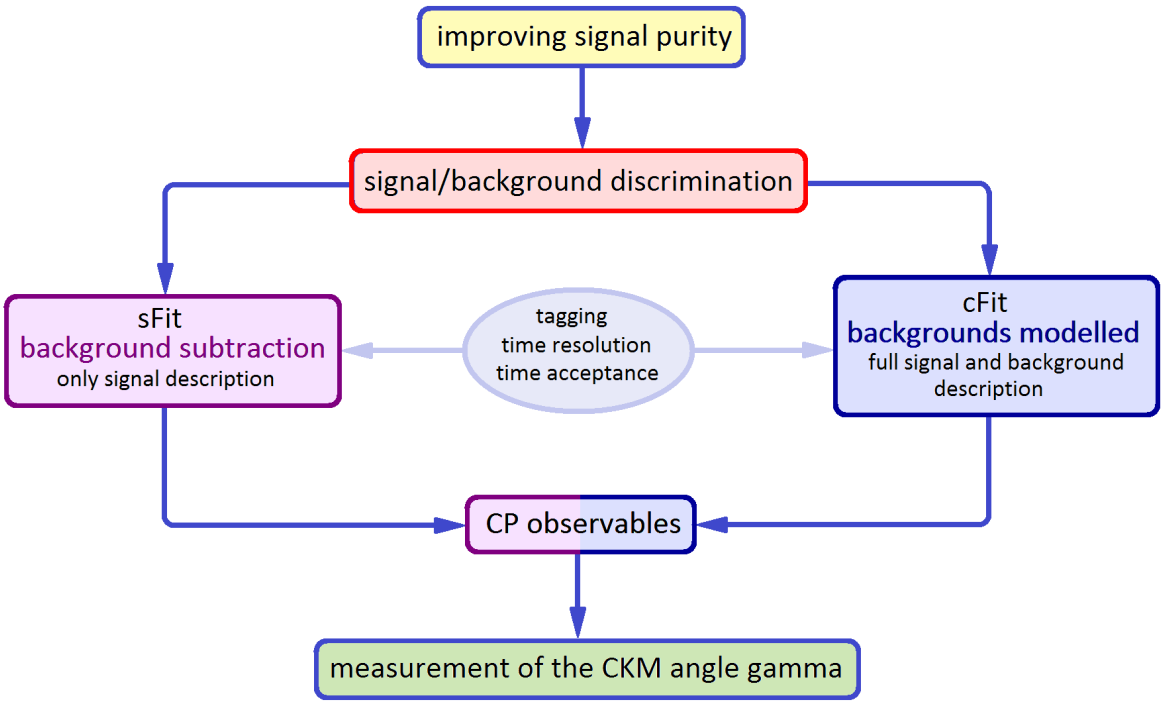

Figure 4.2: Strategy of the measurement the CKM angle γ from the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay.

Chapter 5

Selection

This chapter first describes the parameters relevant for the selection of signal candidates. This is followed by a short discussion of background contributions. Finally the subsequent steps of the selection are briefly described.

5.1 Definitions of kinematic variables

Due to the vast amount of background falling in the signal mass window, the selection of pure hadronic decays such as $B_s^0 \to D_s^{\mp} K^{\pm}$ is challenging. An efficient procedure requires using a lot of variables which help distinguish the signal from the background. Figure 5.1 shows the topology of $B_s^0 \to D_s^{\mp} K^{\pm}$ decay with a definition of the relevant concepts and parameters:

- primary vertex (PV) as defined in Sec. 3.2.1: the position of the pp collision,
- secondary vertex (SV) as defined in Sec. 3.2.1: formed by the decay of long-lived particles,
- mother particle vertex (MV): for the secondary particle it denotes the vertex where the particle comes from, for example the B_s^0 vertex comprises MV for D_s^- meson.
- Impact parameter (IP): the minimum perpendicular distance between the reconstructed track and the primary vertex. In the LHCb experiment the χ^2 impact parameter is commonly used, given by $\mathrm{IP}\chi^2 \sim \frac{\mathrm{IP}^2}{\sigma^2_{\mathrm{IP}}}$, where σ_{IP} is uncertainty on the impact parameter.
- Track χ^2/ndf: the fit χ^2 for a track divided by the number of degrees of freedom (ndf), where ndf depends on the number of hits in the detector.
- Vertex χ^2: as defined in Eq. 3.6. In the LHCb studies it is common to use the Vertex χ^2/ndf, where ndf depends on number of tracks associated with the vertex.
- Flight distance (FD): the distance between the primary vertex (or the mother vertex) and secondary vertex defined as $\mathrm{FD} = |\overrightarrow{r}_{P(M)V} - \overrightarrow{r}_{SV}|$, where $\overrightarrow{r}_{P(M)V} = (x_{P(M)V}, y_{P(M)V}, z_{P(M)V})$ are the coordinates of the primary (mother) vertex and $\overrightarrow{r}_{SV} = (x_{SV}, y_{SV}, z_{SV})$ corresponds to the secondary vertex coordinates. In addition two more variables connected to the flight distance are used:

  - flight distance χ^2 (FDχ^2) given by FD$\chi^2 \sim \frac{FD^2}{\sigma^2_{FD}}$, where FD is the three dimensional flight distance and σ_{FD} denotes its uncertainty,
  - radial flight distance (RFD) which is a projection into the xy plane equal to RFD $= \sqrt{(x_{P(M)V} - x_{SV})^2 + (y_{P(M)V} - y_{SV})^2}$.

- Direction of angle (DIRA): the cosine of the angle between the flight direction, defined as the distance between the primary vertex (or mother particle vertex) and secondary vertex of the particle, and the particle's reconstructed momentum:

$$DIRA = \frac{\overrightarrow{p} \cdot (\overrightarrow{r}_{P(M)V} - \overrightarrow{r}_{SV})}{|\overrightarrow{p}||\overrightarrow{r}_{P(M)V} - \overrightarrow{r}_{SV}|}, \tag{5.1}$$

  where $\overrightarrow{p}$ is the particle's momentum.

- DOCA: the distance of closest approach between two tracks (see Fig. 5.2).
- transverse momentum: momentum projection onto the z axis (see Fig. 5.2).
- cos(θ): where the angle θ is the angle between the momentum and z axis (see Fig. 5.2).
- track ghost probability (TGP): probability that the track is a ghost i.e. it does not have connection to any particle in the detector.

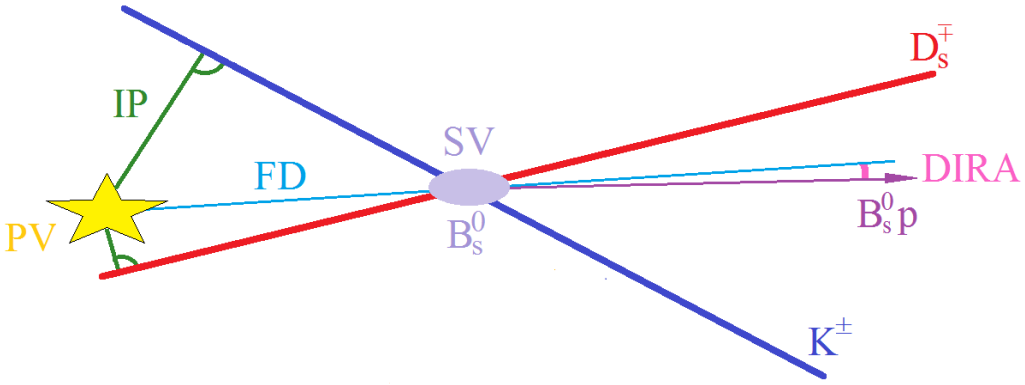

Figure 5.1: Diagram showing the topology of the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay together with the graphical representation of concepts and parameters relevant for the selection: PV - primary vertex, SV - secondary vertex, IP - impact parameter, FD - flight distance, DIRA - direction angle, $B_s^0 p$ - momentum of B_s^0 meson.

5.2 Background components

The following major sources of background have been considered in the analysis:

- **combinatorial background**:
  is composed of candidates where three pions or kaons which do not originate from a real D_s^- decay, but are reconstructed as a D_s^- candidate, are combined together with an additional kaon or pion and identified as a B_s^0 candidate. Alternatively, a true D_s^- is combined with a random track and reconstructed as a B_s^0 candidate,

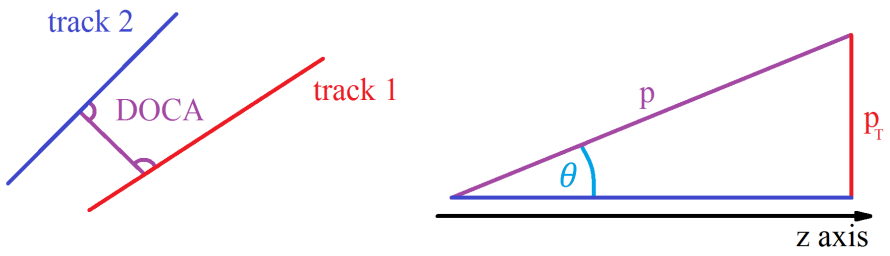

Figure 5.2: Left: definition of DOCA parameter (the distance of closest approach) between two tracks. Right: a particle's momentum projected into z axis with definition of the angle θ and transverse momentum.

- **fully reconstructed background**:
  corresponds to the cases where some of the final state particles are misidentified (misID) and/or where they originate from the decay of other particle i.e. not B_s^0/D_s^-,

- **partially reconstructed background**:
  consists of candidates for which one or more final state particles are missed, and others might be misidentified. It typically peaks in the lower mass range, but also may extend into the signal region.

Both fully and partially reconstructed backgrounds are composed of contributions due to several specific decays. Suppression of each of the above background categories in the data sample requires using different approaches, which are discussed later in this chapter.

5.3 Multivariate Analysis: Boosted Decision Trees Method

Multivariate Analysis (MVA) [94] is a family of methods based on statistical techniques that allow the use of information from more than one variable at time and take their correlations into account. Neural Networks (NN), Genetic Algorithms (GA) or Boosted Decision Trees (BDT) are examples of MVA techniques. The set of input variables (V_1,V_2,...,V_N) defines an N-dimensional space. Every candidate from the data set corresponds to a point in that space and can be classified as signal or background. The input information is combined into a single output variable (MVA Response), as shown in Fig. 5.3, which can be used in later steps of an overall selection. To find the best separation in N-dimensional space, MVA methods require a learning process in which pure signal and background samples are used as patterns. The data sample is split into "training" and "test" subsamples, where the first one is used in the learning process, while the latter to obtain the optimal cut point.

A BDT is used in each step of the selection, since this approach gives the best performance compared to other MVA methods. This method works as a binary tree classifier as shown in Fig. 5.4. The classifier finds the variable with the best separation power. Then, the data is split according to a cut on that the variable. The procedure is repeated among all available observables. The optimal cut value is found by the training algorithm and optimised the best significance. Depending on the majority of training candidates that end up in the final leaf node, the candidates are classified as signal-like or background-like. Boosting refers to use many active decision trees at the same time.

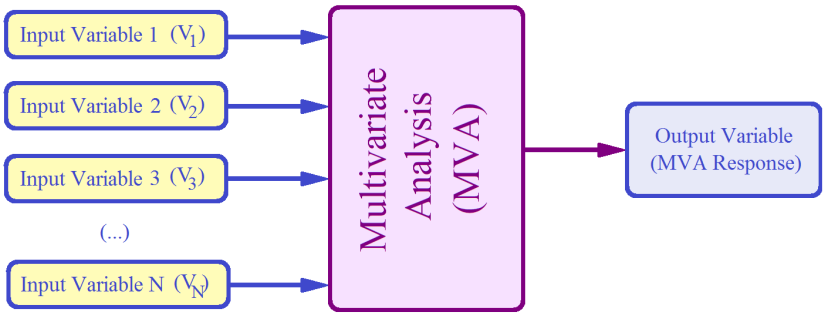

Figure 5.3: A flowchart of the overall multivariate analysis.

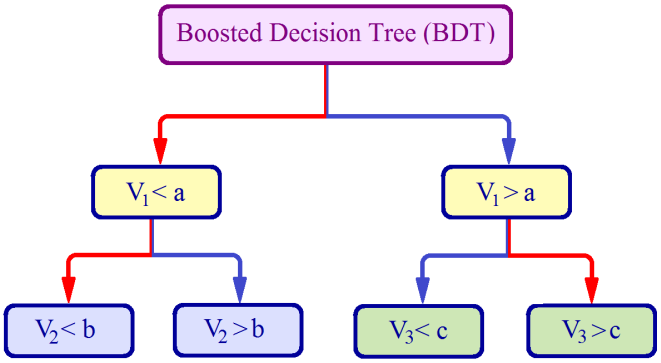

Figure 5.4: An example of a Boosted Decision Tree. The red (blue) arrows denote signal-like (background-like) decision.

This protects the decision tree against statistical fluctuations in the input samples and the situation when a candidate is wrongly assigned to a leaf. To stabilise the above procedure, the final response is a continuous number computed by combining the responses of the individual trees. In addition, the method has to be protected against overtraining, which occurs when the BDT is trained to recognised statistical fluctuations of the training sample. For this purpose the performance in both the training and test samples is compared, since any difference between results can indicate the overtraining. More information about BDTs can be found in Ref. [95, 96].

Several types of BDT approach exist. In the trigger selection the Bonsai BDT (BBDT) is applied, which allows the use of discrete input variables, while for the offline selection the best separation power was found for the Gradient BDT (BDTG).

The BDT approaches used in the trigger selection and preselection steps are common tools applied in Moore and DaVinci software, respectively. On the other hand, the offline selection BDTG is dedicated to provide the best signal/background separation for the $B^0_{(s)} \to D^{\mp}_{(s)} h^{\pm}$ decays. Here the classifier was trained specifically for this measurement and used by other analyses with similar decay chains Ref [2–6]. It is described in detail in Sec. 5.6.1.

5.4 Trigger selection

No explicit requirement is made on L0 triggers and all candidates with a positive L0 decision are kept. In addition, all candidates are required to be TOS or TIS on:

- the `1TrackAllL0` decision in the HLT1 trigger,

- the 2-, 3-, or 4-body `TopoBBDT` decisions or the inclusive `IncPhi` ($\phi^0 \to K^+K^-$) decision in the HLT2 trigger.

Detailed information about the requirements applied in each stage of the trigger selection can be found in App. C. It was checked that the response of the trigger decisions was stable during 2011 data taking.

Since the trigger selection can affect the decay-time acceptance, its correct description is crucial for the analysis. The decay-time acceptance is obtained based on a mixture of $B_s^0 \to D_s^- \pi^+$ data and simulated signal candidates, where all relevant effects are taken into account.

5.5 Preselection

The candidates which passed the trigger selection are processed in the preselection based on kinematic requirements. In that stage final state particles are combined and reconstructed into intermediate resonances ($D^-_{(s)}$, $\phi^0(1020)$, etc) as well as the $B^0_{(s)}$ meson. In addition, the decay-time and B_s^0 mass resolutions are improved by performing a kinematic fit [87] in which the B_s^0 candidate is constrained to originate from its associated proton-proton interaction, i.e. the one with the smallest IPχ^2 with respect to the B_s^0 candidate, and the B_s^0 mass is computed with a constraint on the nominal D_s^- mass as given by the PDG [63]. The preselection requirements are collected in Tab. C.4 in App. C. Figure 5.5 shows the $m(D_s^\mp h^\pm)$ invariant mass and decay-time distributions after the preselection for the $B_s^0 \to D_s^\mp h^\pm$ sample, where about 38 million candidates are reconstructed. In this stage the $B_s^0 \to D_s^\mp h^\pm$ sample is mostly composed of $B_s^0 \to D_s^- \pi^+$ candidates. The preselection efficiency was determined on signal $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ simulation samples and found to be of 5.6% and 5.5%, respectively.

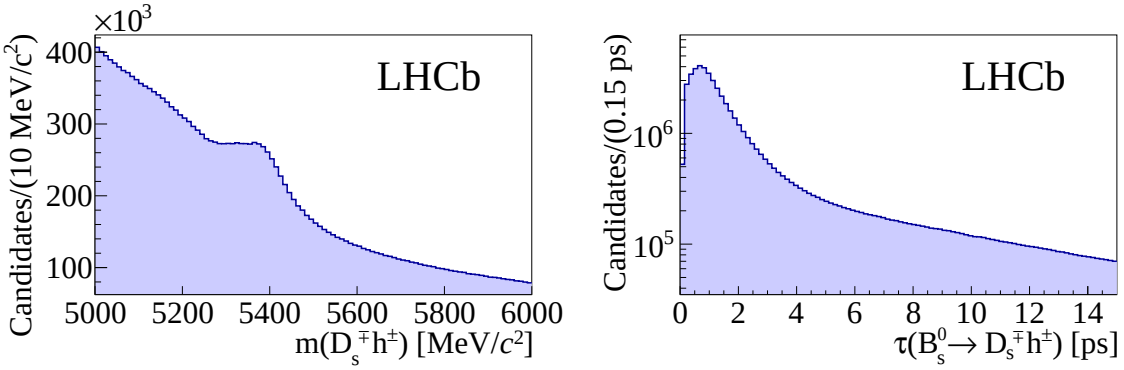

Figure 5.5: Distributions of two key quantities: $m(D_s^\mp h^\pm)$ invariant mass (left) and decay-time (right), obtained after the preselection for $B_s^0 \to D_s^\mp h^\pm$ data sample.

5.6 Offline selection

The offline selection comprises kinematic and geometric selection criteria. The core of the offline selection relies on a BDTG designed to suppress combinatorial background. An important role is played by the PID requirements, in particular the crucial cut on the particle identification of the bachelor particle (PIDK) which defines the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ samples. The selection consists of following stages:

1. BDTG training in a data-driven way,
2. mass and lifetime requirements,
3. definition of specific background vetoes,
4. definition of PID cuts on the D_s^- daughters,
5. optimisation of both the cut on the BDTG response and the PID cut on the bachelor track.

5.6.1 Boosted Decision Trees offline selection

In order to suppress combinatorial background, a set of kinematic variables is combined in a BDTG, implemented in the TMVA software package [94].

To train the BDTG, a data-driven approach is employed, using the 1.0 fb^{-1} $B_s^0 \to D_s^- \pi^+$, $D_s^- \to K^- K^+ \pi^-$ sample with an applied preselection and an additional loose selection described in detail in Tab.5.1. The data sample is split into training and test samples, both of which consist of equal parts of magnet up and magnet down data. The $m(D_s^- \pi^+)$ invariant mass after selection for both the training and test samples are shown in Fig. 5.6. Since all channels in this analysis are kinematically similar it is expected that the resulting BDT performs well also on the other samples.

Table 5.1: Additional preselection requirements applied to $B_s^0 \to D_s^- \pi^+$, $D_s^- \to K^- K^+ \pi^-$ used in BDTG optimisation.

Description	Requirement
Bachelor	PIDK < 0
Both kaons	PIDK > 0
$m(h^- h^+ h^\pm)$ mass	$[1940, 1990]\,\mathrm{MeV}/c^2$
D^- veto:	
PIDK of same charge K, or	> 10, or
$m(h^- h^+ h^\pm)$ under $m(K^\pm \pi^- \pi^+)$ hypothesis	below $1850\,\mathrm{MeV}/c^2$
$\overline{\Lambda}_c^-$ veto:	
p veto, same charge K, or	PIDK$-$PIDp > 5, or
$m(h^- h^+ h^\pm)$ under $m(\overline{p} K^+ \pi^\pm)$ hypothesis	not in $[2250, 2320]\,\mathrm{MeV}/c^2$

A training sample representative of the combinatorial background is selected from the upper B_s^0 mass data sideband defined as $m(D_s^- \pi^+) > 5445\,\mathrm{MeV}/c^2$. For the signal training sample an *sWeighted* data sample is used. The *sWeights* are extracted by fitting the $m(D_s^- \pi^+)$ invariant mass distributions shown in Fig. 5.6, using the full visible range of $[5100, 5800]\,\mathrm{MeV}/c^2$. The fit model is a very much simplified version of the final model, that will be described in Ch. 6, and consists of only two components, a Gaussian signal and an exponential background. Finally candidates are selected from a signal window in the $m(D_s^- \pi^+)$ invariant mass of $[5310, 5430]\,\mathrm{MeV}/c^2$, their *sWeights* rescaled to the range $[0, 1]$, and used in TMVA. The list of input BDTG variables is given in Tab. 5.2. Figure 5.7 shows the resulting distribution of the BDTG Response in the full [-1,1] range for both training and test samples, where the signal-like and background-like components are compared based on *sWeighted* data. The BDTG was checked for overtraining, for

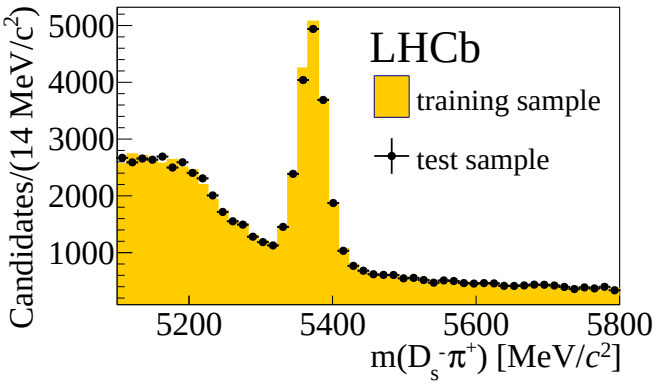

Figure 5.6: The $m(D_s^-\pi^+)$ invariant mass distribution after the selection for the $B_s^0 \to D_s^-\pi^+$ data samples used in the BDTG optimisation. The orange area corresponds to the training sample, while black points denote the test sample.

Table 5.2: Input variables to the BDTG method applied to suppress the combinatorial background.

B_s^0	D_s^-	Bachelor	D_s^- children
DIRA wrt. PV	DIRA wrt. PV	min IP_{χ^2}	min p_T
$\text{IP}\chi^2$	DIRA wrt. B_s^0 vertex	p_T	min $\text{IP}\chi^2$
RFD wrt. PV	min $\text{IP}\chi^2$	$\cos(\theta)$	
Vertex χ^2/ndf	RDF wrt. PV		
Vertex χ^2/ndf (D_s^- constrained)	Vertex χ^2/ndf	max TGP	

which the performance of the BDTG on both the training and test samples was compared. The performance was similar for both samples. In addition, the BDTG was compared to other BDT methods implemented in the TMVA package, and found to have the best performance.

5.6.2 Invariant masses and decay-time requirements

The B_s^0 and D_s^- candidates are required to be within $m(D_s^\mp h^\pm) \in [5300, 5800]\,\text{MeV}/c^2$ and $m(h^-h^+h^\pm) \in [1930, 2015]\,\text{MeV}/c^2$, respectively. In addition, the D_s^- decay-time with respect to the B_s^0 has to be greater than 0 ps, while for B_s^0 the decay-time should exceed 0.4 ps.

The $D_s^- \to K^-K^+\pi^-$ mode is split into three submodes. Candidates in which the K^+K^- pair falls within $20\,\text{MeV}/c^2$ of the $\phi^0(1020)$ mass are identified as belonging to $D_s^- \to \phi\pi^-$ decay. Candidates within a $50\,\text{MeV}/c^2$ window around the $K^{*0}(892)$ mass are identified as a $D_s^- \to K^{*0}K^-$ decay. It is kinematically impossible for a candidate to satisfy both $K^{*0}(892)$ and the $\phi^0(1020)$ requirements. The remaining candidates are referred to as non resonant decays.

The summary of the above requirements is given in Tab. 5.3.

5.6.3 Background vetoes

Specific background vetoes are necessary to reject the abundant D^- backgrounds coming from $B^0 \to D^-\pi^+$, where the $D^- \to K^+\pi^-\pi^-$ decay is reconstructed as a D_s^- by misiden-

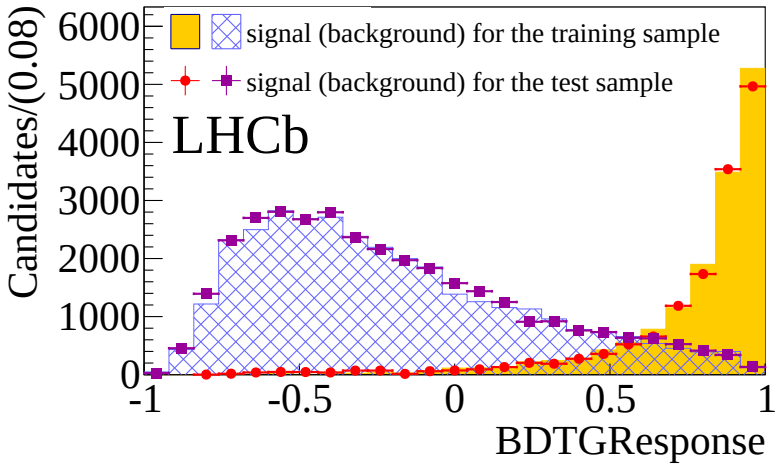

Figure 5.7: The BDTG Response distributions for the training and test samples. The orange area (red points) corresponds to signal weighted by the signal *sWeights*, while the blue dashed one (violet points) shows background weighted by the combinatorial background *sWeights* for training (test) sample.

Table 5.3: Summary of invariant mass and decay-time requirements for B_s^0 and D_s^- candidates.

Applied to	Description	Requirement
All modes	$m(D_s^\mp h^\pm)$ mass	$[5300, 5800]\,\text{MeV}/c^2$
	$m(h^-h^+h^\pm)$ mass	$[1930, 2015]\,\text{MeV}/c^2$
	B_s^0 decay-time	$> 0.4\,\text{ps}$
	D_s^- decay-time (wrt. B_s^0 vertex)	$> 0\,\text{ps}$
$D_s^- \to K^-K^+\pi^-$		
$D_s^- \to \phi\pi^-$	$m(K^+K^-)$	in $[1000, 1040]\,\text{MeV}/c^2$
$D_s^- \to K^{*0}K^-$	$m(K^+K^-)$	not in $[1000, 1040]\,\text{MeV}/c^2$
	$m(\pi^-K^+)$	in $[842, 942]\,\text{MeV}/c^2$
$D_s^- \to (KK\pi)_{\text{nonres}}$	$m(K^+K^-)$	not in $[1000, 1040]\,\text{MeV}/c^2$
	$m(\pi^-K^+)$	not in $[842, 942]\,\text{MeV}/c^2$

tifying a pion as a kaon. These vetoes also reject backgrounds from $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- h^+$ decays, where the proton from $\overline{\Lambda}_c^- \to \overline{p}K^+\pi^-$ is misidentified as a kaon. In addition, D^0 decays such as $D^0 \to K^+K^-$ are vetoed in order to remove *e.g.* $B^0 \to D^0K^{(*)0}$, $D^0 \to K^+K^-$ modes.

Furthermore, charmless backgrounds are particularly dangerous as they can peak directly under the region of the $m(D_s^\mp h^\pm)$ invariant mass. A requirement on the D_s^- separation χ^2 from the B_s^0 vertex is applied in order to suppress them. In the case of $D_s^- \to \pi^-\pi^+\pi^-$ and $D_s^- \to K^-\pi^+\pi^-$, a tighter requirement is used (> 9) compared to $D_s^- \to K^-K^+\pi^-$ (> 2) because the charmless decay $B_s^0 \to KKK\pi$ is suppressed through CKM factors relative to $B_s^0 \to KK\pi\pi$. Moreover, in the case of the $K\pi\pi\pi$ final state not only the B_s^0 decay but also the B^0 decay has to be considered. It can leak into the B_s^0 signal region once the D_s^- mass constraint is applied to the $\pi\pi\pi$ system. In order to study the possible size of any remaining charmless backgrounds the $m(h^-h^+h^\pm)$ sidebands

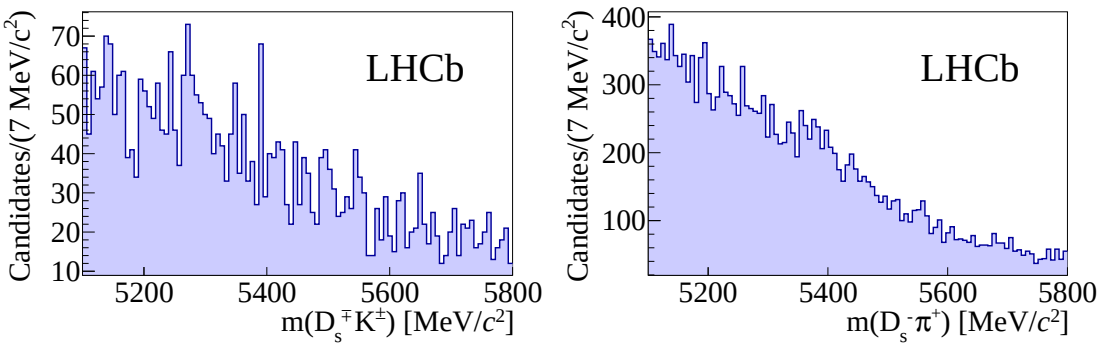

Figure 5.8: The invariant mass $m(D_s^{\mp}h^{\pm})$ from the D_s^- sideband defined as $m(h^-h^+h^{\pm}) = [1915, 1930] \cup [2030, 2050]\,\mathrm{MeV}/c^2$ obtained without a D_s^- mass constraint. Left: $B_s^0 \to D_s^{\mp}K^{\pm}$, right: $B_s^0 \to D_s^-\pi^+$ sample.

$(m(h^-h^+h^{\pm}) = (1915, 1930) \cup (2030, 2050)\,\mathrm{MeV}/c^2)$ are checked. The mass distributions produced without the D-mass constraint are shown in Fig. 5.8, where the lower $m(D_s^{\mp}h^{\pm})$ invariant mass region has been widened to 5100 MeV/c^2 for clarity. For the $B_s^0 \to D_s^{\mp}K^{\pm}$ sample no peaking charmless background is observed. No significant peaking backgrounds are seen in the $B_s^0 \to D_s^-\pi^+$ mass distributions either. At this level of statistical precision, there is no need to take charmless backgrounds into account in the measurement of $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^{\mp}K^{\pm}$ decays.

The specific background vetoes are collected in Tab. 5.4.

5.6.4 Particle identification

The candidates selected using the BDTG are filtered further using PID criteria on the final state tracks. The PID performance depends on the track momenta, so that a PID cut will change the momentum distribution of the track in question. As a consequence, the invariant mass distributions of the misidentified backgrounds may change. In is therefore important to determine the performance of the PID criteria as a function of particle's momentum.

The PID performance is determined from data using a sample of D^* decays [65]. For each PID cut applied in the analysis, its efficiency and misidentification probability on this D^* sample are obtained. This is evaluated in bins of the particle's momentum of and event occupancy as it mentioned in Sec. 3.2.2 and showed in Fig. 3.9.

The PID requirements were optimised for purity by looking in the $m(h^-h^+h^{\pm})$ sideband regions and tightening requirements until no significant peaking structures could be seen. In order to show that the cuts are reasonable, the candidates selected by the preselection which pass/fail the PID selection are shown in Fig. 5.9 split into $D_s^- \to K^-K^+\pi^-$, $D_s^- \to K^-\pi^+\pi^-$, and $D_s^- \to \pi^-\pi^+\pi^-$ candidates. This part of the selection is necessarily different for each D_s^- decay mode, as described below.

- For the $D_s^- \to \pi^-\pi^+\pi^-$ decay, none of the possible misidentified backgrounds fall inside the $m(h^-h^+h^{\pm})$ invariant mass range. Loose PID requirements are nevertheless used to identify the D_s^- decay products as pions in order to suppress combinatorial background.

- For the $D_s^- \to K^-\pi^+\pi^-$ mode, the relevant peaking backgrounds are $\overline{\Lambda}_c^- \to \overline{p}\pi^+\pi^-$ in which the antiproton is misidentified, and $D^- \to K^+\pi^-\pi^-$ in which both the

Table 5.4: Specific background vetoes applied in the selection.

Applied to	Description	Requirement
All	D_s^- FD (wrt. B_s^0 vertex)	> 0
$D_s^- \to K^-K^+\pi^-$	D_s^- FDχ^2 (wrt. B_s^0 vertex)	> 2
	D^0 veto:	
	$m(K^+K^-)$	$< 1840\,\mathrm{MeV}/c^2$
	D^- veto:	
	PIDK of same charge K, or	> 10, or
	$m(h^-h^+h^\pm)$ invariant mass	
	under $m(K^\pm\pi^-\pi^+)$ hypothesis	not in $[1840, 1900]\,\mathrm{MeV}/c^2$
	$\overline{\Lambda}_c^-$ veto:	
	p veto, same charge K, or	PIDK $-$ PIDp > 5, or
	$m(h^-h^+h^\pm)$ invariant mass	
	under $m(\overline{p}K^+\pi^\pm)$ hypothesis	not in $[2255, 2315]\,\mathrm{MeV}/c^2$
$D_s^- \to K^-\pi^+\pi^-$	D_s^- FDχ^2 (wrt. B_s^0 vertex)	> 9
	D^0 veto:	
	$m(K^+\pi^-)$	$< 1750\,\mathrm{MeV}/c^2$
	D^- veto:	
	PIDK of opposite charge π, or	< -10, or
	PIDK of opposite charge K, or	> 20, or
	$m(h^-h^+h^\pm)$ invariant mass	
	under $m(K^\pm\pi^-\pi^+)$ hypothesis	not in $[1839, 1899]\,\mathrm{MeV}/c^2$
	$\overline{\Lambda}_c^-$ veto:	
	p veto, same charge K, or	PIDK $-$ PIDp > 5, or
	$m(h^-h^+h^\pm)$ invariant mass	
	under $m(\overline{p}K^+\pi^\pm)$ hypothesis	not in $[2255, 2315]\,\mathrm{MeV}/c^2$
$D_s^- \to \pi^-\pi^+\pi^-$	D_s^- FDχ^2 (wrt. B_s^0 vertex)	> 9
	D^0 veto:	
	Both $m(\pi^+\pi^-)$	$< 1700\,\mathrm{MeV}/c^2$

kaon and a pion are misidentified. As this process is characterised by the smallest branching fraction to compare with the other D_s^- decay modes studied here, and hence it is the one which is the most affected by background, all D_s^- decay products are required to pass tight PID requirements.

- For the $D_s^- \to \phi\pi^-$ decay, since the invariant mass requirement suppresses most of the specific and combinatorial background, only loose PID requirements are needed.

- For the $D_s^- \to K^{*0}K^-$ mode, there is non-negligible background from misidentified $D^- \to K^+\pi^-\pi^-$ and $\overline{\Lambda}_c^- \to \overline{p}K^+\pi^-$ decays which are suppressed through tight PID requirements on the D_s^- kaon with the same charge as the D_s^- pion.

- For the $D_s^- \to (KK\pi)_{\mathrm{nonres}}$ decay, tight PID requirements are needed to suppress

background contributions.

The details of the PID selection are given in Tab. 5.5.

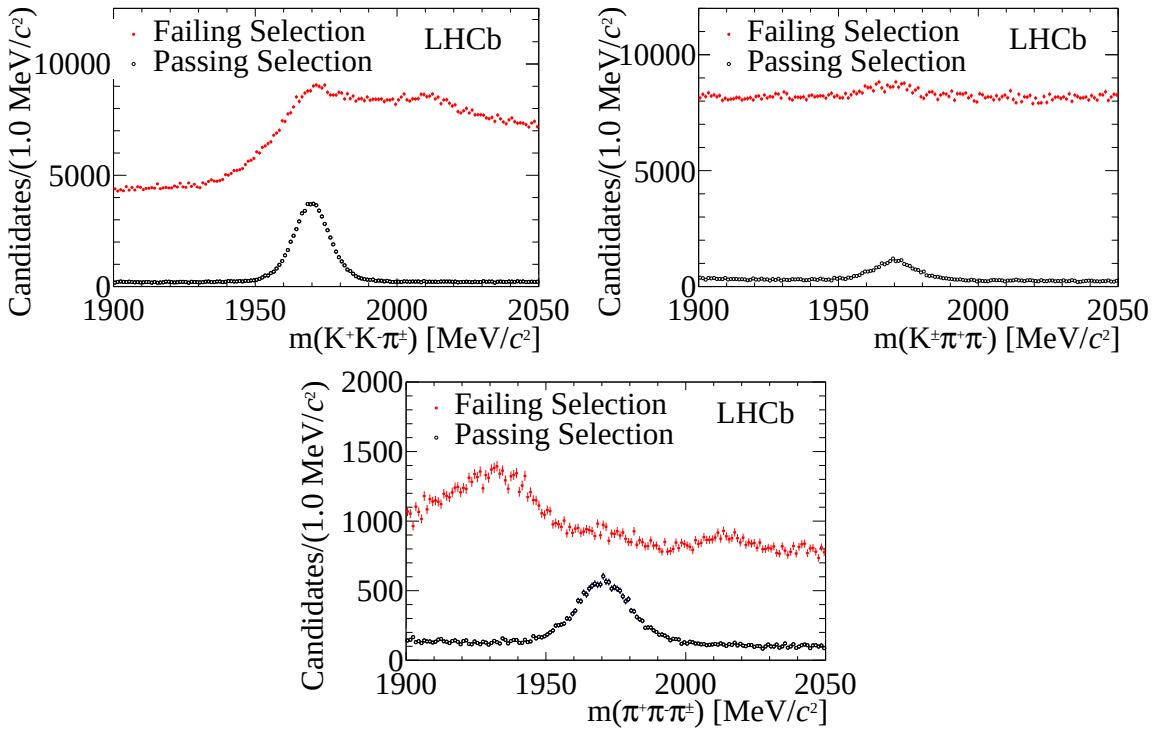

Figure 5.9: Candidates failing and passing the PID selection on applied on the $D^-_{(s)}$ decay products for $D^-_s \to K^-K^+\pi^-$, $D^-_s \to K^-\pi^+\pi^-$, and $D^-_s \to \pi^-\pi^+\pi^-$ (top left to bottom).

Table 5.5: The PID requirements.

Applied to	Description	Requirement
Bachelor track	$B^0_s \to D^-_s\pi^+$	PIDK < 0
	$B^0_s \to D^\mp_s K^\pm$	PIDK > 5
$D^-_s \to K^-K^+\pi^-$		
$D^-_s \to \phi\pi^-$	Both kaons	PIDK > -2
$D^-_s \to K^{*0}K^-$	Same charge kaon	PIDK > 5
	Opposite charge kaon	PIDK > -2
$D^-_s \to (KK\pi)_{\text{nonres}}$	Both kaons	PIDK > 5
$D^-_s \to K^-\pi^+\pi^-$	Kaon	PIDK > 10
	Both pions	PIDK < 5
	Both pions	PID$p < 10$
$D^-_s \to \pi^-\pi^+\pi^-$	All pions	PIDK < 10
	All pions	PID$p < 10$

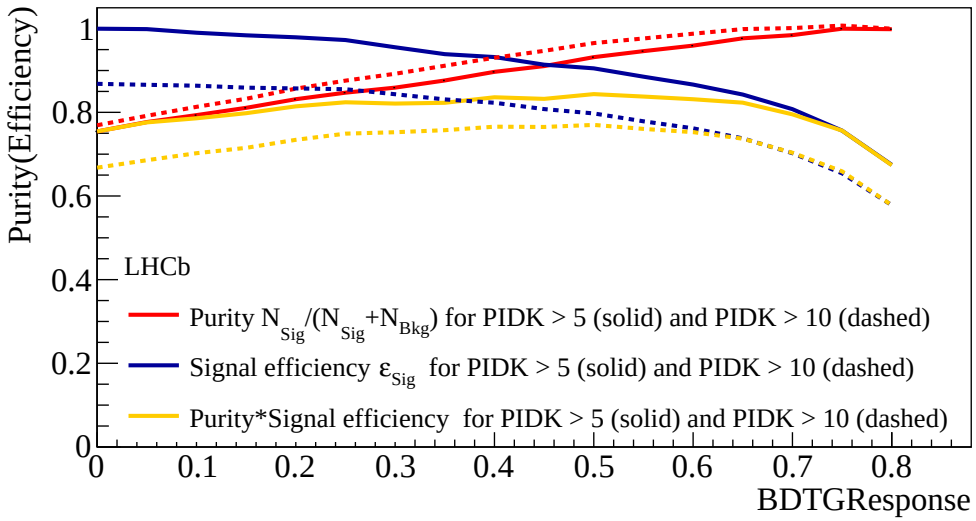

Figure 5.10: The optimisation curve (in the range $[0, 1]$) of the BDTG Response variable shown for all D_s^- final states.

5.6.5 Selection optimisation

The most important requirements for the study in question are optimised for maximum significance. These cuts are the BDTG Response, which suppresses combinatorial background; and the bachelor PID, which mostly controls the specific backgrounds, but also suppresses combinatorial background. The signal yield purity is defined as:

$$P_{Sig} = \frac{N_{Sig}}{N_{Sig} + N_{Bkg}} \ , \tag{5.2}$$

where N_{Sig} and N_{Bkg} are the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal and the sum of all background yields in the $m(D_s^{\mp} K^{\pm})$ invariant mass region [5320,5420] MeV/c^2, respectively. The optimisation curve is given by:

$$O_{Sig} = P_{Sig} \times \epsilon_{Sig}, \tag{5.3}$$

where ϵ_{Sig} is the signal efficiency in order to take into account efficiency effects. Figure 5.10 shows the optimisation curve and its components in a range of the BDTG Response $\in [0, 1]$ and taking into account all D_s^- final states. The best performance is found for PIDK > 5. In addition, since the optimisation curve has a plateau, BDTG Response > 0.3 and PIDK > 5 are set as nominal cut values.

5.6.6 Key distributions after offline selection

In Fig. 5.11 and Fig. 5.12 the distributions of $m(D_s^{\mp} h^{\pm})$ invariant mass and decay-time are shown after the full offline selection for $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ decays, respectively. The 6113 ($B_s^0 \to D_s^{\mp} K^{\pm}$) and 40260 ($B_s^0 \to D_s^- \pi^+$) candidates are reconstructed. The signal efficiency of the offline selection measured on the signal simulation sample and with respect to preselection is 58% (54%) for $B_s^0 \to D_s^- \pi^+$ ($B_s^0 \to D_s^{\mp} K^{\pm}$) decay mode.

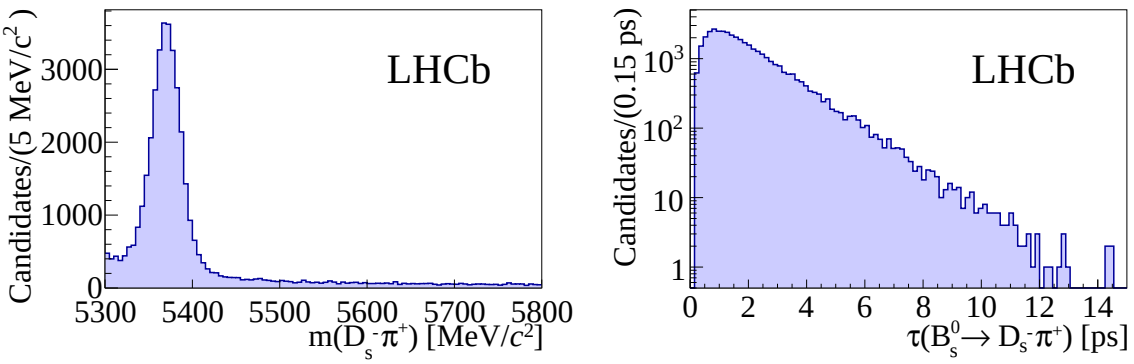

Figure 5.11: Distributions of two key quantities: $m(D_s^-\pi^+)$ invariant mass (left) and decay-time (right), obtained after the offline selection for the $B_s^0 \to D_s^-\pi^+$ data sample.

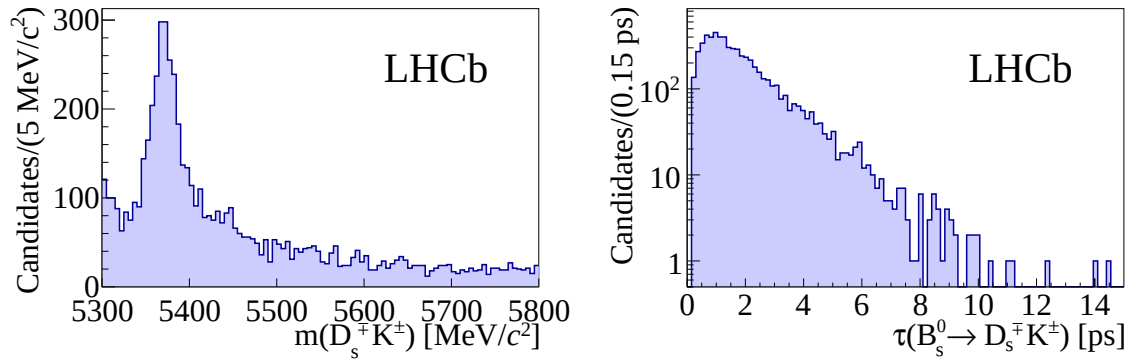

Figure 5.12: Distributions of two key quantities: $m(D_s^\mp K^\pm)$ invariant mass (left) and decay-time (right), obtained after the preselection selection for the $B_s^0 \to D_s^\mp K^\pm$ data sample.

Chapter 6

Multidimensional fit

The Multidimensional fit (MD fit) is performed to the $m(D_s^{\mp}h^{\pm})$ and $m(h^-h^+h^{\pm})$ invariant masses and bachelor PIDK. The following list summaries the key points of the MD fit strategy:

- All fits are unbinned extended maximum likelihood fits implemented in the `B2DXFitters` package [89].
- The distinction between the signal and backgrounds with different charm particles such as D_s^-, D^- or $\overline{\Lambda}_c^-$ hadrons is increased by fitting to the $m(h^-h^+h^{\pm})$ invariant mass.
- The discrimination between modes with different bachelor particles such as kaon, pion or proton is significantly enhanced by using the PIDK variable as an observable.
- The multidimensional fit is performed to both of the $B_s^0 \to D_s^{\mp}h^{\pm}$ samples.
- The subsamples defined by the D_s^- final states are fitted simultaneously.
- As shown in Fig. 6.1 the correlation between variables are negligible, therefore the full PDF for each component can be factorised:

$$\begin{aligned}\mathrm{PDF}(a,b,c;m(D_s^{\mp}h^{\pm}),m(h^-h^+h^{\pm}),\mathrm{PIDK}) = \\ \mathrm{PDF}(a;m(D_s^{\mp}h^{\pm}))\times\mathrm{PDF}(b;m(h^-h^+h^{\pm}))\times\mathrm{PDF}(c;\mathrm{PIDK}).\end{aligned} \tag{6.1}$$

  where a, b and c are parameters, whereas $m(D_s^{\mp}h^{\pm})$, $m(h^-h^+h^{\pm})$ and PIDK denote observables.
- The considered fit ranges are $m(D_s^{\mp}h^{\pm}) \in [5300, 5800]\,\mathrm{MeV}/c^2$, $m(h^-h^+h^{\pm}) \in [1930, 2015]\,\mathrm{MeV}/c^2$. Since the PID selection is different between $D_s^-\pi^+$ and $D_s^{\mp}K^{\pm}$ the following range are used: $D_s^-\pi^+$: $-\mathrm{PIDK} \in [0, 150]$, $D_s^{\mp}K^{\pm}$: $\ln(\mathrm{PIDK}) \in [\ln(5), \ln(150)]$.
- The different PID performance in both magnet polarities is taken into account as well as different background contributions to the D_s^- modes.
- The PDFs of the signal mass distributions are determined by the detector resolution and the radiative character of the $B_{(s)}^0$ and $D_{(s)}^-$ meson decays. For these reasons, the signal invariant mass distributions contain a tail extending towards lower masses.

The description requires using Crystal Ball [97] functions. A single Crystal Ball function is defined as:

$$\mathrm{PDF}_{CB}(\alpha, n, \mu, \sigma; x) = \begin{cases} e^{-\frac{(x-\mu)^2}{2\sigma^2}} & \text{for } \frac{(x-\mu)^2}{2\sigma^2} > -\alpha, \\ A \times (B - \frac{x-\mu}{\sigma})^{-n} & \text{for } \frac{(x-\mu)^2}{2\sigma^2} \leq -\alpha, \end{cases} \tag{6.2}$$

where A and B are given by:

$$A = \left(\frac{n}{|\alpha|}\right) \times e^{-\frac{1}{2}|\alpha|^2}, \qquad B = \frac{n}{|\alpha|} - |\alpha|. \tag{6.3}$$

and (α, n, μ, σ) is set of free parameters. A signal PDF is parametrised by a double Crystal Ball function with: common mean μ, oppositely oriented tails ($\alpha^1 > 0$, $\alpha^2 < 0$) and depends in total on eight free parameters:

$$\begin{aligned} \mathrm{PDF}_{dCB}(\alpha^1, n^1, \sigma^1, \alpha^2, n^2, \sigma^2, \mu, f; x) = f \times \mathrm{PDF}_{CB}(\alpha^1, n^1, \sigma^1, \mu; x) \\ + (1-f) \times \mathrm{PDF}_{CB}(\alpha^2, n^2, \sigma^2, \mu; x), \end{aligned} \tag{6.4}$$

where f denotes fraction between components.

- The functional form of the combinatorial PDF is taken from the $m(D_s^{\mp} h^{\pm})$ invariant mass sideband.
- The specific background PDFs make heavy use of so-called templates, which are predefined shapes that have no free parameter other than the normalisation and mostly obtained using kernel method [98] implemented in the `RooKeysPdf` class in the RooFit framework [75].
- When computing all PDFs the magnet polarities are summed and weighted according to the collected integrated luminosity.
- The differences between data and simulation are reduced using the procedure described in Sec. 6.1.1.
- Based on the MD fit results, for each candidate the *sWeight* given by Eq. 4.2 is obtained.

6.1 $B^0 \to D^- \pi^+$ control samples

The $B^0 \to D^- \pi^+$, $D^- \to K^+ \pi^- \pi^-$ control mode is used in several aspects of the analysis. For this purpose, two $B^0 \to D^- \pi^+$ samples are obtained, called "control sample A" and "control sample B", which differ only in an extra cut of PIDK < 0 for the bachelor pion in the case of control "sample B". The remaining selection criteria are common and listed in Tab. D.1 (App. D).

Control sample A is used to reduce data/simulation discrepancies. The PID requirement on the bachelor particle is not applied since it would affect its kinematics. This is described in Sec. 6.1.1. Control sample B is used to (i) define the widths of the nominal B_s^0 and D_s^- signal PDFs, and (ii) to determine the expected number of $B^0 \to D^- \pi^+$ and $B^0 \to D^- K^+$ background yields in the nominal $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ samples.

Both control samples use a two dimensional fit to the $m(D^- \pi^+)$ and $m(K^+ \pi^- \pi^-)$ invariant mass distributions in the ranges $m(D^- \pi^+) \in [5000, 6000]\,\mathrm{MeV}/c^2$ and $m(K^+ \pi^- \pi^-) \in$

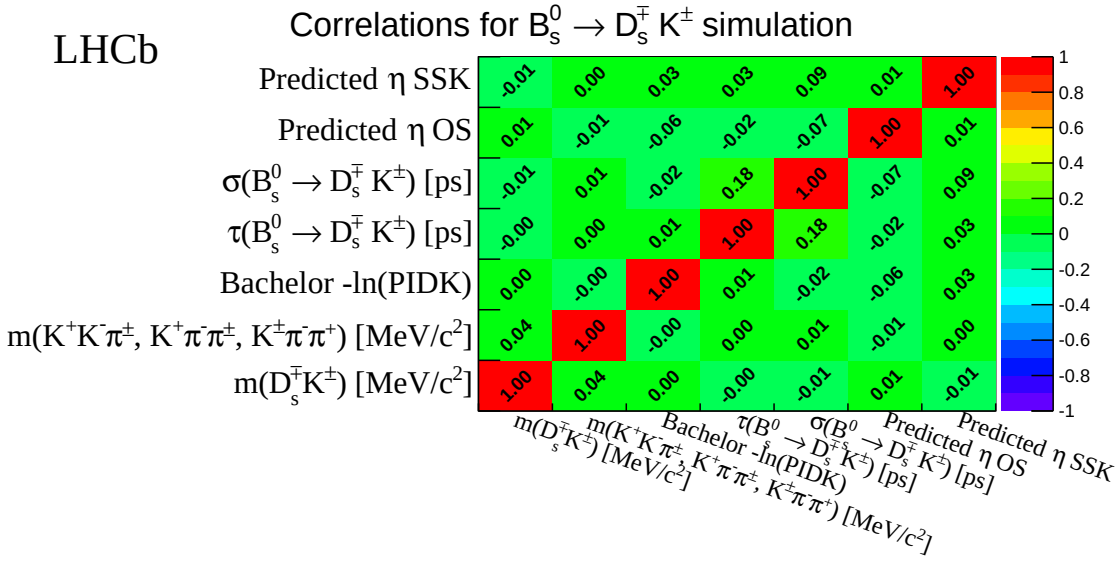

Figure 6.1: Correlations between the variables used in the analysis, as obtained on the signal $B_s^0 \to D_s^\mp K^\pm$ simulation sample.

$[1830, 1910]\,\mathrm{MeV}/c^2$. Because of the additional PID requirement, the two control samples require separate signal and background PDFs, however much of the strategy is common.

The signal PDFs are taken to be double Crystal Ball functions:

- the $m(D^-\pi^+)$ invariant mass:

$$\begin{aligned}\mathrm{PDF}(p_{B^0}; m(D^-\pi^+))_{B^0\to D^-\pi^+} = \\ \mathrm{PDF}_{dCB}(\alpha^1_{B^0}, n^1_{B^0}, \sigma^1_{B^0}, \alpha^2_{B^0}, n^2_{B^0}, \sigma^2_{B^0}, f_{B^0}, \mu_{B^0}; m(D^-\pi^+)),\end{aligned} \tag{6.5}$$

- the $m(K^+\pi^-\pi^-)$ invariant mass:

$$\begin{aligned}\mathrm{PDF}(p_{D^-}; m(K^+\pi^-\pi^-))_{B^0\to D^-\pi^+} = \\ \mathrm{PDF}_{dCB}(\alpha^1_{D^-}, n^1_{D^-}, \sigma^1_{D^-}, \alpha^2_{D^-}, n^2_{D^-}, \sigma^2_{D^-}, f_{D^-}, \mu_{D^-}; m(K^+\pi^-\pi^-)).\end{aligned} \tag{6.6}$$

whose parameters are obtained from a fit to simulated signal $B^0 \to D^-\pi^+$ candidates. In the nominal fit the mean $\mu_{B^0(D^-)}$ and two widths $\sigma^1_{B^0(D^-)}$ and $\sigma^2_{B^0(D^-)}$ are left free. As an example the parametrisations of the signal PDFs for control sample A are shown in Fig. 6.2. The numerical results for both control samples are collected in Tab. D.2 for the $m(D^-\pi^+)$ invariant mass and in Tab. D.3 for the $m(K^+\pi^-\pi^-)$ invariant mass (App. D).

In the model three fully and two partially reconstructed backgrounds are considered. The first group is composed of $B^0 \to D^-K^+$, $B_s^0 \to D_s^-\pi^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^-\pi^+$ decays. In case of the $B^0 \to D^-K^+$ mode, a bachelor kaon K^+ is misidentified as a pion π^+, for $B_s^0 \to D_s^-\pi^+$, $D_s^- \to K^-K^+\pi^-$ decays, a kaon K^- from the D_s^- meson is misidentified as a π^-. Finally, for $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^-\pi^+$, $\overline{\Lambda}_c^- \to \overline{p}K^+\pi^-$ mode, a double misidentification is considered: $\overline{p} \to K^-$ and $K^+ \to \pi^+$. The partially reconstructed backgrounds are due to $B^0 \to D^-\rho^+$ and $B^0 \to D^{*-}\pi^+$ decays. In both cases the neutral particle is missed from $\rho^+ \to \pi^+\pi^0$ and $D^{*-} \to D^-(\pi^0, \gamma)$ decays, respectively. The specific background PDFs are generally obtained using the kernel method from dedicated simulation samples. For the backgrounds which decay to D^- or D^{*-} ($B^0 \to D^-K^+$, $B^0 \to D^-\rho^+$ and $B^0 \to D^{*-}\pi^+$) the signal description in the $m(K^+\pi^-\pi^-)$ invariant mass is used. The resulting PDFs for control sample A are shown in Fig. D.2 (App. D).

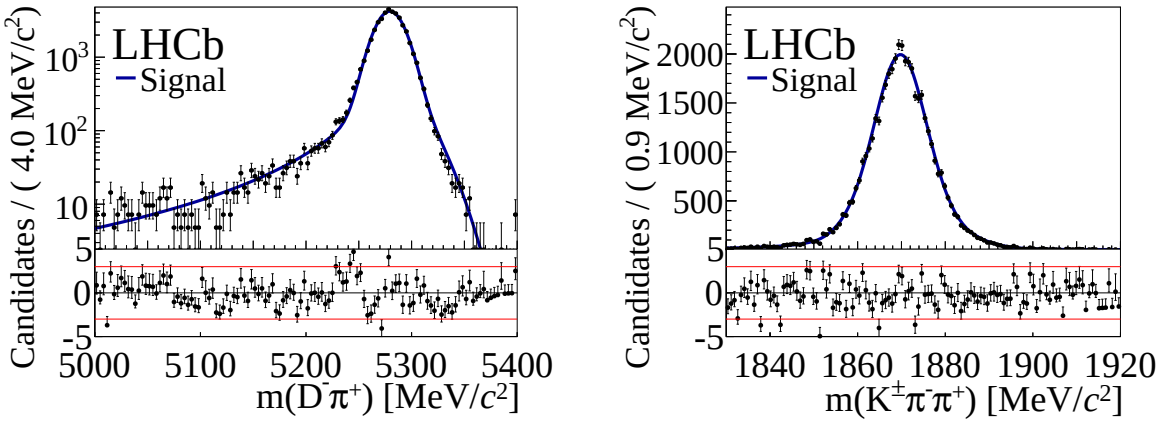

Figure 6.2: Distributions of the $m(D^-\pi^+)$ (left) and $m(K^\pm\pi^-\pi^+)$ (right) invariant masses as obtained on simulation samples with combined magnet polarities (data points). The solid, blue curve corresponds to the fit described in the text.

The combinatorial background is parametrised by a double exponential in the $m(D^-\pi^+)$ invariant mass dimension:

$$\begin{aligned}\text{PDF}(s^1_{B^0}, s^2_{B^0}, fC_{B^0}; m(D^-\pi^+))_{Comb} = \\ fC_{B^0} \times e^{-s^1_{B^0}\times m(D^-\pi^+)} + (1 - fC_{B^0}) \times e^{-s^2_{B^0}\times m(D^-\pi^+)},\end{aligned} \tag{6.7}$$

where $s^1_{B^0}$ ($s^2_{B^0}$) is the slope of the first (second) exponential and fC_{B^0} is fraction between them. To describe the combinatorial component that contains a true D^-, the $m(K^+\pi^-\pi^-)$ invariant mass PDF is given by a single exponential plus the signal double Crystal Ball function:

$$\begin{aligned}\text{PDF}(p_{D^-}, s_{D^-}, fC_{D^-}; m(K^+\pi^-\pi^-))_{Comb} = \\ fC_{D^-} \times e^{-s_{D^-}\times m(h^-h^+h^\pm)} + (1 - fC_{D^-}) \times \text{PDF}_{dCB}(p_{D^-}; m(K^+\pi^-\pi^-)),\end{aligned} \tag{6.8}$$

where fC_{D^-} denotes fraction between components, s_{D^-} is slope of the exponential and p_{D^-} are shared parameters with the signal description. The result of fitting the combinatorial shapes to the B^0 sideband is given in Fig. D.1 (App. D) as an example for control sample A. This result is used to define a fit model and reasonable starting parameters for the combinatorial background, but the actual parameters float in the fits.

6.1.1 Control sample A: Data/Simulation corrections

The description of many of the components in the MD fit relies on simulated samples, which must be reweighted in order to improve their agreement with the data. Control "sample A" is used to measure data/simulation weights in bins of two sensitive kinematic variables: ln(nTracks) in the candidates, where nTracks denotes track multiplicity, and logarithm of the bachelor transverse momentum $\ln(p_T)$. The resulting weight matrix is applied to all other simulation samples used in this analysis.

The overall result is shown in 6.3. The obtained numerical results are collected in Tab. D.4 (App. D).

To obtain the pure $B^0 \to D^-\pi^+$ sample the fitted candidates are weighted by resulting *sWeights*. Then two-dimensional histograms of $\ln(p_T)$ vs. ln(nTracks) with 20

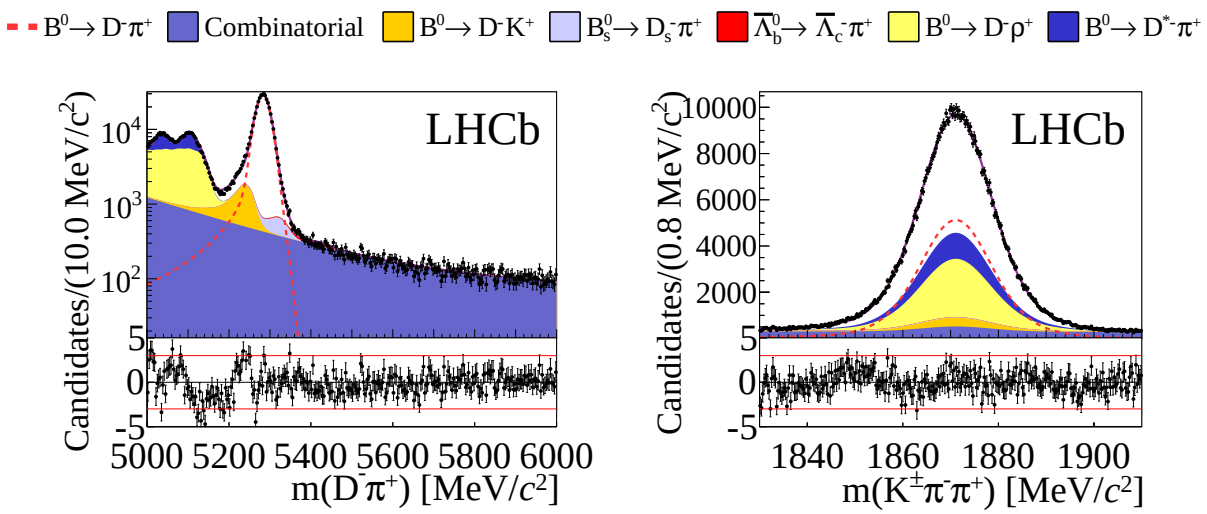

Figure 6.3: Distributions of the $m(D^-\pi^+)$ (left) and $m(K^\pm\pi^-\pi^+)$ (right) invariant masses with combined magnet polarities (data points). The solid, blue curve corresponds to the fit described in the text. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend above the plots.

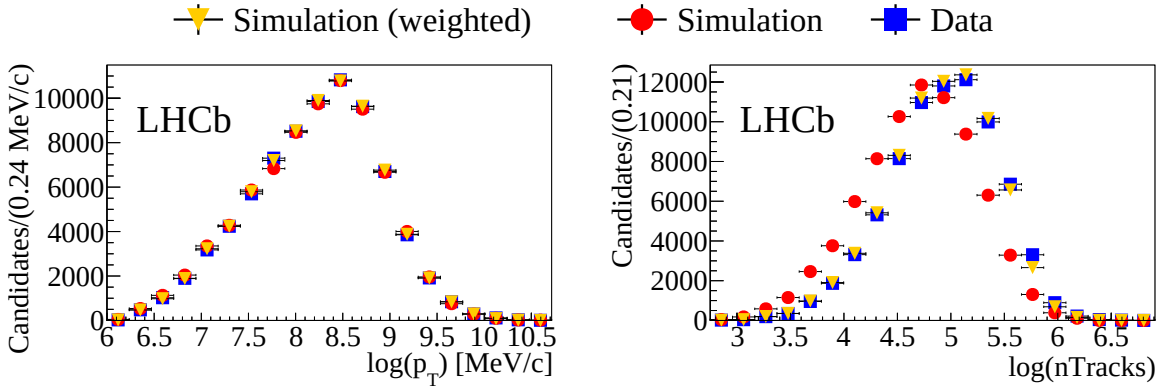

Figure 6.4: One-dimensional projections illustrating the two-dimensional data/simulation weighting, left for $\ln(p_T)$ and right for ln(nTracks) variables.

bins in each direction are compared between data and simulation. Their ratios form the desired correction matrix. The procedure is repeated separately for each magnet polarity. One-dimensional projections, together with the distributions of weighted simulation are presented in Fig. 6.4. Some of the bins receive large corrections when reweighting the simulation, which demonstrates that this procedure is indeed needed.

6.1.2 Control sample B: inputs to $B_s^0 \to D_s^\mp h^\pm$ sample fits

The fit to the $B^0 \to D^-\pi^+$ control "sample B" is used to determine the ratios of widths of the nominal B_s^0 and D_s^- signal PDFs, and a number of background yields that need to be fixed in the nominal $B_s^0 \to D_s^\mp h^\pm$ fits. For this purpose all signal and background templates were redone taking into account the influence of the PIDK < 0 requirement. In the "control $B^0 \to D^-\pi^+$ sample B", the number of $B^0 \to D^-K^+$ candidates needs to be fixed. It is estimated from a prediction based on the efficiencies of the involved PID cuts and branching fractions, and fixed to be $N_{B^0 \to D^-K^+} = 1825$. The overall results of the

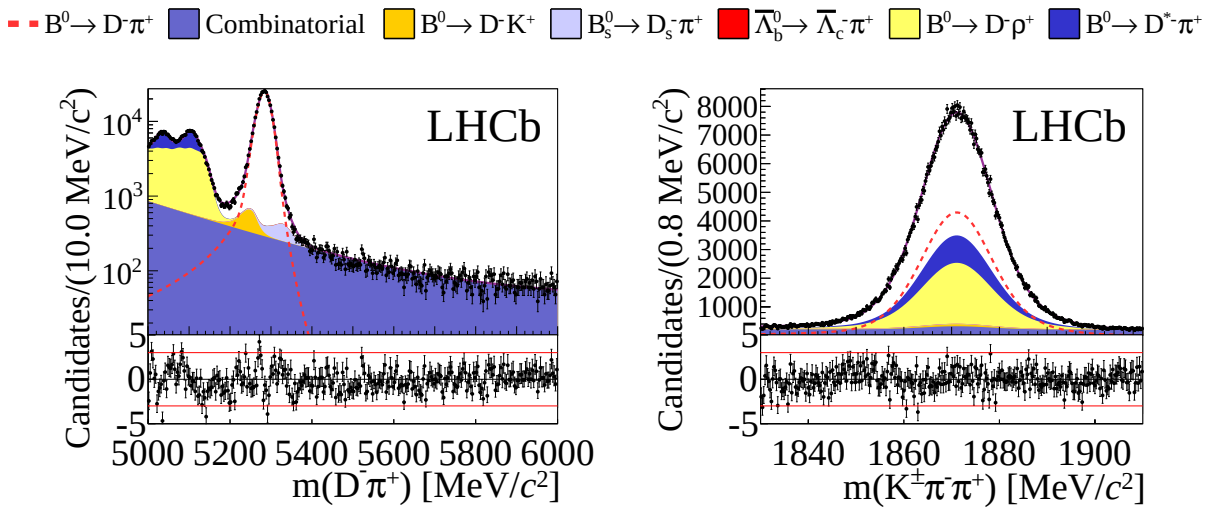

Figure 6.5: Distributions of the $m(D^-\pi^+)$ (left) and $m(K^\pm\pi^-\pi^+)$ (right) invariant masses with combined magnet polarities (data points). The solid, blue curve corresponds to the fit described in the text. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend above the plots.

Table 6.1: Widths of the double Crystal Ball function in $B^0 \to D^-\pi^+$ data and in the respective simulation together with the resulting scale factors.

Parameter	Fitted value in simulation [MeV/c^2]	Fitted value in data [MeV/c^2]	Scale factors
$\sigma^1_{B^0}$	8.60 ± 0.75	11.02 ± 0.25	1.28 ± 0.11
$\sigma^2_{B^0}$	14.54 ± 0.11	17.69 ± 0.08	1.22 ± 0.01
$\sigma^1_{D^-}$	9.98 ± 1.49	11.58 ± 0.07	1.16 ± 0.17
$\sigma^2_{D^-}$	5.63 ± 0.33	6.68 ± 0.28	1.19 ± 0.07

control $B^0 \to D^-\pi^+$ sample B are shown in Fig. 6.5 and collected in Tab. D.5 (App. D). The signal $B^0 \to D^-\pi^+$ yield is reconstructed to be $N_{B^0 \to D^-\pi^+} = 109420$. The resulting signal widths are summarised in Tab. 6.1, together with the desired data/simulation scale factors.

6.2 Signal and background description

The signal and background description is crucial for performing an unbiased fit. Their parametrisations are taken either from data or simulated candidates. A special description based on data is provided for the $B^0 \to D^-\pi^+$ mode as a background to the $B^0_s \to D^-_s\pi^+$ decay and the $B^0_s \to D^-_s\pi^+$ mode as a background to the $B^0_s \to D^\mp_s K^\pm$ decay. In addition, sideband data are used for the combinatorial background. Other parametrisation are obtained on simulated candidates. The discrepancy between the simulated candidates and data samples is a common issue in the correct modelling of PDFs. The most affected variables are kinematic quantities such as momentum, track multiplicity etc. In order to improve the data-simulation agreement, the correction described in Sec. 6.1.1 is used. In addition, the full trigger, preselection and offline kinematic selection chain is applied. Simulation samples are also corrected using misidentification and/or efficiency of the

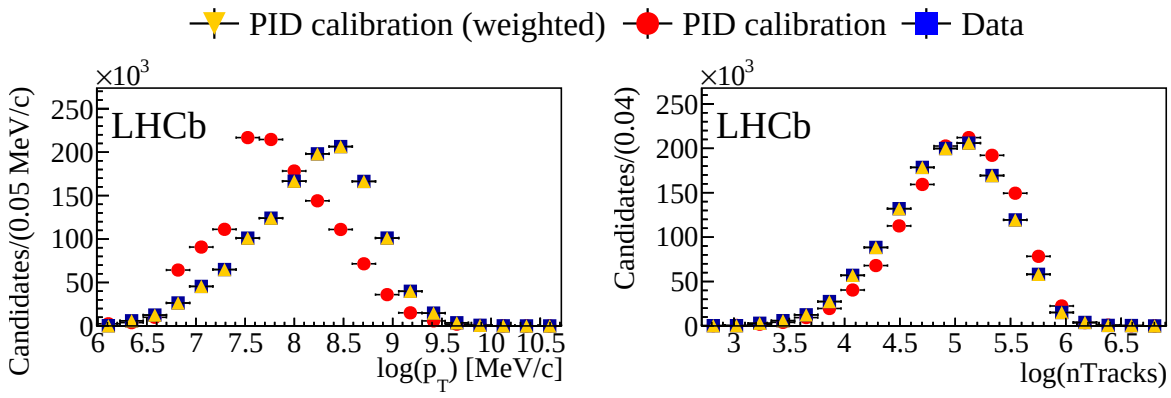

Figure 6.6: One-dimensional projections of the two-dimensional weighting of the PID calibration samples to match the $B^0 \to D^-\pi^+$ decay kinematics. Left: $\ln(p_T)$. Right: ln(nTracks).

applied PIDK requirements.

The descriptions of the $m(D_s^{\mp}h^{\pm})$ and $m(h^-h^+h^{\pm})$ invariant masses are taken directly from the above sample, while the parametrisation of PIDK distribution needs its own special treatment. The respective PIDK PDFs are taken from the high statistics PID calibration samples of control modes such as $D^{*\pm} \to D^0(K^{\mp}\pi^{\pm})\pi^{\pm}$ (for pure kaon and pion samples) and $\overline{\Lambda}_c^- \to \overline{p}K^+\pi^-$ (for the proton sample) [65]. Their backgrounds are subtracted using *sWeights*. In addition, the PIDK requirement from the offline selection is applied to them: PIDK < 0 for the $B_s^0 \to D_s^-\pi^+$ sample and PIDK > 5 for the $B_s^0 \to D_s^{\mp}K^{\pm}$ sample.

Then, the PID calibration samples have to be weighted to match the signal kinematics. For this purpose two variables are used: $\ln(p_T)$ and ln(nTracks). The weighting is done separately for each signal and background component, as well as for both magnet polarities. The result of the weighting is illustrated, taking as an example the $B^0 \to D^-\pi^+$ background component, in Fig. 6.6.

Next, the PDFs are combined for both magnet polarities according to the integrated luminosity in both samples. From the resulting samples the template PDFs for the PIDK distribution are constructed without any free parameter other than the normalisation.

In the following subsections, signal and background descriptions are discussed for both considered modes: $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^{\mp}K^{\pm}$. However, the resulting PDFs are shown only for $B_s^0 \to D_s^{\mp}K^{\pm}$, while corresponding results for $B_s^0 \to D_s^-\pi^+$ are presented in App. E.

6.2.1 Signal

As mentioned at the beginning of the chapter, the signal mass distributions are determined by the detector resolution and the radiative character of the B_s^0 and D_s^- meson decays. Due to low energy photons which are not reconstructed, the invariant mass distributions contain a tail extending towards lower masses, which can be described by a double Crystal Ball function for:

- the $m(D_s^{\mp}h^{\pm})$ invariant mass:

$$\begin{aligned}\mathrm{PDF}(p_{B_s^0}; m(D_s^{\mp}h^{\pm}))_{B_s^0 \to D_s^{\mp}h^{\pm}} = \\ \mathrm{PDF}_{dCB}(\alpha^1_{B_s^0}, n^1_{B_s^0}, \sigma^1_{B_s^0}, \alpha^2_{B_s^0}, n^2_{B_s^0}, \sigma^2_{B_s^0}, f_{B_s^0}, \mu_{B_s^0}; m(D_s^{\mp}h^{\pm})),\end{aligned} \tag{6.9}$$

Table 6.2: Parameters for the sum of the double Crystal Ball function describing the signal $m(D_s^{\mp}h^{\pm})$ invariant mass of $B_s^0 \to D_s^{\mp}K^{\pm}$ sample, obtained from signal simulation, separately for each D_s^- final state.

Parameter	$D_s^- \to (KK\pi)_{\text{nonres}}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$\mu_{B_s^0}$ [MeV/c^2]	5366.3 ± 0.02	5366.5 ± 0.03	5366.4 ± 0.03	5366.0 ± 0.08	5366.2 ± 0.05
$\sigma^1_{B_s^0}$ [MeV/c^2]	10.72 ± 0.12	11.23 ± 0.08	10.77 ± 0.09	11.27 ± 0.30	11.39 ± 0.10
$\sigma^2_{B_s^0}$ [MeV/c^2]	16.01 ± 0.11	17.02 ± 0.10	15.34 ± 0.08	19.41 ± 0.93	17.65 ± 0.10
$\alpha^1_{B_s^0}$	2.21	2.21	2.05	2.40	2.09
$\alpha^2_{B_s^0}$	-2.42	-2.19	-2.03	-3.42	-2.33
$n^1_{B_s^0}$	1.00 ± 0.02	1.12 ± 0.02	1.21 ± 0.02	0.98 ± 0.05	1.27 ± 0.02
$n^2_{B_s^0}$	3.15 ± 0.22	3.61 ± 0.19	6.57 ± 0.46	0.52 ± 0.29	4.02 ± 0.36
$f_{B_s^0}$	0.62 ± 0.02	0.70 ± 0.02	0.58 ± 0.13	0.78 ± 0.05	0.70 ± 0.01

Table 6.3: Parameters for the sum of the double Crystal Ball function describing the signal $m(h^-h^+h^{\pm})$ invariant mass of $B_s^0 \to D_s^{\mp}K^{\pm}$ sample, obtained from signal simulation, separately for each D_s^- final state.

Parameter	$D_s^- \to (KK\pi)_{\text{nonres}}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$\mu_{D_s^-}$ [MeV/c^2]	1968.8 ± 0.03	1968.8 ± 0.01	1968.9 ± 0.04	1968.9 ± 0.05	1968.8 ± 0.06
$\sigma^1_{D_s^-}$ [MeV/c^2]	5.35 ± 0.13	8.24 ± 0.11	6.08 ± 0.93	8.85 ± 0.23	8.09 ± 0.15
$\sigma^2_{D_s^-}$ [MeV/c^2]	5.18 ± 0.13	4.49 ± 0.04	5.13 ± 0.26	5.21 ± 0.20	7.38 ± 0.18
$\alpha^1_{D_s^-}$	1.23 ± 0.09	1.98 ± 0.05	1.13 ± 0.14	1.71 ± 0.11	0.91 ± 0.21
$\alpha^2_{D_s^-}$	-1.12 ± 0.04	-3.05 ± 0.14	-1.38 ± 0.27	-2.53 ± 0.42	-1.11 ± 0.05
$n^1_{D_s^-}$	4.66 ± 0.08	1.49 ± 0.20	13.3 ± 1.10	2.02 ± 0.47	11.5 ± 1.63
$n^2_{D_s^-}$	69.9 ± 10.1	0.61 ± 0.18	11.0 ± 1.18	1.09 ± 0.77	40.0 ± 5.32
$f_{D_s^-}$	0.48 ± 0.05	0.40 ± 0.01	0.40 ± 0.06	0.55 ± 0.05	0.49 ± 0.08

- the $m(h^-h^+h^{\pm})$ invariant mass:

$$\begin{aligned}\text{PDF}(p_{D_s^-};m(h^-h^+h^{\pm}))_{B_s^0\to D_s^{\mp}h^{\pm}} = \\ \text{PDF}_{dCB}(\alpha^1_{D_s^-}, n^1_{D_s^-}, \sigma^1_{D_s^-}, \alpha^2_{D_s^-}, n^2_{D_s^-}, \sigma^2_{D_s^-}, f_{D_s^-}, \mu_{D_s^-}; m(h^-h^+h^{\pm})).\end{aligned} \tag{6.10}$$

In order to find the best parametrisation, pure signal simulation samples are fitted for both magnet polarities but each D_s^- final state is considered separately.

Since the $\alpha^{1(2)}_{B_s^0}$ parameters cause some instability if left free a two stage procedure is used for $m(D_s^{\mp}h^{\pm})$ invariant mass description. The shape is first obtained with all parameters floating, then a second fit is performed, where the $\alpha^{1(2)}_{B_s^0}$ parameters are fixed to the best fit value from previous step. The central values from both determinations are equal, but the second has a correct covariance matrix. The results are given in Tab. 6.2. In case of the $m(h^-h^+h^{\pm})$ invariant mass all parameters are left free and the results are given in Tab. 6.3. For both B_s^0 and D_s^- descriptions in the fit to the data, only the means $\mu_{B_s^0}$ and $\mu_{D_s^-}$ are floating, while all other shape parameters are fixed to the values in the tables, where the widths are scaled by the data/simulation ratio of Tab. 6.1.

The PDFs of the PIDK variable are obtained by weighting pure kaon PID calibration sample taken from $D^{*\pm}$ decays. The resulting parametrisation for $m(D_s^{\mp}h^{\pm})$ and $m(h^-h^+h^{\pm})$ invariant masses and PIDK variable are shown in Fig. 6.7.

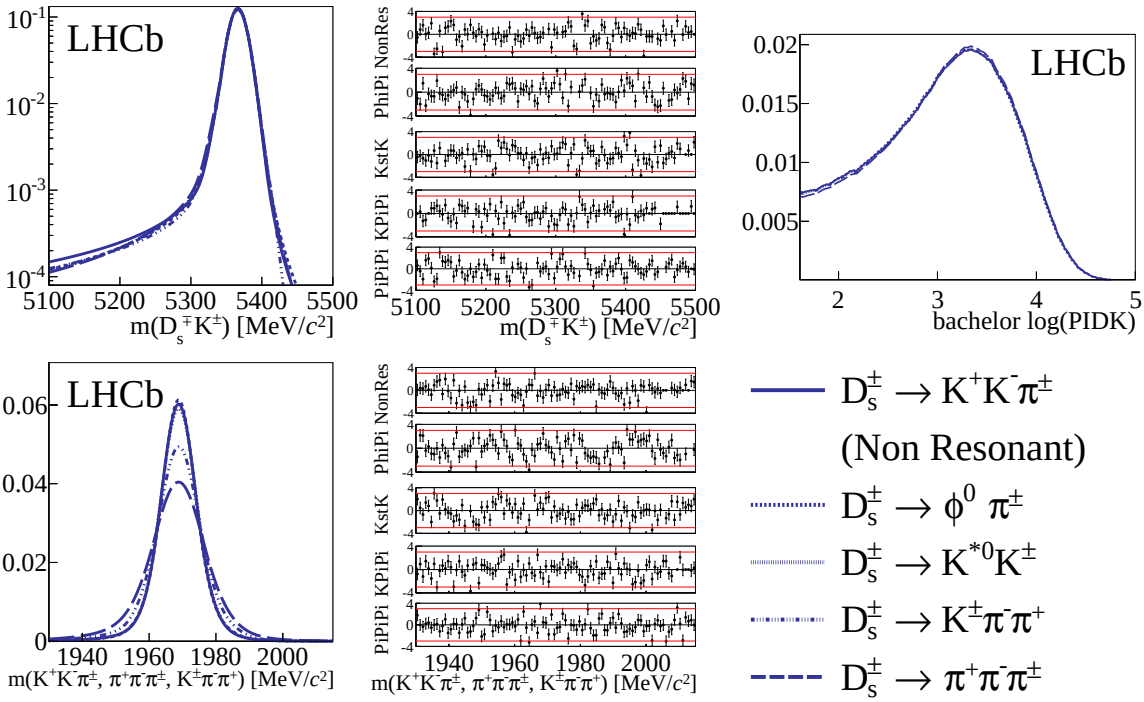

Figure 6.7: Signal PDFs of $B_s^0 \to D_s^\mp K^\pm$ sample as evaluated from simulation on the joint magnet up and down samples, separately for each D_s^- final states. Different D_s^- final states are shown as blue lines as described in the legend placed in the bottom right. Top left: the $m(D_s^\mp K^\pm)$ invariant mass together with the pull distributions, top right: the ln(PIDK) variable, bottom: the $m(h^-h^+h^\pm)$ invariant mass together with the pull distributions.

6.2.2 Combinatorial background

The functional form of the combinatorial background is taken from the $m(D_s^\mp h^\pm)$ sideband $m(D_s^\mp h^\pm) > 5800\,\text{MeV}/c^2$, but with the full range for $m(h^-h^+h^\pm)$ invariant mass: $m(h^-h^+h^\pm) = (1930,\ 2015)\ \text{MeV}/c^2$. Each D_s^- mode is considered independently; however magnet up and down polarities are combined. In order not to distort the $m(D_s^\mp h^\pm)$ invariant mass shape of the combinatorial background, the mass constraint on the D_s^- is removed for this study.

The combinatorial background in the $m(D_s^\mp h^\pm)$ invariant mass is described as follows:

- for $B_s^0 \to D_s^- \pi^+$ ($D_s^- \to (KK\pi)_{\text{nonres}}, D_s^- \to \phi\pi^-, D_s^- \to K^{*0}K^-$) by a single exponential plus flat distribution:

$$\text{PDF}(s_{B_s^0}, fC_{B_s^0}; m(D_s^\mp h^\pm))_{Comb} = fC_{B_s^0} \times e^{-s_{B_s^0} \times m(D_s^\mp h^\pm)} + (1 - fC_{B_s^0}) \times 1,$$

  where $s_{B_s^0}$ is the slope of the exponential and $fC_{B_s^0}$ denotes fraction between components.

- for $B_s^0 \to D_s^- \pi^+$ ($D_s^- \to K^-\pi^+\pi^-, D_s^- \to \pi^-\pi^+\pi^-$) and $B_s^0 \to D_s^\mp K^\pm$ (all D_s^- modes) by a single exponential:

$$\text{PDF}(s_{B_s^0}; m(D_s^\mp h^\pm))_{Comb} = e^{-s_{B_s^0} \times m(D_s^\mp h^\pm)},$$

  where $s_{B_s^0}$ is the slope of the exponential.

The results for $B_s^0 \to D_s^\mp K^\pm$ are shown in the top left part of Fig. 6.8. The functional form of the applied parametrisation is sufficiently complex to be able to describe the data. In order to cover any difference between shapes in the sideband $m(D_s^\mp h^\pm) > 5800\,\mathrm{MeV}/c^2$ and in the nominal fit region: $m(D_s^\mp h^\pm) \in (5300, 5800)\,\mathrm{MeV}/c^2$, the slopes of the exponential are left free.

The combinatorial background in the $m(h^-h^+h^\pm)$ invariant mass dimension absorbs every component which does not peak in the $m(D_s^\mp h^\pm)$ invariant mass. This means true combinatorial, in which all tracks are random tracks, and, in some D_s^- channels, backgrounds that contain a true D_s^-, but a random bachelor track. The combinatorial background in the range $m(h^-h^+h^\pm) = (1930, 2015)\,\mathrm{MeV}/c^2$ can be described by:

- an exponential plus signal D_s^- mass shape for the $D_s^- \to (KK\pi)_{\mathrm{nonres}}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$ D_s^- final states:

$$\begin{aligned}\mathrm{PDF}(p_{D_s^-}, s_{D_s^-}, fC_{D_s^-}; m(h^-h^+h^\pm))_{Comb} = fC_{D_s^-} \times e^{-s_{D_s^-} \times m(h^-h^+h^\pm)} \\ + (1 - fC_{D_s^-}) \times \mathrm{PDF}(p_{D_s^-}; m_{D_s^-})_{B_s^0 \to D_s^\mp K^\pm},\end{aligned}$$

  where $s_{D_s^-}$ is the slope of the exponential, $fC_{D_s^-}$ corresponds to fraction between components and $p_{D_s^-}$ are shared with signal description parameters of double Crystal Ball.

- a single exponential for $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$:

$$\mathrm{PDF}(s_{D_s^-}; m(h^-h^+h^\pm)) = e^{-s_{D_s^-} \times m(h^-h^+h^\pm)},$$

  where $s_{D_s^-}$ is the slope of the exponential.

The results for $B_s^0 \to D_s^\mp K^\pm$ are shown in the bottom of Fig. 6.8. To cover differences between the sideband and nominal $m(D_s^\mp h^\pm)$ mass regions, the slope parameters and also, for those channels that have it, the fraction between peaking and exponential components are left free. Since the shape of the peaking component coincides with the signal shape, its mean is floating, too, as described in the previous Section 6.2.1.

The combinatorial background contains contributions from many particle species, namely kaons, pions, and protons. Each of these components has their own, distinctive PIDK shape. In the nominal fit each of these parametrisation is considered separately. However they share with the same $m(D_s^\mp h^\pm)$ and $m(h^-h^+h^\pm)$ invariant masses models.

The PDFs of the PIDK variable are obtained based on the $D^{*\pm}$ and $\overline{\Lambda}_c^-$ calibration samples, which are weighted to match the $\ln(p_T)$ and ln(nTracks) spectra of the combinatorial background in the fit range ($m(D_s^\mp h^\pm) = [5300, 5800]\,\mathrm{MeV}/c^2$). Since the pure combinatorial background sample can only be obtained from the sideband of $m(D_s^\mp h^\pm)$, its kinematics must be extrapolated into the fit region. The upper $m(D_s^\mp h^\pm)$ invariant mass sideband, $m(D_s^\mp h^\pm) = [5600, 6800]\,\mathrm{MeV}/c^2$, is partitioned into 10 bins. In each of these bins the transverse momentum p_T and track multiplicity nTracks spectra are fitted by a Landau function [99]. The output of the fit is used to predict the evolution of the Landau's mean and width in bins inside the fit region. Finally summing all bins in the fit region with appropriate weights gives the wanted $\ln(p_T)$ and ln(nTracks) spectra in the fit region.

The combinatorial PDF for the $B_s^0 \to D_s^-\pi^+$ decay contain only pion and kaon contributions:

$$\mathrm{PDF}(fC_{\mathrm{PIDK}}; \mathrm{PIDK})_{\mathrm{Comb}} = \mathrm{fC_{PIDK}} \times \mathrm{PDF(PIDK)}_\pi + (1 - \mathrm{fC_{PIDK}}) \times \mathrm{PDF(PIDK)_K},$$

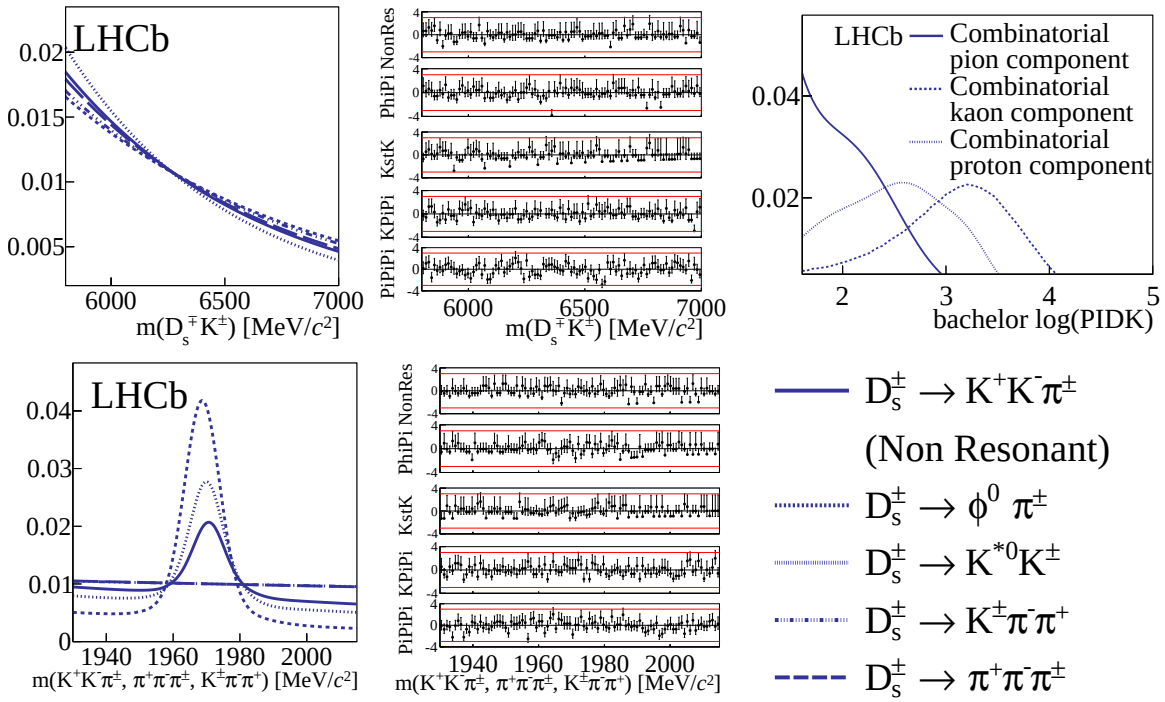

Figure 6.8: Combinatorial background PDFs of $B_s^0 \to D_s^{\mp} K^{\pm}$ sample as evaluated from the $m(D_s^{\mp} h^{\pm})$ sideband on the joint magnet up and down samples, separately for each D_s^- final state. For the $m(D_s^{\mp} h^{\pm})$ and $m(h^- h^+ h^{\pm})$ invariant masses, different D_s^- final states are shown as blue lines as described in the legend placed on the plots. For the PIDK variable pion, kaon and proton components are described in the legend placed in the figure. Top left: the $m(D_s^{\mp} K^{\pm})$ invariant mass together with the resulting pull distributions, top right: the ln(PIDK) variable, bottom: the $m(h^- h^+ h^{\pm})$ invariant mass together with the resulting pull distributions.

where $fC_{\rm PIDK}$ is the fraction between components, whereas PDF(PIDK)$_\pi$ and PDF(PIDK)$_{\rm K}$ denote the pion and kaon components, respectively. For the $B_s^0 \to D_s^{\mp} K^{\pm}$ in addition to the two above contributions, the proton component has to be considered:

$$\begin{aligned}\text{PDF}(fC_{\rm PIDK1}, fC_{\rm PIDK2};\text{PIDK})_{\rm Comb} = \text{fC}_{\rm PIDK} \times \text{PDF(PIDK)}_\pi + (1 - \text{fC}_{\rm PIDK1}) \times \\ [fC_{\rm PIDK2} \times \text{PDF(PIDK)}_{\rm K} + (1 - \text{fC}_{\rm PIDK2}) \times \text{PDF(PIDK)}_{\rm p}],\end{aligned}$$

where $fC_{\rm PIDK1}, fC_{\rm PIDK2}$ are fractions, and PDF(PIDK)$_\pi$, PDF(PIDK)$_{\rm K}$ and PDF(PIDK)$_{\rm p}$ denotes the pion, kaon and proton components, respectively. The resulting PDFs for the $B_s^0 \to D_s^{\mp} K^{\pm}$ combinatorial background are shown in Fig. 6.8. The normalisation ($fC_{\rm PIDK}$ in $B_s^0 \to D_s^- \pi^+$ sample and $fC_{\rm PIDK1}, fC_{\rm PIDK2}$ in $B_s^0 \to D_s^{\mp} K^{\pm}$ sample) is left floating in the fit.

6.2.3 Fully Reconstructed Backgrounds

The fully reconstructed backgrounds are the dominant source of specific background components in the $B_s^0 \to D_s^{\mp} h^{\pm}$ samples.

The contributions to the $B_s^0 \to D_s^- \pi^+$ sample are due to: $B^0 \to D_s^- \pi^+$, $B_s^0 \to D_s^{\mp} K^{\pm}$, $B^0 \to D^- \pi^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ decays. Fully reconstructed backgrounds considered in the $B_s^0 \to D_s^{\mp} K^{\pm}$ sample are due to: $B^0 \to D_s^- K^+$, $B_s^0 \to D_s^- \pi^+$, $\Lambda_b^0 \to D_s^- p$, $B^0 \to D^- K^+$, $B^0 \to D^- \pi^+$, $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- K^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ decays. The detailed informations about

Table 6.4: Fully reconstructed backgrounds considered in the $B_s^0 \to D_s^- \pi^+$ sample.

Decay chain	Background to	How it contributes to the signal
$B^0 \to D_s^- \pi^+$	$B_s^0 \to D_s^- \pi^+$	Same final state
$B_s^0 \to D_s^- K^+$	$B_s^0 \to D_s^- \pi^+$	$K^+ \to \pi^+$ (bachelor)
$B^0 \to D^- \pi^+ \to (K^+\pi^-\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (K^+K^-\pi^-)\pi^+$	$\pi^- \to K^-$ (D^-)
$B^0 \to D^- \pi^+ \to (K^+\pi^-\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (\pi^+K^-\pi^-)\pi^+$	$K^+ \to \pi^+$ and $\pi^- \to K^-$ (D^-)
$\overline{\Lambda_b^0} \to \overline{\Lambda_c^-} \pi^+ \to (\overline{p}K^+\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (K^+K^-\pi^-)\pi^+$	$\overline{p} \to K^-$ (D^-)
$\overline{\Lambda_b^0} \to \overline{\Lambda_c^-} \pi^+ \to (\overline{p}K^+\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (\pi^+K^-\pi^-)\pi^+$	$\overline{p} \to K^-$ ($\overline{\Lambda_c^-}$) and $K^+ \to \pi^+$ ($\overline{\Lambda_c^-}$)

Table 6.5: Fully reconstructed backgrounds considered in the $B_s^0 \to D_s^\mp K^\pm$ sample.

Decay chain	Background to	How it contributes to the signal
$B^0 \to D_s^- K^+$	$B_s^0 \to D_s^- K^+$	Same final state
$B_s^0 \to D_s^- \pi^+$	$B_s^0 \to D_s^- K^+$	$\pi^+ \to K^+$ (bachelor)
$\Lambda_b^0 \to D_s^- p$	$B_s^0 \to D_s^- K^+$	$p \to K^+$ (bachelor)
$B^0 \to D^- K^+ \to (K^+\pi^-\pi^-)K^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\pi^- \to K^-$ (D^-)
$B^0 \to D^- \pi^+ \to (K^+\pi^-\pi^-)\pi^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\pi^- \to K^-$ (D^-) and $\pi^+ \to K^+$ (bachelor)
$\overline{\Lambda_b^0} \to \overline{\Lambda_c^-} K^+ \to (\overline{p}K^+\pi^-)K^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\overline{p} \to K^-$ ($\overline{\Lambda_c^-}$)
$\overline{\Lambda_b^0} \to \overline{\Lambda_c^-} \pi^+ \to (\overline{p}K^+\pi^-)\pi^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\overline{p} \to K^-$ ($\overline{\Lambda_c^-}$) and $\pi^+ \to K^+$ (bachelor)

misidentification types are collected in Tab. 6.4 for the $B_s^0 \to D_s^- \pi^+$ decay and Tab. 6.5 for the $B_s^0 \to D_s^\mp K^\pm$ mode.

The most important of these backgrounds are $B^0 \to D^- \pi^+$ (as background to $B_s^0 \to D_s^- \pi^+$) and $B_s^0 \to D_s^- \pi^+$ (as background to $B_s^0 \to D_s^\mp K^\pm$). Their shapes are determined in a data driven way, as illustrated in the following for the $B_s^0 \to D_s^- \pi^+$:

1. A pure sample of $B_s^0 \to D_s^- \pi^+$ is reconstructed under the correct mass hypothesis for the misidentified particle (i.e. π hypothesis), applying the full selection chain with a signal mass window $m(D_s^- \pi^+) = (5320, 5420)\,\mathrm{MeV}/c^2$ plus a PID requirement to explicitly select this decay: PIDK < 0 on the bachelor pion. Since the mis-ID and efficiency of the PIDK cuts depend on momentum, these PIDK requirements will distort the momentum distribution of the misidentified particle.

2. The efficiency and mis-ID probability of the PIDK cuts applied in the first step, as a function of momentum, are obtained from the PID calibration sample of $D^{*\pm}$ decays. To recover the pre-PID cut momentum distribution the pure sample is weighted using calibration functions.

3. The pure $B_s^0 \to D_s^- \pi^+$ sample is reconstructed as $B_s^0 \to D_s^\mp K^\pm$, which means under the wrong mass hypothesis for the misidentified particle, i.e. the bachelor π^+ now has the K^+ mass. This changes the mass distribution.

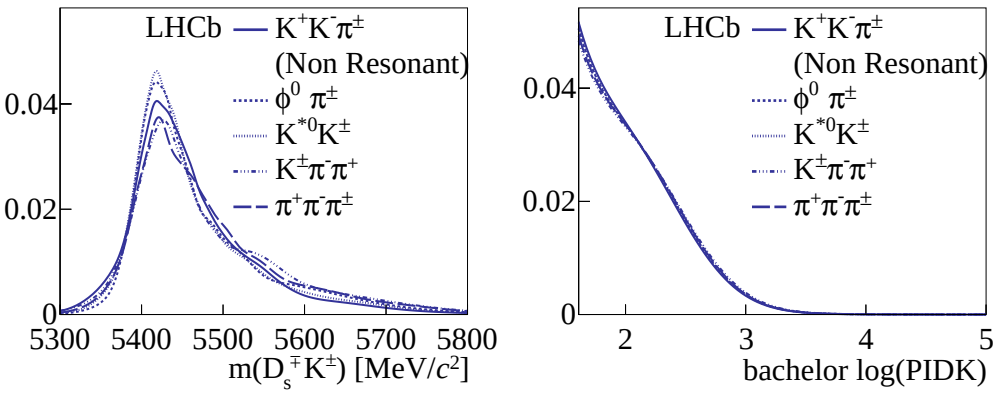

Figure 6.9: The PDFs taken from data for $B_s^0 \to D_s^- \pi^+$ samples, separately for each D_s^- final state, but with combined magnet polarities. Different D_s^- final states are shown as blue lines as described in the legend in the plots. Left: $m(D_s^\mp h^\pm)$ invariant mass, Right: ln(PIDK) variable.

4. The PIDK cuts in the nominal $B_s^0 \to D_s^\mp K^\pm$ selection also modify the momentum distribution. To consider their effect, the resulting sample is again weighted by using the calibration functions, corresponding to the $B_s^0 \to D_s^\mp K^\pm$ PID selection (PIDK > 5).

The PDFs of $m(D_s^\mp h^\pm)$ invariant mass are taken directly from this sample, while for the $m(h^-h^+h^\pm)$ invariant mass the signal $B_s^0 \to D_s^\mp K^\pm$ PDFs (see Fig. 6.7) are used. Finally, the PIDK distribution is obtained based on weighting the pion PID calibration sample to match the kinematics of each $B_s^0 \to D_s^- \pi^+$ sample. The resulting PDFs for the $m(D_s^\mp h^\pm)$ invariant mass and the PIDK are shown in Fig. 6.9.

This procedure is also applied to the $B^0 \to D^- \pi^+$, $D^- \to K^+\pi^-\pi^-$ sample, and the resulting shape is used to parametrise the $B^0 \to D^- \pi^+$ background in the $B_s^0 \to D_s^- \pi^+$ fit in both final states $D_s^- \to K^-K^+\pi^-$ and $D_s^- \to K^-\pi^+\pi^-$. The PDFs for $m(D_s^\mp h^\pm)$ and $m(h^-h^+h^\pm)$ invariant masses are created directly based on above samples. The PIDK distributions are obtained from the pion PID calibration sample to match $B^0 \to D^- \pi^+$ kinematics, as shown in Fig.6.6.

The $m(D_s^\mp h^\pm)$ invariant mass PDFs for the backgrounds with the same final state as the signal, $B^0 \to D_s^- K^+$ and $B^0 \to D_s^- \pi^+$, are assumed to be described by double Crystal Ball functions. Their tail parameters are taken from simulated candidates and fixed in the fit, whereas the mean is taken from $B_s^0 \to D_s^\mp h^\pm$ signal description, and shifted downward by the nominal difference between the B_s^0 and B^0 masses, $\Delta = 86.6\,\text{MeV}/c^2$ [63]. In addition, the widths are obtained from the $B_s^0 \to D_s^\mp h^\pm$ signal shapes and scaled by the width ratio $B_s^0 \to D_s^\mp h^\pm / B^0 \to D_s^\mp h^\pm$ as measured in simulated candidates. Since the final state $D_s^\mp h^\pm$ is the same as signal, the signal $B_s^0 \to D_s^\mp h^\pm$ PDFs are used for the $m(h^-h^+h^\pm)$ invariant mass and PIDK distributions.

The $m(D_s^\mp h^\pm)$ invariant mass PDFs of the remaining fully reconstructed backgrounds are taken from simulation. The fully reconstructed backgrounds which contain a correctly reconstructed D_s^- meson behave just like the signal in the $m(h^-h^+h^\pm)$ invariant mass distribution, so the signal PDFs (see Fig. 6.7) are used. The remaining distributions ($\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ in the $B_s^0 \to D_s^- \pi^+$ fit, and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- h^+$ and $B^0 \to D^- h^-$ in the $B_s^0 \to D_s^\mp K^\pm$ fit) are obtained from selected and corrected simulation samples. The PIDK PDFs for the fully reconstructed backgrounds are obtained using the procedure discussed

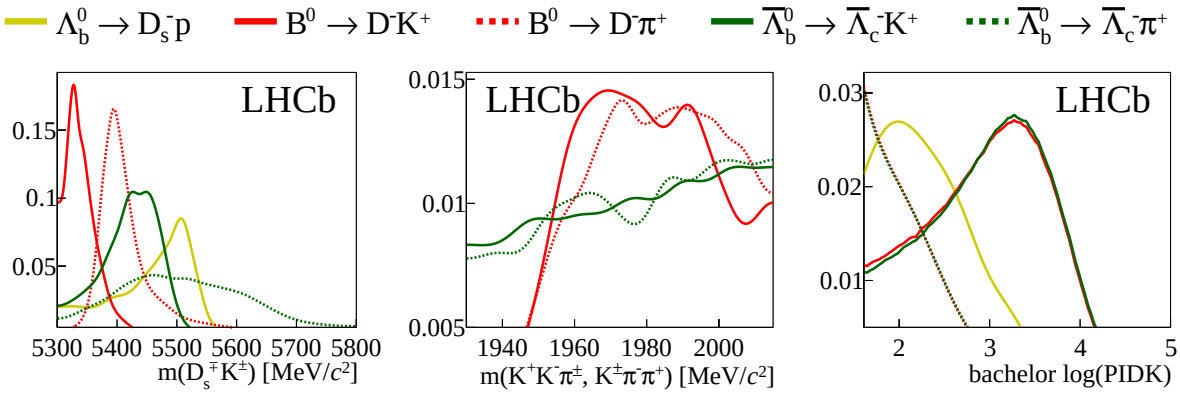

Figure 6.10: The PDFs taken from simulation for the remaining fully reconstructed backgrounds in the $B_s^0 \to D_s^\mp K^\pm$ sample. Different background contributions are shown as coloured lines as described in the legend above the plots. Left: $m(D_s^\mp h^\pm)$ invariant mass, middle: $m(h^-h^+h^\pm)$ invariant mass, right: ln(PIDK) variable.

in the beginning of this section (Sec. 6.2). They are taken from the PID calibration samples, which got weighted to match the kinematics found in the relevant simulation samples. The contributing backgrounds are described using the pion, kaon or proton PID calibration sample depending on the bachelor particle. The resulting PDFs of remaining fully reconstructed backgrounds are shown in Fig. 6.10.

6.2.4 Partially Reconstructed Backgrounds

The partially reconstructed backgrounds typically peak in the lower mass range, but do extend into the signal region. The backgrounds included in the nominal fit are listed in Tab. 6.6. The $B_s^0 \to D_s^{*-}\pi^+$ background contributes to both $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ samples due to the missing photon from the $D_s^{*-} \to D_s^-\gamma$ decay. In addition, two more decays have to be taken into account as backgrounds to $B_s^0 \to D_s^\mp K^\pm$. The first one is $B_s^0 \to D_s^-\rho^+$, when the π^0 is missed from the $\rho^+ \to \pi^+\pi^0$ decay and pion misidentification occurs ($\pi^+ \to K^+$). The second component is $\Lambda_b^0 \to D_s^{*-}p$ when the photon from $D_s^{*-} \to D_s^-\gamma$ is missed and in addition the proton is misidentified as the kaon ($p \to K^+$).

The $m(D_s^\mp h^\pm)$ invariant mass PDFs of the partially reconstructed backgrounds are taken from simulation. For the $m(h^-h^+h^\pm)$ invariant mass distributions the $B_s^0 \to D_s^-\pi^+$ or $B_s^0 \to D_s^\mp K^\pm$ signal PDFs are used. Finally, the PIDK PDFs are obtained based on the PID calibration samples and depending on nature of bachelor particle: either from the pion or proton PID calibration sample. The resulting PDFs of partially reconstructed backgrounds are shown in Fig. 6.11.

Table 6.6: Partially reconstructed backgrounds considered in the $B_s^0 \to D_s^\mp h^\pm$ samples.

Decay chain	Background to	How it contributes to the signal
$B_s^0 \to D_s^{*-}\pi^+$	$B_s^0 \to D_s^-\pi^+$	missed photon from $D_s^{*-} \to D_s^-\gamma$
$B_s^0 \to D_s^-\rho^+$	$B_s^0 \to D_s^-K^+$	missed π^0 from $\rho^+ \to \pi^+\pi^0$ and $\pi^+ \to K^+$ (bachelor)
$B_s^0 \to D_s^{*-}\pi^+$	$B_s^0 \to D_s^-K^+$	missed photon from $D_s^{*-} \to D_s^-\gamma$ and $\pi^+ \to K^+$ (bachelor)
$\Lambda_b^0 \to D_s^{*-}p$	$B_s^0 \to D_s^-K^+$	missed photon from $D_s^{*-} \to D_s^-\gamma$ and $p \to K^+$ (bachelor)

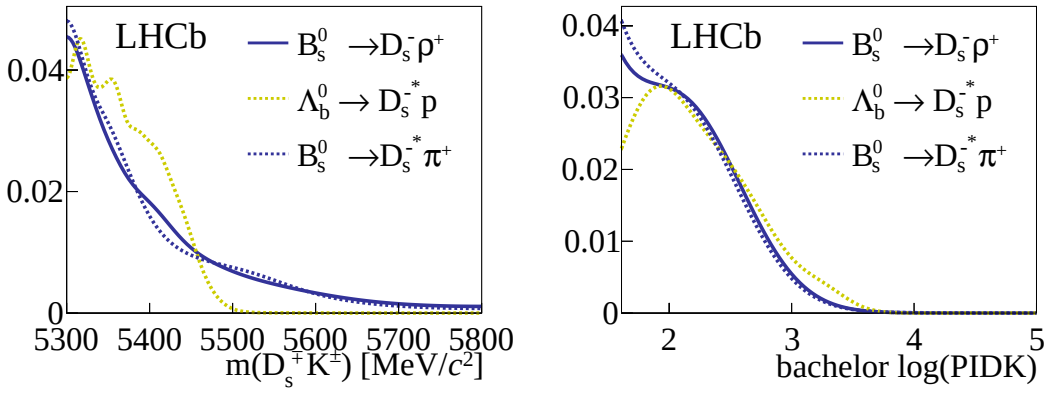

Figure 6.11: The PDFs taken from simulation for the partially reconstructed backgrounds in the $B_s^0 \to D_s^\mp K^\pm$ sample. Different background contributions are shown as coloured lines as described in the legend in the plots. Left: $m(D_s^\mp h^\pm)$ invariant mass, Right: ln(PIDK) variable.

6.3 Expected background yields

The backgrounds whose yields are below 2% of the signal yield are fixed from known branching fractions and relative efficiencies measured using simulated candidates.

6.3.1 Background due to $B^0 \to D^- \pi^+$ and $B^0 \to D^- K^+$ decays

The number of $B^0 \to D^- \pi^+$ background candidates remaining in the $B_s^0 \to D_s^- \pi^+$ samples can be estimated from data, using the sample described in Sec. 6.1.2. To obtain the expected yield of $B^0 \to D^- \pi^+$ without applying a PID requirement, the fitted $B^0 \to D^- \pi^+$ yield ($N_{B^0 \to D^- \pi^+} = 109420$) is corrected due to the different selections applied to the $B^0 \to D^- \pi^+$ and $B_s^0 \to D_s^- \pi^+$ samples. The correction is computed separately for each D_s^- final state: $D_s^- \to (KK\pi)_{\text{nonres}}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0} K^-$ and $D_s^- \to K^- \pi^+ \pi^-$. It contains the following contributions: efficiencies due to the choice of the signal mass window, the vetoes applied in the D^-/D_s^- child phase space, the different PID efficiencies of the D^-/D_s^- children (partly depending on the phase space position due to different PID requirements in the veto regions) as well as efficiencies of the cut on FDχ^2. In addition, the misidentification rate is needed, which is calculated by weighting the momentum spectrum of the $B^0 \to D^- \pi^+$ simulated samples by the misidentification of relevant PIDK requirements. Then, using the above information the expected $B^0 \to D^- \pi^+$ yield in the $B_s^0 \to D_s^- \pi^+$ are calculated and listed in the first row of Tab. 6.8.

These yields are also exploited to fix the $B^0 \to D^- \pi^+$ and $B^0 \to D^- K^+$ contributions in the $B_s^0 \to D_s^\mp K^\pm$ fit, using a factor 1/15 to account for the relative branching fractions of the $B^0 \to D^- K^+$ and $B^0 \to D^- \pi^+$ modes and correcting for the PID efficiencies of the PIDK < 0 and PIDK > 5 requirements on the bachelor track. The resulting contributions from $B^0 \to D^- \pi^+$ and $B^0 \to D^- K^+$ expected in the $B_s^0 \to D_s^\mp K^\pm$ fit is given in Tab. 6.11.

6.3.2 Background due to $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- K^+$ decays

For the evaluation of the number of expected $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ candidates in the $B_s^0 \to D_s^- \pi^+$ sample, the latter is reconstructed under the $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ hypothesis and fitted. The

selection for the reconstruction follows nominal $B_s^0 \to D_s^- \pi^+$ selection, however the $\overline{\Lambda}_c^-$ veto is removed and the D_s^- combination is required to be within 30 MeV/c^2 of the $m(\overline{\Lambda}_c^- \pi^+)$ invariant mass. In addition, the kaon track from the D_s^- combination satisfies the proton PID selection criteria PIDp − PIDK > 5 and PID$p > 0$. The relative misidentifications between these and the nominal selections are obtained.

In the fits to the $m(\overline{\Lambda}_c^- \pi^+)$ invariant mass only two contributions are considered: the $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ component, parametrised by the $B_s^0 \to D_s^- \pi^+$ signal double Crystal Ball functions and combinatorial background, described by a single exponential. To compute the number of expected $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ candidates in the $B_s^0 \to D_s^- \pi^+$ sample the fitted yield, shown in Fig. 6.12 is multiplied by the relative misidentifications. The resulting yields are listed in second row of Tab. 6.8.

These yields are also exploited to fix the $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- K^+$ contributions in the $B_s^0 \to D_s^\mp K^\pm$ fit, using again the factor 1/15 of the relative branching fractions of the $B^0 \to D^- K^+$ and $B^0 \to D^- \pi^+$ modes, assuming that the ratio of $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- K^+$ to $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ is the same, and correcting for the PID efficiencies of the PIDK < 0 and PIDK > 5 requirements on the bachelor track. The obtained results are collected in Tab. 6.11.

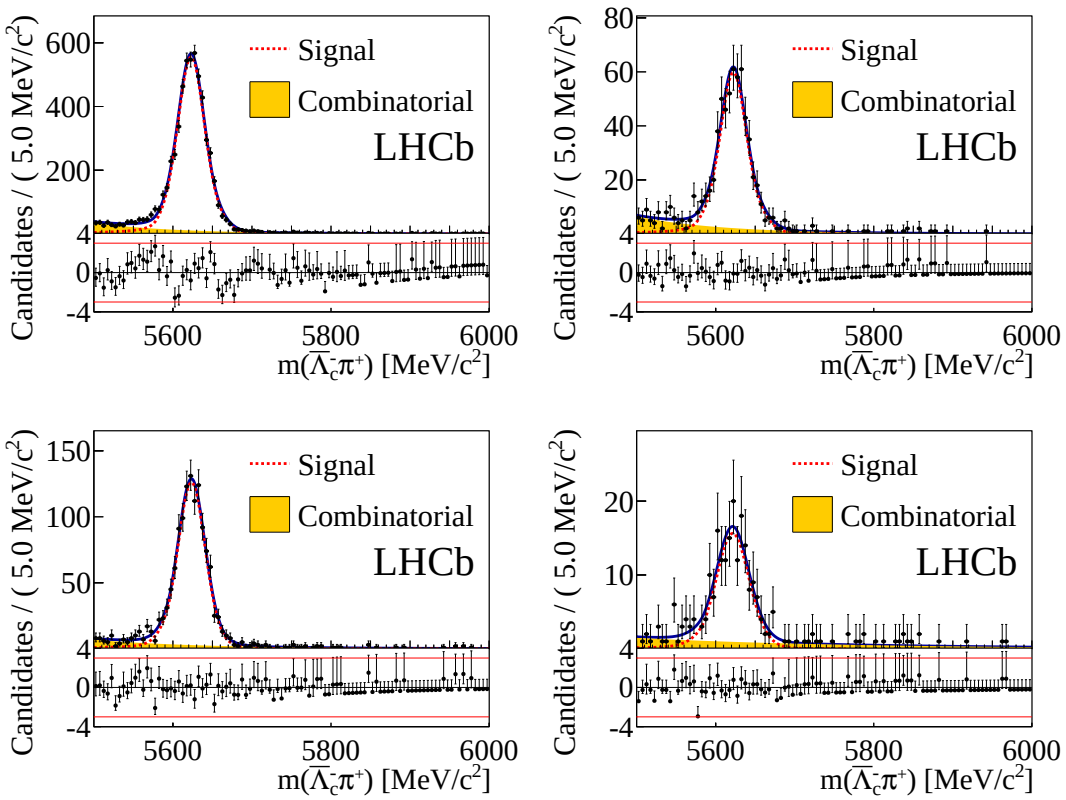

Figure 6.12: Result of the fit to the $B_s^0 \to D_s^- \pi^+$ sample, reconstructed under the $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ hypothesis. Distributions of $m(\overline{\Lambda}_c^- \pi^+)$ invariant mass for D_s^- final states with combined magnet polarities, from top left to bottom right: $D_s^- \to (KK\pi)_{\text{nonres}}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0} K^-$, $D_s^- \to K^-\pi^+\pi^-$. The combinatorial background component is shown as an orange area, while the signal is presented as a dashed red line.

6.3.3 Background due to $B_s^0 \to D_s^\mp K^\pm$ decays

To obtain the expected $B_s^0 \to D_s^\mp K^\pm$ yield in the $B_s^0 \to D_s^- \pi^+$ fit without applying any PID requirement on the bachelor, the $B^0 \to D^- \pi^+$ yields are corrected by relative branching and fragmentation fractions. The relative PID efficiency of the PIDK < 0 and PIDK > 5 cuts on the bachelor track as well as the misidentification of the PIDK > 5 cut are taken into account. The expected number of $B_s^0 \to D_s^\mp K^\pm$ background candidates in the $B_s^0 \to D_s^- \pi^+$ sample are given in Tab. 6.8.

6.4 Multidimensional fit to the $B_s^0 \to D_s^- \pi^+$ sample

The $B_s^0 \to D_s^- \pi^+$ MD fit uses the $m(D_s^\mp h^\pm)$ invariant mass, $m(h^- h^+ h^\pm)$ invariant mass, and the PIDK variable in a three-dimensional unbinned extended likelihood fit. It builds on the PDF shapes described in the previous sections and fits five subsamples, defined as D_s^- final states, simultaneously. The backgrounds are arranged into five groups collected in Tab. 6.7.

Table 6.7: Definition of background groups in the $B_s^0 \to D_s^- \pi^+$ MD fit.

Group	Decay channels
1	$B_s^0 \to D_s^{*-} \pi^+$, $B^0 \to D_s^- \pi^+$
2	$B^0 \to D^- \pi^+$
3	$\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$
4	$B_s^0 \to D_s^\mp K^\pm$
5	Combinatorial background

Table 6.8: Fixed yields in the $B_s^0 \to D_s^- \pi^+$ MD fit.

Background	$D_s^- \to (KK\pi)_{\text{nonres}}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0} K^-$	$D_s^- \to K^- \pi^+ \pi^-$	$D_s^- \to \pi^- \pi^+ \pi^-$
$B^0 \to D^- \pi^+$	374.0	6.0	93.0	30.0	0.0
$\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$	290.0	36.0	69.0	1.0	0.0
$B_s^0 \to D_s^\mp K^\pm$	40.0	47.0	40.0	8.0	21.0

Group 1 contains partially reconstructed backgrounds. From these backgrounds only $B_s^0 \to D_s^{*-} \pi^+$ falls into the nominal B_s^0 mass window $m(D_s^\mp h^\pm) = [5300, 5800]\,\text{MeV}/c^2$. In addition, the small, fully reconstructed $B^0 \to D_s^- \pi^+$ background is considered. For this group the contributions are implemented with fractions:

$$\begin{aligned} \text{PDF}(f_1; m(D_s^\mp h^\pm), m(h^- h^+ h^\pm), \text{PIDK})_{\text{Group1}} = \\ f_1 \times \text{PDF}(m(D_s^\mp h^\pm), m(h^- h^+ h^\pm), \text{PIDK})_{\text{B}_\text{s}^0 \to \text{D}_\text{s}^{*-} \pi^+} \\ + (1 - f_1) \times \text{PDF}(m(D_s^\mp h^\pm), m(h^- h^+ h^\pm), \text{PIDK})_{\text{B}^0 \to \text{D}_\text{s}^- \pi^+}, \end{aligned} \tag{6.11}$$

which means that f_1 of the group 1 background yield is due to $B_s^0 \to D_s^{*-} \pi^+$, and the rest due to $B^0 \to D_s^- \pi^+$. This fraction floats in the fit, shared among the five data samples. **Group 2, 3 and 4** refers to the fully reconstructed, but misidentified backgrounds due to: $B^0 \to D^- \pi^+$, $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ decays, respectively, whose yields are fixed from the predictions presented in Tab. 6.8. **Group 5** comprises the combinatorial background discussed in detail in Sec. 6.2.2

For each D_s^- final state, the total PDF is composed of:

$$\mathrm{PDF}_{Tot}(a;m(D_s^{\mp}h^{\pm}),m(h^-h^+h^{\pm}),\mathrm{PIDK}) = \sum_j N_j \times \mathrm{PDF}(a_j;m(D_s^{\mp}h^{\pm}),m(h^-h^+h^{\pm}),\mathrm{PIDK})_{\mathrm{j}}, \tag{6.12}$$

where $j = \{1,..,6\}$ denotes the signal and five groups of backgrounds, N_j are yields and a_j are additional floating parameters: $a_{(B_s^0\to D_s^-\pi^+)} = (\mu_{B_s^0}, \mu_{D_s^-})$ for the signal, $a_{(Comb)} = (s_{B_s^0}, s_{D_s^-}, fC_{B_s^0}, fC_{D_s^-}, fC_{\mathrm{PIDK}})$ for the combinatorial background and $a_{(Group1)} = f_1$ for Group 1 background. In total 15 yields and 20 other parameters are left free in the fit.

The fit results to the $B_s^0 \to D_s^-\pi^+$ mass distribution, separately for each considered sample are shown in App. F. The combined result for all subsamples are presented Fig. 6.13. Finally, the fitted parameters are listed in Tab. 6.9.

Table 6.9: Fitted values of the parameters for the MD fit to $B_s^0 \to D_s^-\pi^+$ sample. The N_i are the yields of the signal and background contributions. The means $\mu_{B_s^0}$ and $\mu_{D_s^-}$ are the parameters of the double Crystal Ball used to describe the signal in the $m(D_s^{\mp}h^{\pm})$ and $m(h^-h^+h^{\pm})$ invariant masses, respectively. The parameter f_1 is the fraction between the modes in the group 1: $B_s^0 \to D_s^{*-}\pi^+$, $B^0 \to D_s^-\pi^+$. The $s_{B_s^0(D_s^-)}$ parameters are the slopes of combinatorial background in 10^{-3} units. The fractions $fC_{B_s^0}$ are the fractions between the exponential and flat distributions in the $m(D_s^{\mp}h^{\pm})$ invariant mass, whereas $fC_{D_s^-}$ are the fractions between the exponential and signal shapes in the $m(h^-h^+h^{\pm})$ invariant mass. The fC_{PIDK} is the fraction between the pion and kaon components in the PIDK combinatorial shape, shared among all subsamples.

Parameter	$D_s^- \to (KK\pi)_{\mathrm{nonres}}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$N_{B_s^0\to D_s^-\pi^+}$	4667 ± 76	10284 ± 108	7304 ± 91	1671 ± 44	4338 ± 73
N_{Group1}	88 ± 19	112 ± 22	89 ± 18	31 ± 11	54 ± 16
N_{Comb}	2587 ± 67	1153 ± 57	1173 ± 51	1096 ± 38	3021 ± 65
$N_{B_s^0\to D_s^-\pi^+}$	total yield: 28264 ± 182				
$s_{B_s^0}$	-6.20 ± 0.66	-8.75 ± 1.19	-8.41 ± 0.91	-1.98 ± 0.25	-2.19 ± 0.16
$s_{D_s^-}$	-4.53 ± 0.93	-2.75 ± 1.77	-4.81 ± 1.50	-4.24 ± 1.29	-2.58 ± 0.78
$fC_{B_s^0}$	0.78 ± 0.06	0.66 ± 0.05	0.77 ± 0.05	not fitted	not fitted
$fC_{D_s^-}$	0.97 ± 0.02	0.62 ± 0.03	0.84 ± 0.03	not fitted	not fitted
$\mu_{B_s^0}$ [MeV/c^2]	5369.9 ± 0.1				
$\mu_{D_s^-}$ [MeV/c^2]	1969.8 ± 0.1				
f_1	0.807 ± 0.787				
fC_{PIDK}	0.913 ± 0.009				

6.5 Multidimensional fit to the $B_s^0 \to D_s^{\mp}K^{\pm}$ sample

The fit to the $B_s^0 \to D_s^{\mp}K^{\pm}$ distribution is constructed analogously to the $B_s^0 \to D_s^-\pi^+$ one, with five D_s^- final states fitted simultaneously. The backgrounds are arranged into five groups, defined in Tab. 6.10.

Group 1 encompasses the decays to the same final state: $B^0 \to D_s^-K^+$. **Group 2** comprises B_s^0 decays with a bachelor pion: $B_s^0 \to D_s^-\pi^+$, $B_s^0 \to D_s^{*-}\pi^+$ and $B_s^0 \to D_s^-\rho^+$.

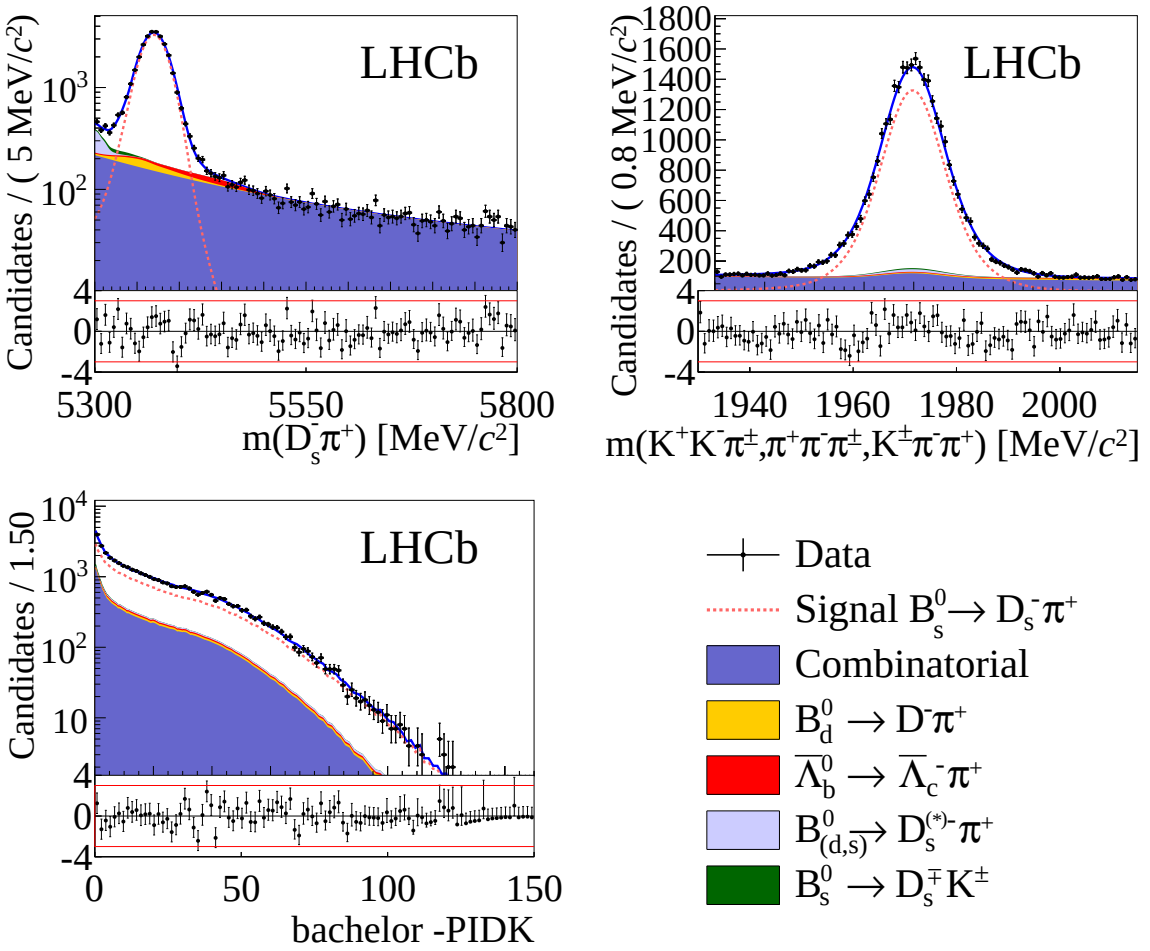

Figure 6.13: The simultaneous fit to the $B_s^0 \to D_s^-\pi^+$ candidates for both magnet polarities and all D_s^- final states combined, from top left to bottom: $m(D_s^\mp h^\pm)$ invariant mass, $m(h^-h^+h^\pm)$ invariant mass, $-$PIDK variable. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend placed in the bottom right.

For this group the various contributions are implemented with recursive fractions:

$$\begin{aligned}\mathrm{PDF}(f_{21};&m(D_s^\mp h^\pm), m(h^-h^+h^\pm), \mathrm{PIDK})_{\mathrm{Group2}} = \\ &f_{21} \times \mathrm{PDF}(m(D_s^\mp h^\pm), m(h^-h^+h^\pm), \mathrm{PIDK})_{\mathrm{B_s^0 \to D_s^-\pi^+}} \\ &+ (1-f_{21}) \times [f_{22} \times \mathrm{PDF}(m(D_s^\mp h^\pm), m(h^-h^+h^\pm), \mathrm{PIDK})_{\mathrm{B_s^0 \to D_s^{*-}\pi^+}} \\ &+ (1-f_{22}) \times \mathrm{PDF}(m(D_s^\mp h^\pm), m(h^-h^+h^\pm), \mathrm{PIDK})_{\mathrm{B_s^0 \to D_s^-\rho^+}}]. \end{aligned} \quad (6.13)$$

The fraction f_{21} is floated and shared among the subsamples, however due to the similar behaviour of $B_s^0 \to D_s^{*-}\pi^+$ and $B_s^0 \to D_s^-\rho^+$ decays the fraction f_{22} is set to be 0.5.

Group 3 is composed of $\overline{\Lambda}_b^0$ decays with a proton misidentified as kaon ($p \to K^+$) and with the PDF defined as:

$$\begin{aligned}\mathrm{PDF}(m(D_s^\mp h^\pm),&m(h^-h^+h^\pm), \mathrm{PIDK})_{\mathrm{Group3}} = \\ &f_{31} \times \mathrm{PDF}(m(D_s^\mp h^\pm), m(h^-h^+h^\pm), \mathrm{PIDK})_{\mathrm{\Lambda_b^0 \to D_s^- p}} \\ &+ (1-f_{31}) \times \mathrm{PDF}(m(D_s^\mp h^\pm), m(h^-h^+h^\pm), \mathrm{PIDK})_{\mathrm{\Lambda_b^0 \to D_s^{*-} p}}, \end{aligned} \quad (6.14)$$

Table 6.10: Definition of the background groups in the $B_s^0 \to D_s^{\mp} K^{\pm}$ MD fit.

Group	Type	Decay channels
1	Bachelor K	$B^0 \to D_s^- K^+$
2	Bachelor π	$B_s^0 \to D_s^- \pi^+$, $B_s^0 \to D_s^{*-} \pi^+$, $B_s^0 \to D_s^- \rho^+$
3	Λ_b^0 decays	$\Lambda_b^0 \to D_s^- p$, $\Lambda_b^0 \to D_s^{*-} p$
4	–	$B^0 \to D^- K^+$, $B^0 \to D^- \pi^+$
5	–	$\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- K^+$, $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$
6	Combinatorial	–

Table 6.11: Fixed yields in the $B_s^0 \to D_s^{\mp} K^{\pm}$ MD fit.

Background	$D_s^- \to (KK\pi)_{\text{nonres}}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0} K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$B^0 \to D^- K^+$	17.0	0.0	5.0	0.0	0.0
$B^0 \to D^- \pi^+$	14.0	0.0	3.0	3.0	0.0
$\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- K^+$	15.0	2.0	4.0	0.0	0.0
$\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$	11.0	1.0	3.0	0.0	0.0

where f_{31} is fixed from prediction to be 0.75 and it is shared among all subsamples. For groups 2 and 3 the global, shared, fraction f_{51} is introduced which helps to control $\Lambda_b^0 \to D_s^{(*)-} p$ yield in the fitter. The global PDF of **Group 23** is constructed as follows:

$$\begin{aligned} \text{PDF}(f_{51}; m(D_s^{\mp} h^{\pm}), m(h^- h^+ h^{\pm}), \text{PIDK})_{\text{Group23}} = \\ f_{51} \times \text{PDF}(m(D_s^{\mp} h^{\pm}), m(h^- h^+ h^{\pm}), \text{PIDK})_{\text{Group2}} \\ + (1 - f_{51}) \times \text{PDF}(m(D_s^{\mp} h^{\pm}), m(h^- h^+ h^{\pm}), \text{PIDK})_{\text{Group3}}, \end{aligned} \tag{6.15}$$

with floating fraction f_{51}. **Groups 4 and 5** refers to the fully reconstructed, but misidentified backgrounds due to: $B^0 \to D^- h^-$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- h^+$ decays. Since their predicted yields are below 2% of the signal, they are fixed to the values from Tab. 6.11. Finally **Group 6** contains the combinatorial background as discussed in Sec. 6.2.2.

For each D_s^- final state, the total PDF is obtained using Eq. 6.12. Since groups 2 and 3 are merged in one extended PDF, the sum in Eq. 6.12 has six components: $j = \{0, 1, 23, 4, 5, 6\}$, where $j = 0$ denotes the signal and $j = \{1, 23, 4, 5, 6\}$ are groups of backgrounds. The floating parameters are $a_{(B_s^0 \to D_s^{\mp} K^{\pm})} = (\mu_{B_s^0}, \mu_{D_s^-})$ for the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal, and for the combinatorial background: $a_{(Comb)} = (s_{B_s^0}, s_{D_s^-}, fC_{D_s^-}, fC_{\text{PIDK1}}, fC_{\text{PIDK2}})$. In addition, $a_{(Group2)} = f_{21}$ for Group 2 and $a_{(Group23)} = f_{51}$ for the merged Group 23. In total 20 yields and 19 other parameters are left free in the fit.

The fit results to the $B_s^0 \to D_s^{\mp} K^{\pm}$ sample, separately for each considered D_s^- final state, and combined, are shown in Figs. 6.14, 6.15, 6.16, and 6.17. The fitted parameters are listed in Tab. 6.12.

Table 6.12: Fitted values of the parameters for the MD fit to $B_s^0 \to D_s^{\mp} K^{\pm}$ sample. The N_i are the yields of the signal and background contributions. The means $\mu_{B_s^0}$ and $\mu_{D_s^-}$ are the parameters of the double Crystal Ball used to describe the signal in $m(D_s^{\mp} h^{\pm})$ and $m(h^- h^+ h^{\pm})$ invariant masses, respectively. The parameter f_{21} is the fraction between the modes in group 2: $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{*-} \pi^+ / B_s^0 \to D_s^- \rho^+$. The parameter f_{51} is the fraction between groups 2 and 3. The $s_{B_s^0(D_s^-)}$ parameters are the slopes of combinatorial background in 10^{-3} units. The fractions $fC_{D_s^-}$ are the fractions between the exponential and signal shapes in the $m(h^- h^+ h^{\pm})$ invariant mass. The $fC_{\mathrm{PIDK}(1,2)}$ are the fractions between the pion, kaon and proton components in the PIDK combinatorial shape, shared among all subsamples.

Par.	$D_s^- \to (KK\pi)_{\mathrm{nonres}}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$N_{B_s^0 \to D_s^{\mp} K^{\pm}}$	309 ± 21	576 ± 27	475 ± 24	107 ± 13	301 ± 22
N_{Group1}	18 ± 7	34 ± 8	39 ± 9	9 ± 5	27 ± 8
$N_{Group23}$	225 ± 21	498 ± 30	327 ± 24	89 ± 16	258 ± 25
N_{Comb}	487 ± 27	311 ± 26	258 ± 22	428 ± 24	946 ± 37
$N_{B_s^0 \to D_s^{\mp} K^{\pm}}$	total yield: 1768 ± 49				
$s_{B_s^0}$	-3.17 ± 0.41	-1.82 ± 0.63	-2.91 ± 0.61	-1.09 ± 0.38	-1.55 ± 0.27
$s_{D_s^-}$	-4.29 ± 2.05	-2.92 ± 3.54	-3.39 ± 3.10	-0.00 ± 1.02	-1.97 ± 1.36
$fC_{D_s^-}$	0.95 ± 0.05	0.54 ± 0.07	0.80 ± 0.08	not fitted	not fitted
$\mu_{B_s^0}$ [MeV/c^2]	5370.4 ± 0.5				
$\mu_{D_s^-}$ [MeV/c^2]	1969.8 ± 0.1				
f_{21}	0.648 ± 0.034				
f_{51}	0.986 ± 0.049				
fC_{PIDK1}	0.504 ± 0.032				
fC_{PIDK2}	0.346 ± 0.079				

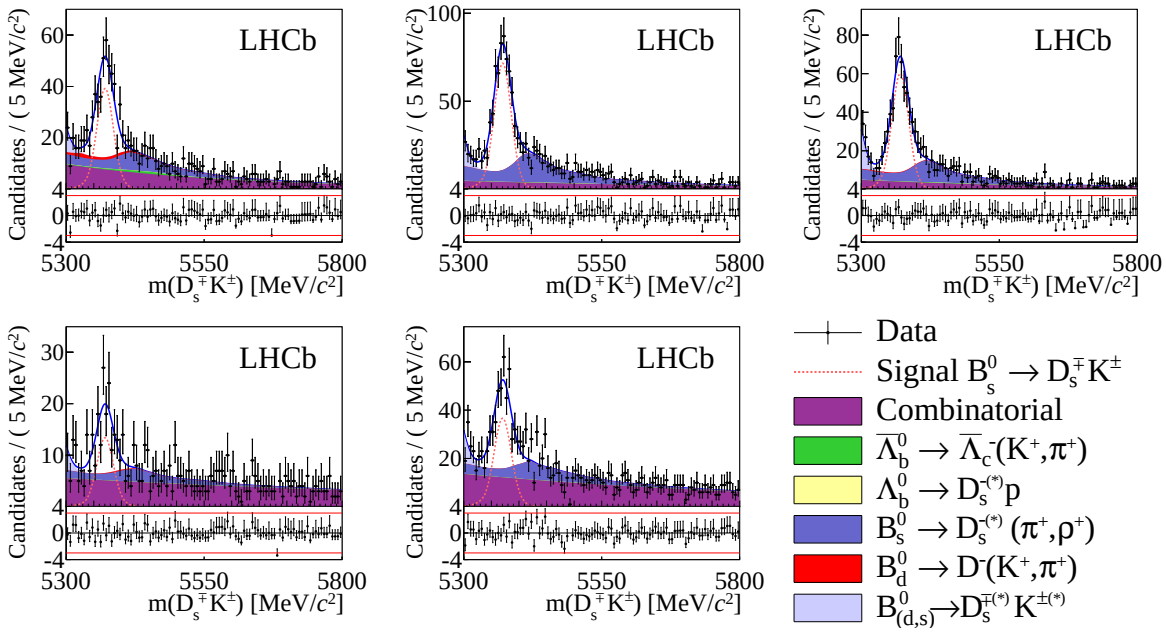

Figure 6.14: Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of $m(D_s^{\mp} h^{\pm})$ invariant mass for D_s^- final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{\mathrm{nonres}}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend placed in the bottom right.

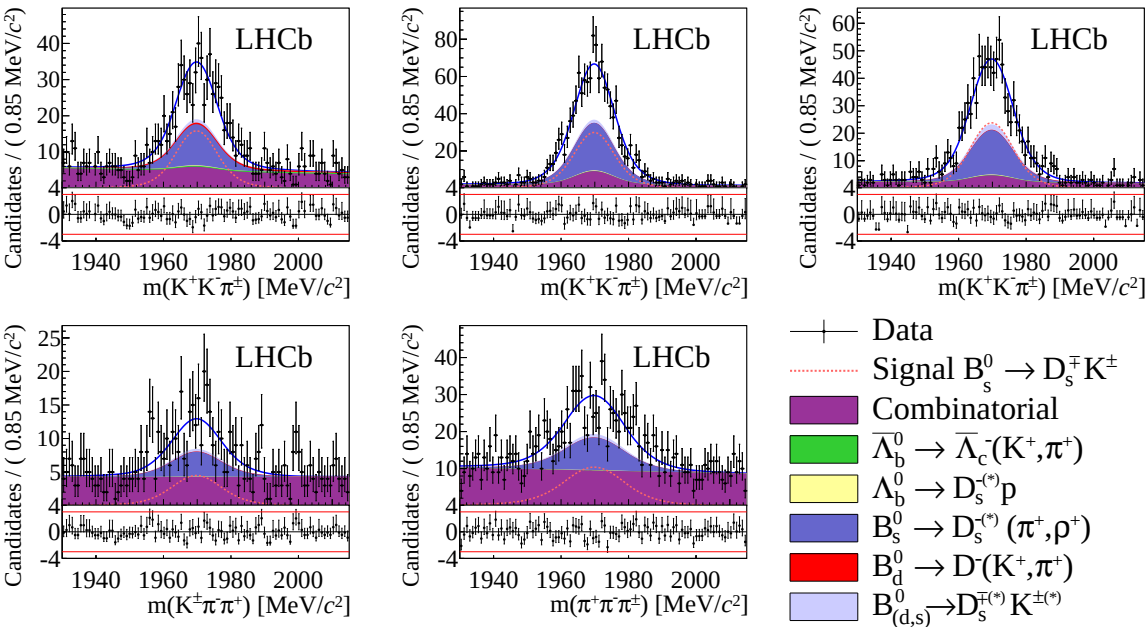

Figure 6.15: Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of $m(h^-h^+h^{\pm})$ invariant mass for D_s^- final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{\rm nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend placed in the bottom right.

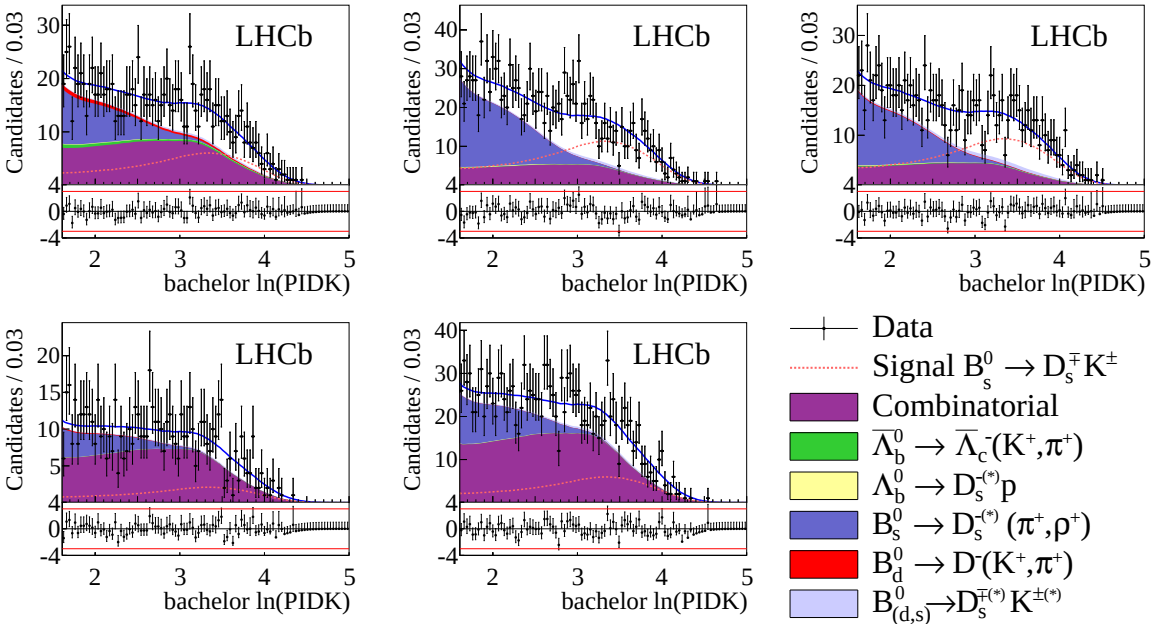

Figure 6.16: Result of the simultaneous fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Distributions of ln(PIDK) for D_s^- final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{\rm nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend placed in the bottom right.

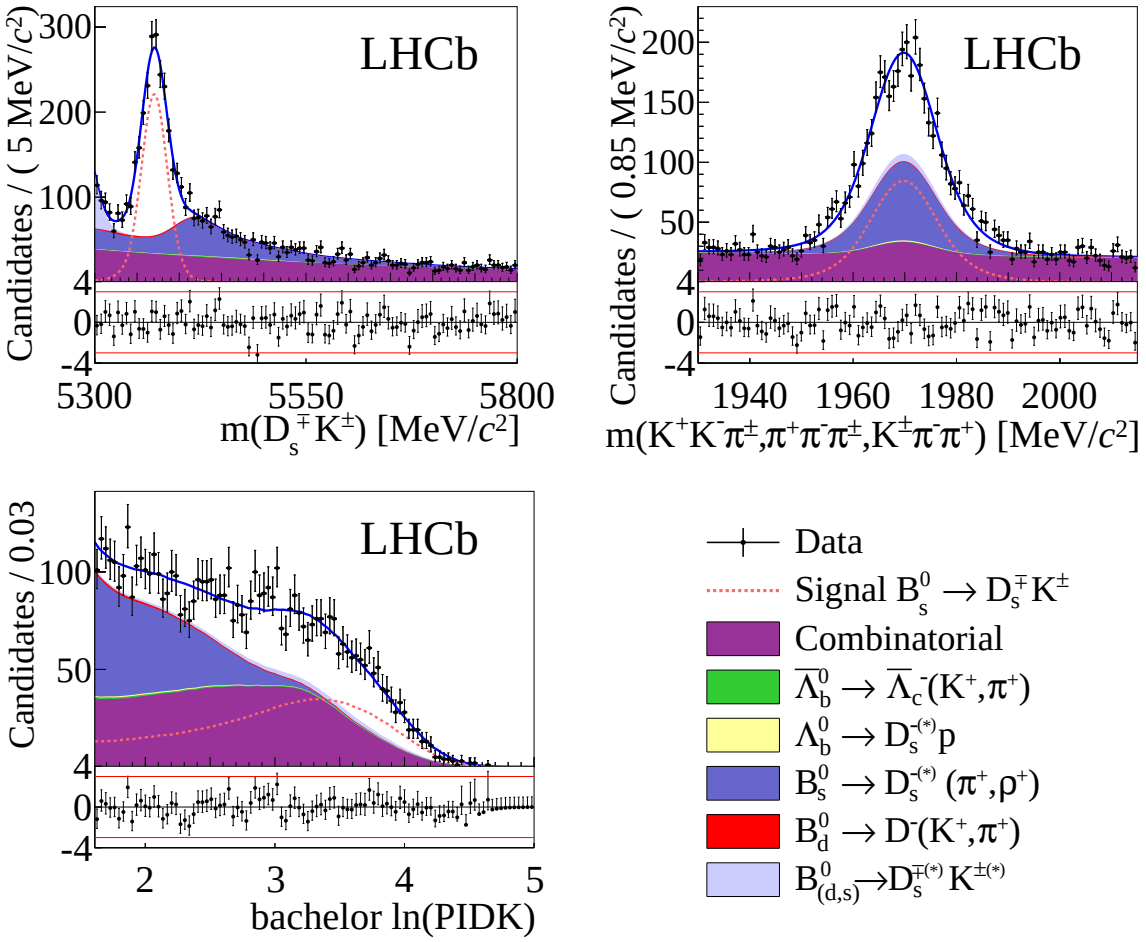

Figure 6.17: The simultaneous fit to the $B_s^0 \to D_s^{\mp}K^{\pm}$ candidates for both magnet polarities and all D_s^- final states combined, from top left to bottom: $m(D_s^{\mp}h^{\pm})$ invariant mass, $m(h^-h^+h^{\pm})$ invariant mass, ln(PIDK) variable. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend placed in the bottom right.

6.6 Multidimensional fit validation

Large ensembles of pseudo experiments are used to both validate the fit procedure and to evaluate systematic uncertainties. These pseudo experiments are generated with a stand-alone code which produces samples in the same format as the data, and these samples are subsequently put through the full data fit procedure: the invariant masses and PIDK distributions are fitted in order to extract yields and *sWeights*, and these results are subsequently fed to the decay-time fitter.

The pseudo experiments are generated to be as data-like as possible, both in terms of the signal yields and in the sense of including all known backgrounds. For both $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ samples, all the backgrounds are generated in roughly (to within about 10% percent) the proportions found in the data fit. The overall sample sizes are also roughly (again, to within about 10% percent) the same as in the data samples. Poisson fluctuations [100] in the yields of all modes between the individual pseudo experiments are allowed. The signal and background shapes are generated using the results of the data MD fit.

The verification of the fitter is based on 1000 pseudo experiments from which the pseudo pull distribution p_α is obtained:

$$p_\alpha = \frac{\alpha_{gen} - \alpha_{fit}}{\sigma_{\alpha;fit}} \tag{6.16}$$

where α denotes the observable, α_{gen} is the generated value, while α_{fit} and $\sigma_{\alpha;fit}$ are the returned by the fitter value and its uncertainty. A Gaussian distribution [101] is fitted to the pseudo pull whose mean μ_p corresponds to the bias between fitted and generated values and a width σ_p is used to uncertainty verification. The $\mu_p \sim 0$ and $\sigma_p \sim 1$ are signatures of an unbiased fitter.

6.6.1 Multidimensional fit validation for $B_s^0 \to D_s^- \pi^+$

The main focus of the $B_s^0 \to D_s^- \pi^+$ fit validation is the verification that the number of signal candidates can be correctly extracted from the fit. As seen in Fig. 6.18 and Tab. 6.13 the pseudo pull distributions for all floating components are reasonably unbiased.

6.6.2 Multidimensional fit validation for $B_s^0 \to D_s^\mp K^\pm$

Additional studies for $B_s^0 \to D_s^\mp K^\pm$ are shown in Fig. 6.19 for the floating yields and the fractions. The results are collected in Tab. 6.14. The relevant pseudo pull for yields are unbiased. A small bias is visible for f_{51}, however it can be explained due to the generated value 0.938, which is close to the limit at 1.0.

Table 6.13: Results of pseudo experiments corresponding to MD fit to the $B_s^0 \to D_s^- \pi^+$ sample.

Parameter	μ_p of pseudo pull	σ_p of pseudo pull	average uncertainty
$N_{B_s^0 \to D_s^- \pi^+}$	0.011 ± 0.036	1.055 ± 0.028	188.7 ± 0.1
N_{Group1}	-0.097 ± 0.031	0.928 ± 0.024	41.27 ± 0.03
N_{Comb}	-0.020 ± 0.033	0.974 ± 0.026	147.4 ± 0.1

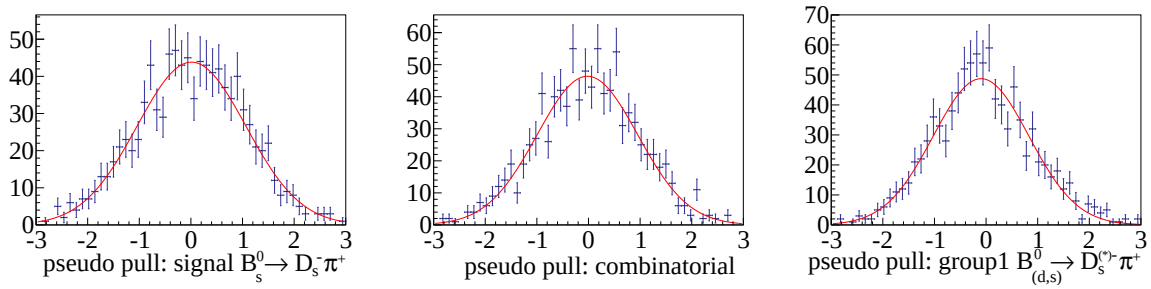

Figure 6.18: Pseudo pulls from pseudo experiments corresponding to MD fit to the $B_s^0 \to D_s^- \pi^+$ sample. From left to right: signal, combinatorial background and Group 1 background.

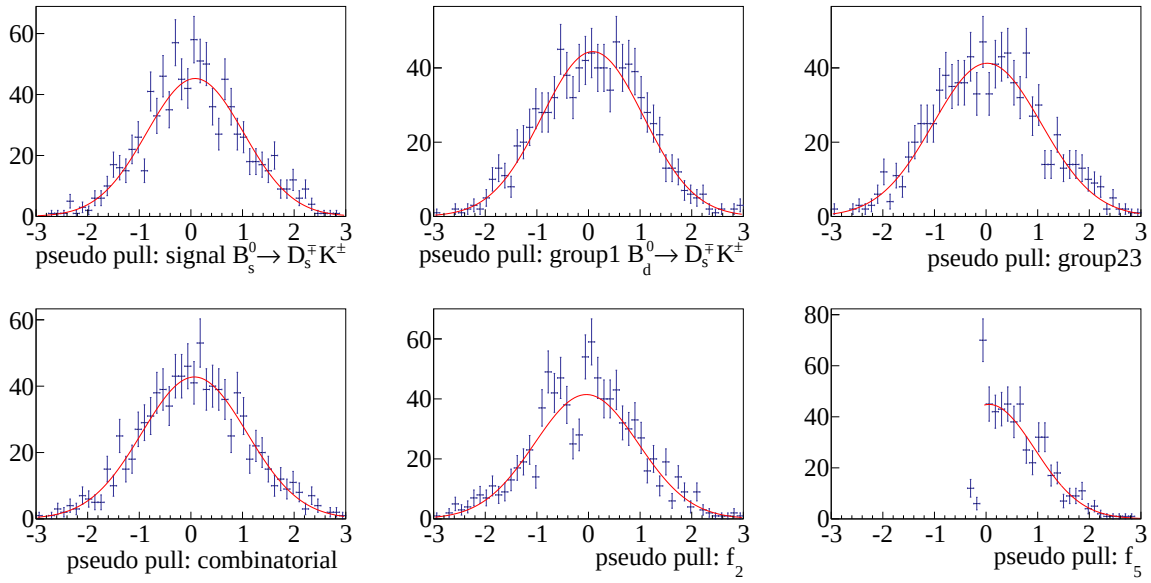

Figure 6.19: Pseudo pulls from pseudo experiments corresponding to MD fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ sample. From top left: signal, group 1, group 23, combinatorial background, fraction f_{21}, f_{51}.

Table 6.14: Results of pseudo experiments corresponding to MD fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ sample.

Parameter	μ_p of pseudo pull	σ_p of pseudo pull	average uncertainty
$N_{B_s^0 \to D_s^{\mp} K^{\pm}}$	-0.026 ± 0.033	0.991 ± 0.026	52.49 ± 0.02
N_{Group1}	-0.015 ± 0.035	1.034 ± 0.028	18.53 ± 0.03
$N_{Group23}$	-0.060 ± 0.036	1.078 ± 0.029	61.07 ± 0.04
N_{Comb}	0.099 ± 0.033	0.994 ± 0.028	82.86 ± 0.06
f_{21}	-0.032 ± 0.033	0.998 ± 0.027	0.0362 ± 0.0001
f_{51}	0.059 ± 0.020	0.891 ± 0.094	0.0421 ± 0.0038

Chapter 7

Flavour Tagging

The decay rates given by Eq. 2.47 and Eq. 2.48 depend on the initially produced flavour eigenstates $|B_s^0(t=0)\rangle$ and $|\overline{B}_s^0(t=0)\rangle$, thus the time-dependent *CP* measurement requires tagging of the initial flavour. That information is provided by the tagging algorithms which determine whether beauty mesons contained a b or $\overline{b}$ quark at production time. The theoretical and experimental aspects of the flavour tagging are described in the following sections.

7.1 Principles of the flavour tagging

The tagging decision is defined as $q_t = \{-1, 0, +1\}$, where $q_t = -1$ ($q_t = 1$) denotes $\overline{B}_s^0$ (B_s^0) in the initial state. For the candidates where the algorithm does not provide tagging information, the value $q_t = 0$ is assigned. The tagging efficiency is given by

$$\epsilon^{tag} = \frac{N_{tag}}{N_{all}}, \tag{7.1}$$

where N_{tag} is the number of candidates with tagging decision $q_t = \pm 1$, and N_{all} denotes all candidates in the sample. The tagging efficiency can be written as a function of the tagging decision:

$$\epsilon^{tag}(q_t) = \begin{cases} \epsilon^{tag}, & \text{for tagged candidates } q_t = \pm 1, \\ (1 - \epsilon^{tag}), & \text{for untagged candidates } q_t = 0, \end{cases} \tag{7.2}$$

In addition, each tagging algorithm has an intrinsic mistag rate ω defined as:

$$\omega = \frac{N_{wrong}}{N_{tag}}, \tag{7.3}$$

where N_{wrong} is the number of incorrectly tagged candidates.

The *CP* violating coefficients that are accessible in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays directly depend on the dilution D due to the mistag probability ω,

$$\mathrm{D} = 1 - 2\omega \ . \tag{7.4}$$

The statistical uncertainty on C, S_f and $S_{\bar{f}}$ scales with $1/\sqrt{\varepsilon_{\text{eff}}}$ defined as:

$$\varepsilon_{\text{eff}} = \epsilon^{tag}(1 - 2\omega)^2. \tag{7.5}$$

The tagging algorithms are optimised for highest ε_{eff} on data using the self-tagging decays such as $B^+ \to J/\psi K^+$ or $B_s^0 \to D_s^- \pi^+$.

7.2 Effect of the flavour tagging to the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay rates

The decay rates given by the Eq. 2.47 and Eq. 2.48 can be written as a combination of the *CP*-even and *CP*-odd parts. The *CP*-even part does not change sign when the *CP* transformation is applied to either the initial state or the final state, while the *CP*-odd part does. Using this convention the decay rates are given by:

$$\frac{d\Gamma_{q_t,q_f}(t)}{dte^{-\Gamma_s t}} \sim \left[E_{q_f}(t) + q_t q_f O_{q_f}(t)\right], \tag{7.6}$$

where:

$$E_{q_f}(t) = \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + D_{q_f} \sinh\left(\frac{\Delta\Gamma_s t}{2}\right), \tag{7.7}$$

$$O_{q_f}(t) = C_{q_f} \cos\left(\Delta m_s t\right) - S_{q_f} \sin\left(\Delta m_s t\right). \tag{7.8}$$

Here q_t is the tagging decision, q_f denotes the bachelor final state: for the positive (negative) charge the $q_f = +1(-1)$ is assigned. Finally the coefficients C_{q_f}, S_{q_f} and D_{q_f} correspond to C_f, S_f and D_f for $q_f = +1$ as well as $C_{\bar{f}}$, $S_{\bar{f}}$ and $D_{\bar{f}}$ for $q_f = -1$.

The full discriminating power between the four interfering decays is achieved by using tagging information and bachelor charge as observables. The four decay rates can be combined to:

$$\begin{aligned}\frac{d\Gamma_{B_s^0 \to D_s^{\mp} K^{\pm}}(t, q_t, q_f)}{dte^{-\Gamma_s t}} &= (2 - |q_t|) \times \epsilon^{tag}(q_t) \times \cosh\left(\frac{\Delta\Gamma_s}{2}\right) \\ &+ (2 - |q_t|) \times \epsilon^{tag}(q_t) \times D(q_f) \times \sinh\left(\frac{\Delta\Gamma_s}{2}\right) \\ &+ q_t q_f \times (1 - 2\omega) \times \epsilon^{tag}(q_t) \times C \times \cos\left(\Delta m_s t\right) \\ &- q_t q_f \times (1 - 2\omega) \times \epsilon^{tag}(q_t) \times S(q_f) \times \sin\left(\Delta m_s t\right),\end{aligned} \tag{7.9}$$

where:

$$D(q_f) = \begin{cases} D_f, & \text{for } q_f = 1, \\ D_{\bar{f}}, & \text{for } q_f = -1, \end{cases} \qquad S(q_f) = \begin{cases} S_f, & \text{for } q_f = 1, \\ S_{\bar{f}}, & \text{for } q_f = -1. \end{cases} \tag{7.10}$$

Figure 7.1 shows a simulation of the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay rates given by Eq. 7.9 in decay-time range $\tau \in [0.2, 3.0]$ ps with perfect and imperfect tagging performance. The perfect tagging performance corresponds to $\omega = 0\%$ and $\epsilon^{tag} = 100\%$, while as an example of imperfect performance the values $\omega = 35\%$ and $\epsilon^{tag} = 60\%$ were chosen. The effects from decay-time acceptance and resolution are not taken into account. In the simulation the *CP* parameters defined by Eq. 2.53 correspond to the CKM angle $\gamma = 70°$, the strong phase $\delta = 30°$ and $r_{D_s K} = 0.37$.

7.3 Flavour tagging issues specific to LHCb experiment

At the LHCb experiment, two types of flavour-tagging algorithms (opposite side and signal side ones) are used [102] as shown in Fig. 7.2. The opposite side (OS) tagging algorithms

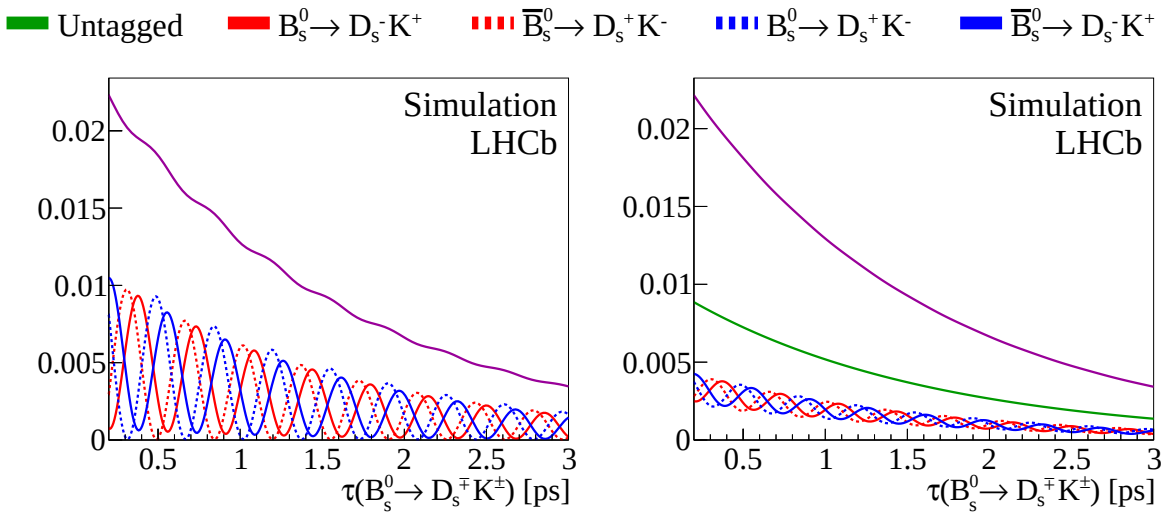

Figure 7.1: Simulation of the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay rates in the decay-time range $\tau \in [0.2, 3.0]$ ps with perfect (left plot) and imperfect (right plot) tagging performance. Different contributions are shown as coloured solid/dashed lines as described in the legend placed in the top.

rely on the pair production of b and $\bar{b}$ quarks and infer the flavour of a given signal beauty meson from the identification of the flavour of the other non-signal b-hadron. In the presented analysis, four OS algorithms are considered:

- **Opposite side kaon tagger (OSK)**,
  which reconstructs candidates from the $b \to c \to s$ decay chain. Therefore $K^-(K^+)$ mesons are signatures of $\bar{b}(b)$-hadrons.

- **Two opposite side lepton taggers for muon (OSμ) and electron (OSe)**,
  which infer the presence of a b quark using the charge of the lepton l, either a muon or an electron, from semileptonic $b \to c l^- \overline{\nu}_l$ decays. The negative (positive) charge of the lepton corresponds to $\bar{b}(b)$-hadrons. Lepton taggers are limited by the branching fraction of the decays in question, but have a high purity.

- **Vertex charge tagger (Q_{vtx})**,
  in which two tracks with high IPχ^2 and transverse momentum p_T are reconstructed as the vertex candidates. Based on geometric and kinematic properties, the probability to originate from a b-hadron decay is obtained for each two tracks combination. Afterwards other tracks, which originate from the vertex in question, are added to the vertex candidate. The overall vertex charge is calculated by summing the weighted charges of the tracks by their transverse momentum. Depending on the value of the overall vertex charge the tagging decision is assigned.

The OS taggers are combined in a single OS response.

Another type of tagger is the same (signal) side kaon tagger (SSK), which determines the signal B_s^0 flavour by exploiting its fragmentation. In the hadronisation of a b quark to a B_s^0 meson, an $s\overline{s}$ quark pair is produced from the vacuum. In about 50% of the cases an additional s quark forms a charged kaon. The positive and negative charges of the kaon correspond to B_s^0 and $\overline{B}_s^0$ meson identification, respectively.

There exist several physical sources of mistag in the above algorithms. In case of the opposite side taggers the crucial one is due to b hadronisation: in about 50% of the cases

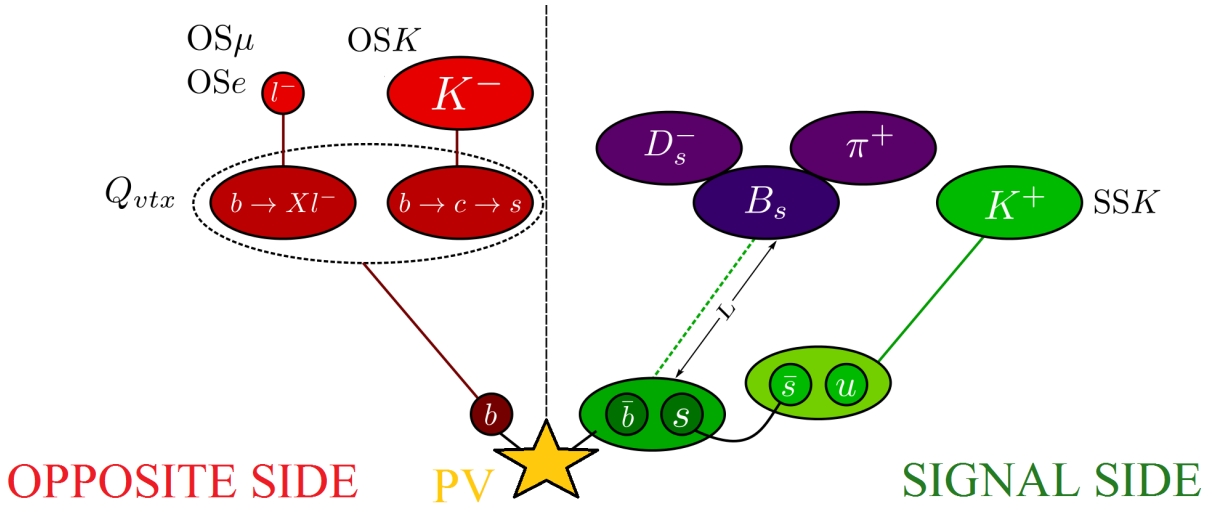

Figure 7.2: The diagram of the flavour algorithms used in the LHCb experiment Ref. [103].

the non-signal b quark hadronises to neutral beauty mesons. As discussed in Sec. 2.4.1 neutral mesons oscillate, thus 19% of B^0 and 50% B_s^0 mesons change their flavour before decaying. Therefore the information about the original b quark is lost. In addition, kaon or lepton taggers can select a track with random charge from the underlying event, which in 50% of cases lead to wrong tagging decision. Moreover, beauty mesons can decay through the chain $b \to c \to sl^+$ instead of the direct semileptonic transition. In this case, the resulting lepton carries a wrong charge and thus the wrong tagging decision for OSμ and OSe. Finally, for the vertex charge tagger, the track can be wrongly assigned to the vertex. The main source of a wrong tagging decision for the SSK is when the associated s quark does not form a charged kaon.

7.4 Tagging calibration

The mistag probability, η, is predicted for each B_s^0 candidate using a neural network trained on simulated candidates. The inputs to the neural network are the kinematic, geometric and PID properties of tagging particles. In the LHCb experiment a dedicated group provides trained neural networks to time-dependent measurements. Due to variations in the properties of tagging tracks for different channels, the predicted mistag probability η has to be calibrated specifically for each analysis, using flavour specific, self-tagging, decays.

The estimated mistag probability η is considered as a per-candidate variable, thus adding an observable to the decay-time fit described in Sec.9. It is used by the fit to assign a larger weight to the candidates that have a lower probability of being tagged wrongly. The calibration of the mistag rate is considered by using the following linear function:

$$\omega = p_0 + p_1 \times (\eta - \langle\eta\rangle) \ , \tag{7.11}$$

where the values of parameters p_0 and p_1 are taken form the measurements done on self-tagging decays. This parametric relation minimises the correlation between the coefficients p_0 and p_1.

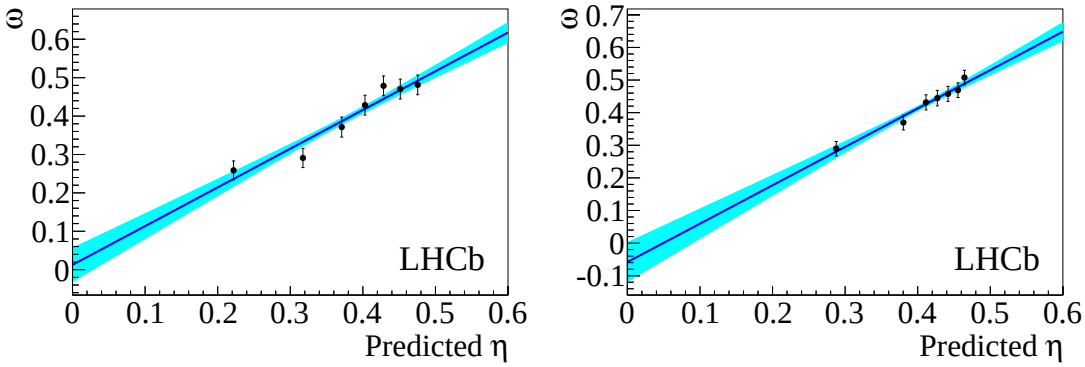

Figure 7.3: The measured mistag rate against the average predicted mistag rate for OS (left) and SSK (right) taggers in $B_s^0 \to D_s^- \pi^+$ decays. The error bars represent only the statistical uncertainties. The solid curve is the linear fit to the data points, the shaded area defines the 68% confidence level region of the calibration function (statistical only).

7.4.1 Opposite Side tagging calibration

The calibration procedure for the OS tagging is done using several control modes: $B^+ \to J/\psi K^+$, $B^+ \to D^0\pi^+$, $B^0 \to D^{*-}\mu^+\nu_\mu$, $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to D_s^-\pi^+$. Among these the calibration obtained using the $B_s^0 \to D_s^-\pi^+$ decays is particularly important for the $B_s^0 \to D_s^\mp K^\pm$ analysis because of the similarities in terms of trigger, reconstruction and kinematics. In this case the tagging calibration is determined using the fit method described in Sec. 9, where the fit to the decay-time is performed with Δm_s fixed to the LHCb measured value of $17.768\,\mathrm{ps}^{-1}$ [4]. The linear dependence of the mistag on the predicted mistag probability, η, is assumed as given by Eq. 7.11. On the left part of Fig. 7.3 the measured mistag ω is plotted against the predicted mistag η. For the OS tagger the weighted average of the fitted $p_0 - \langle\eta\rangle$ and p_1 values in the five control channels is used, since the measurements are compatible. The final average values are quoted in Tab. 7.1 with a statistical uncertainty including the systematics related to each control channel and a systematic uncertainty which accounts for phase space differences among different channels. The correlation between p_0 and p_1, $\rho(p_0, p_1)$ is of 0.13. The resulting calibration parameters for the $B_s^0 \to D_s^\mp K^\pm$ fit are: $p_0 = 0.3834 \pm 0.0014 \pm 0.0040$ and $p_1 = 0.972 \pm 0.012 \pm 0.035$, where the p_0 for each control channel needs to be translated to the $\langle\eta\rangle$ of $B_s^0 \to D_s^-\pi^+$ decay i.e. the channel which is most similar to the signal channel $B_s^0 \to D_s^\mp K^\pm$. This is achieved by the transformation $p_0 \to p_0 + p_1(\langle\eta\rangle - 0.3813)$ in each control channel.

7.4.2 Same Side Kaon tagging calibration

The SSK tagger can be calibrated only in the $B_s^0 \to D_s^-\pi^+$ channel, since other fully reconstructed B_s^0 self-tagging channels have too low statistics, while semileptonic B_s^0 decays have large yields, but are not used for tagging calibration purposes due to the poor time resolution.

The SSK algorithm uses a neural network to select fragmentation particles, giving improved flavour tagging power [103] with respect to earlier cut-based [104] algorithms. It is calibrated on the $B_s^0 \to D_s^-\pi^+$ mode using a similar procedure as for the OS tagger, resulting in $\langle\eta\rangle = 0.4097$, $p_0 = 0.4244 \pm 0.0086 \pm 0.0071$ and $p_1 = 1.255 \pm 0.140 \pm 0.104$, where the first uncertainty is statistical and second systematic. The systematic uncertainties include

Table 7.1: Calibration parameters (described in the text) of the combined OS tagger extracted from different control channels. In each entry the first uncertainty is statistical and the second systematic.

Control channel	$\langle\eta\rangle$	$p_0 - \langle\eta\rangle$	p_1
$B^+ \to J/\psi K^+$	0.3919	0.0008 ± 0.0014 ± 0.0015	0.982 ± 0.017 ± 0.005
$B^+ \to \overline{D}^0\pi^+$	0.3836	0.0018 ± 0.0016 ± 0.0015	0.972 ± 0.017 ± 0.005
$B^0 \to J/\psi K^{*0}$	0.390	0.0090 ± 0.0030 ± 0.0060	0.882 ± 0.043 ± 0.039
$B^0 \to D^{*-}\mu^+\nu_\mu$	0.3872	0.0081 ± 0.0019 ± 0.0069	0.946 ± 0.019 ± 0.061
$B_s^0 \to D_s^-\pi^+$	0.3813	0.0159 ± 0.0097 ± 0.0071	1.000 ± 0.116 ± 0.047
Average	0.3813	0.0021 ± 0.0014 ± 0.0040	0.972 ± 0.012 ± 0.035

Table 7.2: Initial flavour asymmetry for OS tagging extracted from $B^+ \to J/\psi K^+$ decays and for SSK tagging extracted from prompt $D_s^\pm$ reweighting the p_T distribution to $B_s^0 \to D_s^-\pi^+$.

tagger	Δp_0	Δp_1	$\Delta\epsilon^{tag}$ [%]
OS	0.0124 ± 0.0021	0.095 ± 0.024	-0.197 ± 0.126
SSK	-0.020 ± 0.004	-0.01 ± 0.03	0.022 ± 0.004

the error on the decay-time resolution, the $B_s^0 \to D_s^-\pi^+$ fit model, and the backgrounds in the $B_s^0 \to D_s^-\pi^+$ fit. The correlation between p_0 and p_1 is small, less than 0.03. The right part of Fig. 7.3 shows the measured mistag rate ω as a function of the predicted mistag rate η in $B_s^0 \to D_s^-\pi^+$ decays for SSK tagger. The data points show a linear correlation corresponding to the functional form in Eq. 7.11.

7.4.3 Tagging asymmetries

In addition to the calibrations discussed in the previous sections, each tagger was calibrated on subsamples split by initial charge, to address a possible tagging asymmetry. The following relations allow for different calibrations on subsamples:

$$\omega = p_0 + \frac{\Delta p_0}{2} + (p_1 + \frac{\Delta p_1}{2}) \times (\eta - \langle\eta\rangle), \tag{7.12}$$

$$\bar{\omega} = p_0 - \frac{\Delta p_0}{2} + (p_1 - \frac{\Delta p_1}{2}) \times (\eta - \langle\eta\rangle)\,. \tag{7.13}$$

Here, p_0 and p_1 are the calibration constants from the previous sections, and Δp_0 and Δp_1 parametrise the differences. For the OS tagger the initial flavour asymmetry is easily evaluated through the charge of the b meson in the decay $B^+ \to J/\psi K^+$. For the SSK tagger, it is not possible to evaluate p_0 and p_1 separately for B_s^0 and $\overline{B}_s^0$ due to oscillations. Since the initial flavour asymmetry describes the effect on the tagging parameters of the different interactions between matter and anti-matter, is evaluated separately for K^+ and K^- tagging particles. In Tab. 7.2 the initial flavour asymmetry for p_0, p_1 and ϵ^{tag} are reported.

Table 7.3: The Flavour Tagging performances for only OS tagged, only SSK tagged and both OS and SSK tagged candidates for $B_s^0 \to D_s^- \pi^+$.

$B_s^0 \to D_s^- \pi^+$	ϵ^{tag} [%]	ε_{eff} [%]
OS only	19.80 ± 0.23	1.61 ± 0.03 ± 0.08
SSK only	28.85 ± 0.27	1.31 ± 0.22 ± 0.17
both OS-SSK	18.88 ± 0.23	2.15 ± 0.05 ± 0.09
total	67.53	5.07

7.5 The Opposite Side and Same Side Kaon taggers combination

The correlation between the OS and SSK taggers, as shown in Fig. 6.1, is negligible so the taggers and their asymmetries can be combined. It is crucial to note that any such combination only makes sense using calibrated taggers.

The data sample is split in three different subsamples depending on the tagging decision: only OS tagged, only SSK tagged and both OS and SSK tagged candidates. In the third sample the mistag probability ω from the OS and the SSK taggers are combined together to give the tagging decision and a combined ω for each candidate. The tagging performances in the three subsamples are reported in Tab. 7.3 for $B_s^0 \to D_s^- \pi^+$. The total tagging efficiency is 67.53%, while the tagging power is equal to 5.07%. In addition, the uncertainties on the calibration of the OS and SSK tagger are propagated to the combination of both taggers.

Taggers are combined in the following manner. For each tagger $i \in \{OS, SSK\}$ the difference of logarithms DLL is computed:

$$DLL_i = \ln(P(B|q_{ti}, \omega_i(\eta_i))) - \ln(P(\overline{B}|q_{ti}, \omega_i(\eta_i))) = \ln \frac{P(B|q_{ti}, \omega_i(\eta_i))}{P(\overline{B}|q_{ti}, \omega_i(\eta_i))}, \tag{7.14}$$

where q_{ti} denotes the decision of a single tagger, $\omega_i(\eta_i)$ is the calibrated mistag and:

$$P(B|q_{ti}, \omega_i(\eta_i)) = \prod_i \left(\frac{1 - q_{ti}}{2} - q_{ti} \left(1 - \omega_i(\eta_i)\right) \right), \tag{7.15}$$

$$P(\overline{B}|q_{ti}, \omega_i(\eta_i)) = \prod_i \left(\frac{1 + q_{ti}}{2} - q_{ti} \left(1 - \omega_i(\eta_i)\right) \right). \tag{7.16}$$

If $\omega_i(\eta_i)$ is larger or equal to 0.5, DLL_i is set to zero, which corresponds to an untagged candidate. Then, the difference of logarithms for single taggers are summed:

$$DLL = \sum_i DLL_i. \tag{7.17}$$

Finally, the combined mistag ω_c is defined as:

$$\omega_c = \frac{1}{1 + e^{|DLL|}}. \tag{7.18}$$

The combined mistag ω_c PDFs, obtained from $sWeighted$ $B_s^0 \to D_s^- \pi^+$ data, are shown in Fig. 7.4. The combined tagging decisions q_c are collected in Tab. 7.4. The decay rates given by Eq. 7.9 are updated to this convention.

Table 7.4: The combined tagging decision q_c.

	only OS	only SSK	both OS and SSK	untagged
B_s^0	1	2	3	-
$\overline{B}_s^0$	-1	-2	-3	-
untagged	-	-	-	0

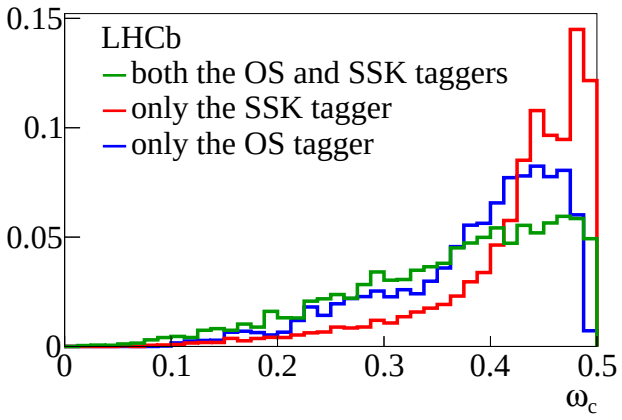

Figure 7.4: The combined mistag ω_c PDFs obtained from *sWeighted* $B_s^0 \to D_s^- \pi^+$ data.

7.6 Comparison of tagging performance in $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ decays

In order to use the mode $B_s^0 \to D_s^- \pi^+$ as a tagging calibration channel for $B_s^0 \to D_s^\mp K^\pm$, it is mandatory to confirm that the taggers perform identically on the two decays. Figure 7.5 shows the comparison between η distributions for $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ samples, as obtained on simulated candidates. A good agreement between the distributions is found.

In addition, the mistag distribution ω as a function of η (given by Eq. 7.11) for simulated $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ candidates is shown in Fig. 7.6 separately for OS and SSK taggers. As expected the dependence of ω on η appears to be linear. The results of calibrations are collected in Tab. 7.5. The fitted values are in good agreement. These fits demonstrate that the calibration parameters from the $B_s^0 \to D_s^- \pi^+$ decay are portable to $B_s^0 \to D_s^\mp K^\pm$. Finally, the comparison between data samples weighted by *sWeights* is shown in Fig. 7.7.

Table 7.5: Results of the calibration obtained on signal $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ simulation. The parameters p_0 and p_1 are described in the text.

	Opposite side		Same side	
	p_0	p_1	p_0	p_1
$B_s^0 \to D_s^- \pi^+$	0.364 ± 0.002	0.848 ± 0.016	0.403 ± 0.001	1.332 ± 0.012
$B_s^0 \to D_s^\mp K^\pm$	0.359 ± 0.003	0.824 ± 0.040	0.404 ± 0.003	1.323 ± 0.003

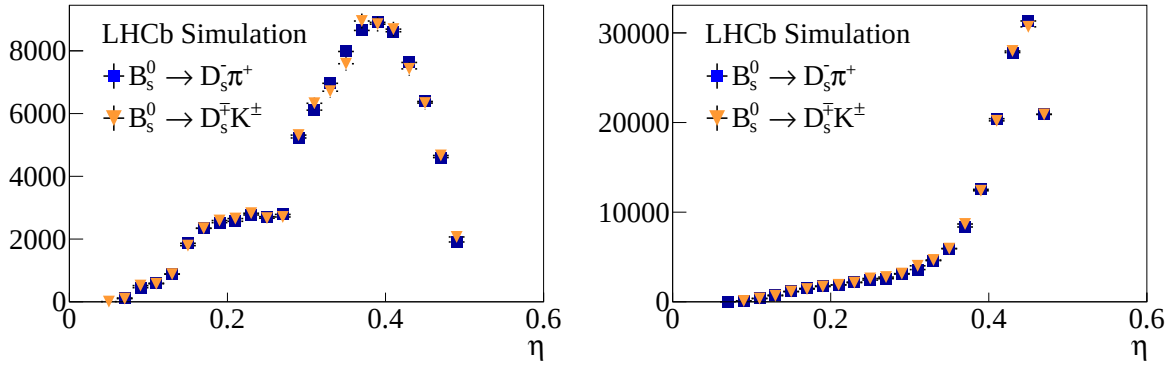

Figure 7.5: Predicted η distributions for the OS combination (left) and the SSK tagger (right) in simulation. The blue (orange) points correspond to the $B_s^0 \to D_s^- \pi^+$ ($B_s^0 \to D_s^\mp K^\pm$) decay.

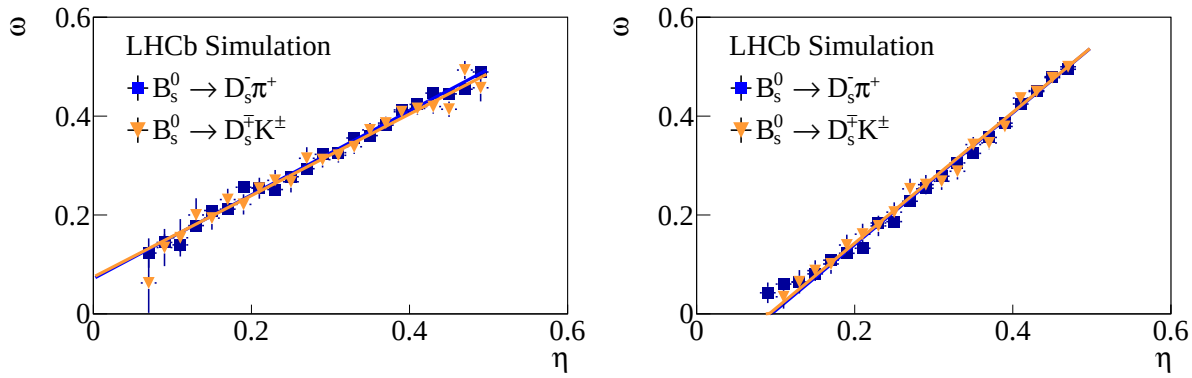

Figure 7.6: Dependence of the observed mistag ω on calibrated mistag η, for selected $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ candidates in simulation. Left is OS tagging, right is SSK. The blue (orange) points correspond to the $B_s^0 \to D_s^- \pi^+$ ($B_s^0 \to D_s^\mp K^\pm$) decay.

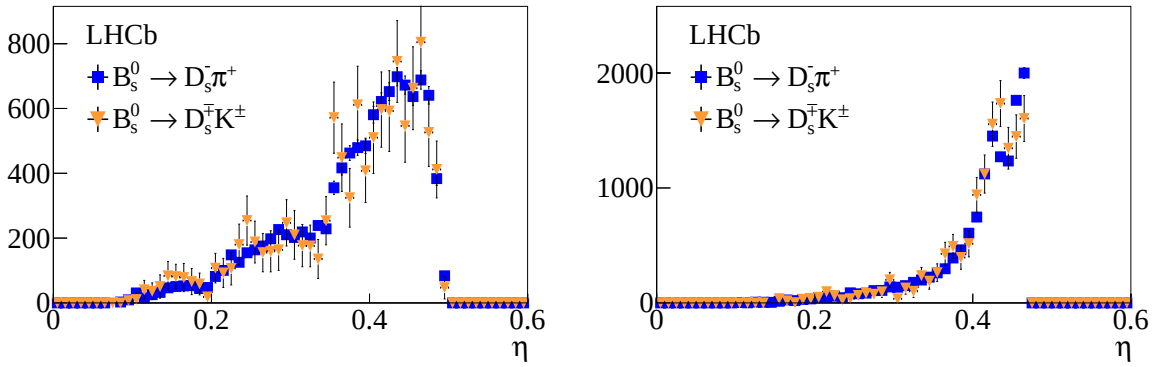

Figure 7.7: Predicted η distributions for the OS combination (left) and SSK tagger (right) in data. The blue (orange) points correspond to the $B_s^0 \to D_s^- \pi^+$ ($B_s^0 \to D_s^\mp K^\pm$) decay.

Chapter 8

Decay-time resolution and acceptance

The fast B_s^0–$\overline{B}_s^0$ oscillations require the finite decay-time resolution of the detector to be taken into account. In addition, the signal selection described in Ch. 5 results in a decay-time acceptance, which distorts the measured decay-time distribution. Any mismodelling of either effect can potentially bias the time-dependent *CP* violation observables. Figure 7.1 from the previous chapter shows the simulation of the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay rates in which the decay-time resolution and acceptance are neglected. These effects are taken into account in Fig. 8.1 where the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay rates are presented with a realistic decay-time acceptance (left figure) or a realistic decay-time resolution (right figure), for clarity with perfect tagging performance. Finally, both effects are applied in Fig. 8.2, where the left (right) plot corresponds to perfect (realistic) tagging performance. Sections 8.1 and 8.2 describe in detail the decay-time resolution and the decay-time acceptance used in the presented measurement.

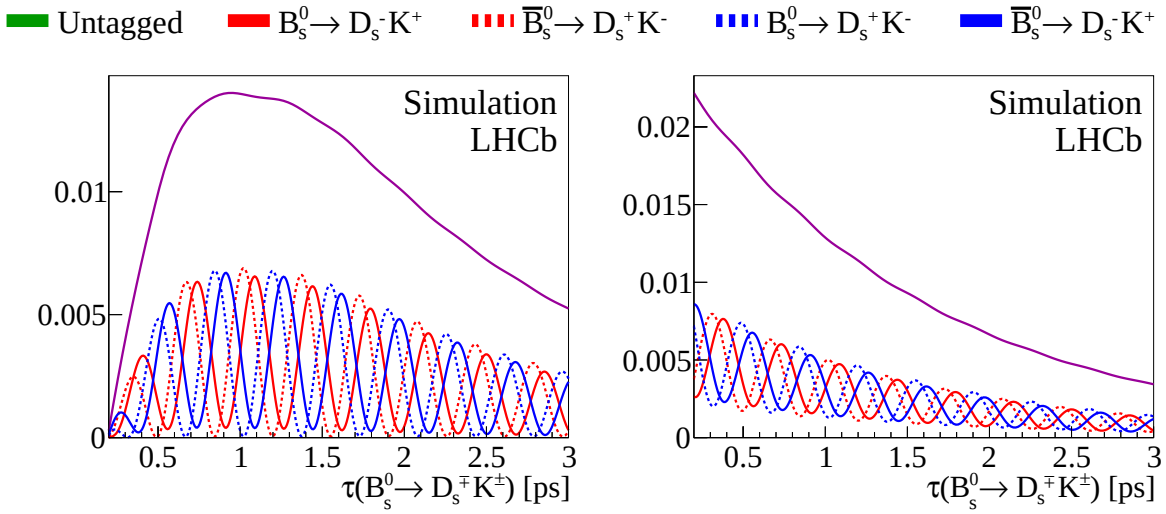

Figure 8.1: Simulation of the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay rates in the decay-time range $\tau \in [0.2, 3.0]$ ps with perfect tagging performance and realistic decay-time acceptance (left plot) or realistic decay-time resolution (right plot). Different contributions are shown as coloured solid/dashed lines as described in the legend placed in the top.

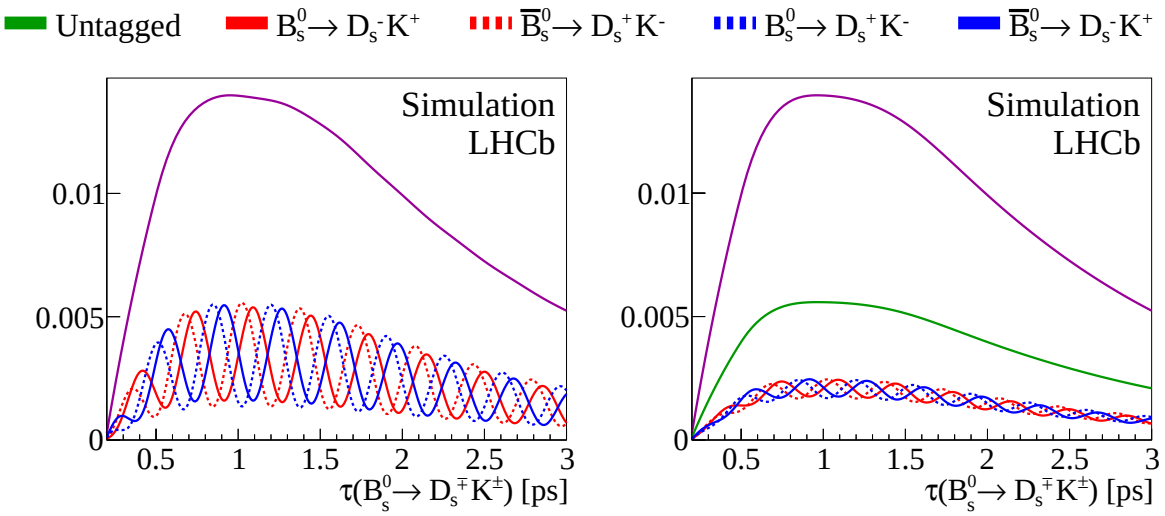

Figure 8.2: Simulation of the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay rates in the decay-time range $\tau \in [0.2, 3.0]$ ps with realistic decay-time acceptance and decay-time resolution. Left (right) plot corresponds to perfect (realistic) tagging performance. Different contributions are shown as coloured solid/dashed lines as described in the legend placed in the top.

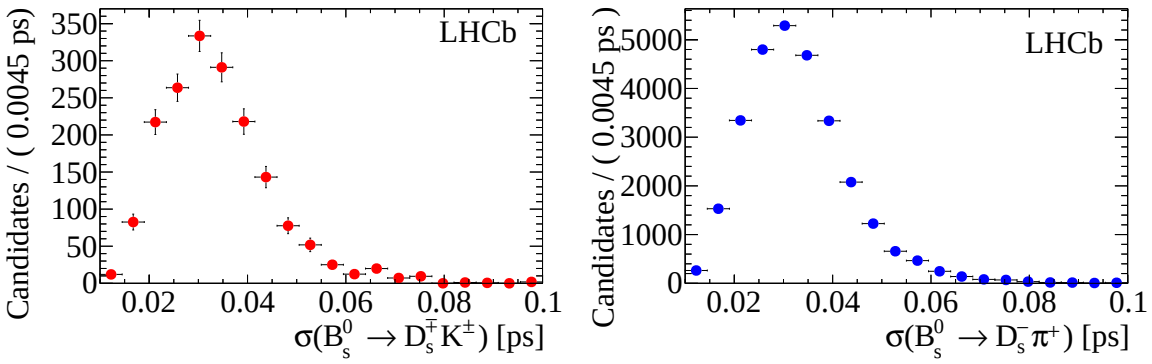

Figure 8.3: The estimated decay-time uncertainty from the *sWeighted* data for $B_s^0 \to D_s^{\mp} K^{\pm}$ (left) and $B_s^0 \to D_s^- \pi^+$ (right) candidates.

8.1 Decay-time resolution

The finite decay time resolution of the detector leads to a dilution of the observable oscillation. The decay-time t of B_s^0 mesons is defined as the time between their production and decay in the rest frame of the B_s^0:

$$t = \frac{(\vec{d} \cdot \vec{p}) m}{|\vec{p}|^2}, \tag{8.1}$$

where m is the invariant mass, $\vec{d}$ denotes the vector between the production and decay, while $\vec{p}$ corresponds to the particle's momentum. For each candidate i a global kinematic fit determines an estimate of the uncertainty σ_t^i of the reconstructed decay-time. The uncertainty depends on the B_s^0 production, the momentum and secondary vertex position. The estimated decay-time uncertainties for signal $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ candidates obtained on *sWeighted* data are shown in Fig. 8.3.

To obtain the experimental decay rates $\frac{d\Gamma(t)_{\rm exp}}{dt}$, the theoretical decay rates, including the

flavour tagging (given Eq. 7.9), have to be convoluted with the decay-time resolution [105]:

$$\frac{d\Gamma(t)_{\text{exp}}}{dt} = \frac{d\Gamma(t)}{dt} \otimes R(t) \equiv \int_{-\infty}^{+\infty} \frac{d\Gamma(t')}{dt'} R(t-t')\Theta(t')dt', \tag{8.2}$$

where $R(t)$ denotes the time resolution and $\Theta(t)$ is the Heaviside step function[7]. Considering the simplified case when $\frac{d\Gamma(t)}{dt} \propto \cos(\Delta m_s t)$ and the decay-time resolution is parametrised by a single Gaussian $R(t) = G(0, \sigma_t; t)$ with mean $\mu = 0$ and width $\sigma = \sigma_t$, the experimental model is given by:

$$\frac{d\Gamma(t)}{dt} \otimes G(0, \sigma_t; t) \propto e^{-\frac{1}{2}\Delta m_s^2 \cdot \sigma_t^2} \cdot \cos(\Delta m_s t) = D_{res} \cdot \cos(\Delta m_s t), \tag{8.3}$$

where the D_{res} is the decay-time dilution factor.

There are two main ways to describe the decay-time resolution in the decay-time fit. The first approach assumes a fixed resolution function, typically a double or triple Gaussian. The second method corresponds to a resolution function that has a different width for each candidate, making use of the per-candidate decay-time uncertainty, estimated during reconstruction (Fig. 8.3). This method is in general more precise, but also more computationally expensive. In the nominal decay-time *sFit* described in Sec. 9, the per-candidate decay-time resolution is used. However, as a cross-check, a triple Gaussian decay-time resolution is also obtained. Detailed information about that approach can be found in App. G. The average decay-time resolution is found to be 47 fs.

Since the estimated uncertainty from the reconstruction does not correspond to the true uncertainty, it has to be properly calibrated. The calibration is taken from the Δm_s measurement [4] and it assumes a scale factor s_{σ_t} between the estimated per-candidate uncertainty σ_t^i and the real resolution σ_t given by:

$$\sigma_t = s_{\sigma_t} \cdot \sigma_t^i. \tag{8.4}$$

The calibration is performed using prompt D_s^- mesons combined with a random track and kinematically weighted to give a sample of so-called fake B_s^0 candidates, whose the secondary vertex is identical to the primary vertex. By construction the true decay-time of the fake B_s^0 candidates is equal to zero. The comparison between the true and fake B_s^0 candidates is shown in Fig. 8.4.

Finally, the distribution of the reconstructed decay-time divided by the per-candidate estimated uncertainty is fitted as a Gaussian. The width of the Gaussian corresponds to the scale factor. A value greater than one indicates that the estimated uncertainty underestimates the true uncertainty. The scale factor is found to be $s_{\sigma_t} = 1.37 \pm 0.10$, where the uncertainty is dominated by the systematic uncertainty on the similarity between the kinematically weighted fake B_s^0 candidates and the signal.

8.2 Decay-time acceptance

In addition to the decay-time resolution, the decay-time of the B_s^0 mesons is distorted by several requirements which are applied in the signal selection described in Ch. 5. Such examples are: FDχ^2, IPχ^2, DIRA or vertex χ^2, where the efficiency of the selection criteria

[7] $\Theta(t) = \begin{cases} 0, & t < 0, \\ 1, & t \geq 0, \end{cases}$

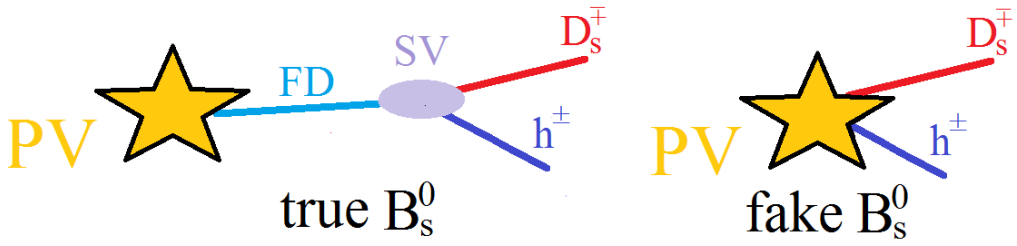

Figure 8.4: The left plot shows a true B_s^0 decay where the beauty meson flight before decaying. The right plot shows a fake B_s^0, with true decay-time equal to zero. The $D_s^{\mp}$ and $h^{\pm}$ mesons are created directly in the PV vertex.

depends on the decay-time. These effects are modelled by the decay-time acceptance $a(t)$. The decay rate equations given by Eq. 8.2 are additionally corrected to:

$$\frac{d\Gamma(t)^{\rm acc}_{\rm exp}}{dt} = \frac{d\Gamma(t)_{\rm exp}}{dt} \times a(t). \tag{8.5}$$

The decay-time acceptance of both $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates has to be correctly described for the fits to return unbiased results. In the case of $B_s^0 \to D_s^- \pi^+$, this can be accomplished by fixing Γ_s to the latest LHCb measurement and floating the acceptance parameters. In the case of $B_s^0 \to D_s^{\mp} K^{\pm}$, however, the acceptance parameters cannot be floated, because they are heavily correlated with the *CP* observables, and in particular any upper decay-time acceptance would be fully correlated with the D_f and $D_{\bar{f}}$ parameters. For this reason, the acceptance in the $B_s^0 \to D_s^{\mp} K^{\pm}$ fit is fixed to that found in the $B_s^0 \to D_s^- \pi^+$ data fit, corrected by the difference between the two channels observed in signal simulation. To obtain these differences, an acceptance function is fitted to both $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ signal simulation samples, where the candidates pass through the signal kinematic selection and are reweighted for data/simulation kinematic differences as well as for the effect of the PID requirement.

In all cases, the decay-time acceptance is described using splines, which can be implemented in an analytic way in the decay-time fit following the method presented in [105]. Splines are cubic polynomials, parametrised by so-called knots which determine their boundaries; n knots corresponds to n+2 base splines $v_i(t)$. Figure 8.5 shows an example of splines.

The decay-time acceptance is defined by 8 basic splines v_i, where $i = [1, ..., 8]$. The spline boundaries were chosen in an ad-hoc manner to produce a smooth acceptance shape, and placed at $t = \{0.5, 1.0, 1.5, 2.0, 3.0, 12.0\}$ ps for the nominal fit configuration. The value of 0.5 ps corresponds to two splines v_1 and v_2, whereas the value 12.0 ps denotes the knot position for v_7 and v_8. There are more knots at smaller decay-times where the function varies more rapidly, and fewer at high decay-times where flat behaviour is expected with a possible slow drop due to an upper acceptance effect from the VELO reconstruction. Due to normalisation one of the basic splines has to be fixed and $v_7 = 1$ is chosen. In addition, due to stabilise the last spline v_8 is also fixed using following formula:

$$v_N = v_{N-1} + \frac{v_{N-2} - v_{N-1}}{t_{N-2} - t_{N-1}} \times (t_N - t_{N-1}), \tag{8.6}$$

where t_{N-2} is the value of the knot position corresponding to the basic function v_{N-2} and, in this particular case, $N = 8$.

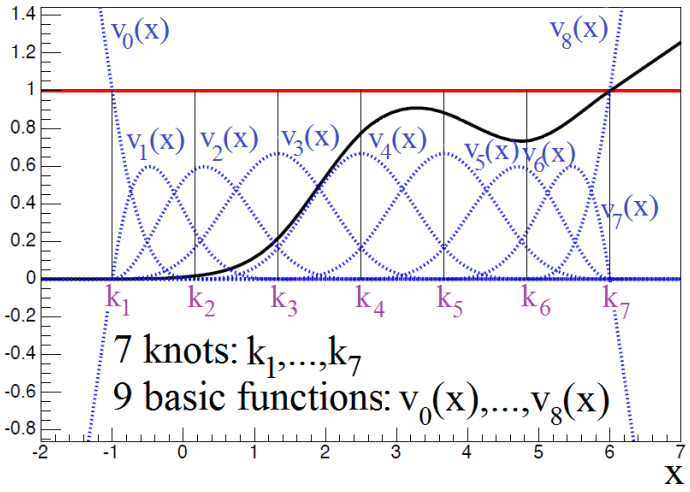

Figure 8.5: A graphical example of how a distribution can be decomposed into splines. Ref. [105].

Since the simulated candidates were produced without *CP* violation, their decay-time distribution is described by:

$$\frac{d\Gamma(t)^{\rm acc}_{\rm simulation}}{dt} = a(t) \times \left(\frac{1}{2}[e^{(-t/\tau_H)} + e^{(-t/\tau_L)}] \otimes R(t) \right), \tag{8.7}$$

where $\tau_H = 1.536875$ ($\tau_L = 1.407125$) ps are generated values, which denotes the decay-time of the heavier (lighter) component. The fits to the simulated events are shown in Fig. 8.6 and Tab. 8.1. Candidates corresponding to different D_s^- decay modes enter into the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ acceptance fits with weights corresponding to the yields observed in the relevant data fits.

The ratio of these two splines is given by the ratio of their coefficients, which is also given in Tab. 8.1. The correlation matrices are given in Tabs. H.2 ($B_s^0 \to D_s^- \pi^+$), H.3 ($B_s^0 \to D_s^\mp K^\pm$), and H.4 (ratios) in App. H. The $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ acceptances are significantly different, although all ratios are within 2σ of unity. A systematic is assigned due to imperfect knowledge of the decay-time acceptance.

Table 8.1: The fit parameters for the acceptance fit to the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ simulation samples, and their ratio.

Parameters	Fit to $B_s^0 \to D_s^- \pi^+$	Fit to $B_s^0 \to D_s^\mp K^\pm$	$B_s^0 \to D_s^\mp K^\pm / B_s^0 \to D_s^- \pi^+$
v_1	0.512 ± 0.008	0.498 ± 0.008	0.973 ± 0.022
v_2	0.745 ± 0.012	0.742 ± 0.013	0.996 ± 0.024
v_3	0.996 ± 0.015	0.981 ± 0.017	0.985 ± 0.023
v_4	1.131 ± 0.018	1.163 ± 0.020	1.028 ± 0.024
v_5	1.231 ± 0.017	1.242 ± 0.019	1.009 ± 0.021
v_6	1.227 ± 0.027	1.285 ± 0.030	1.047 ± 0.034

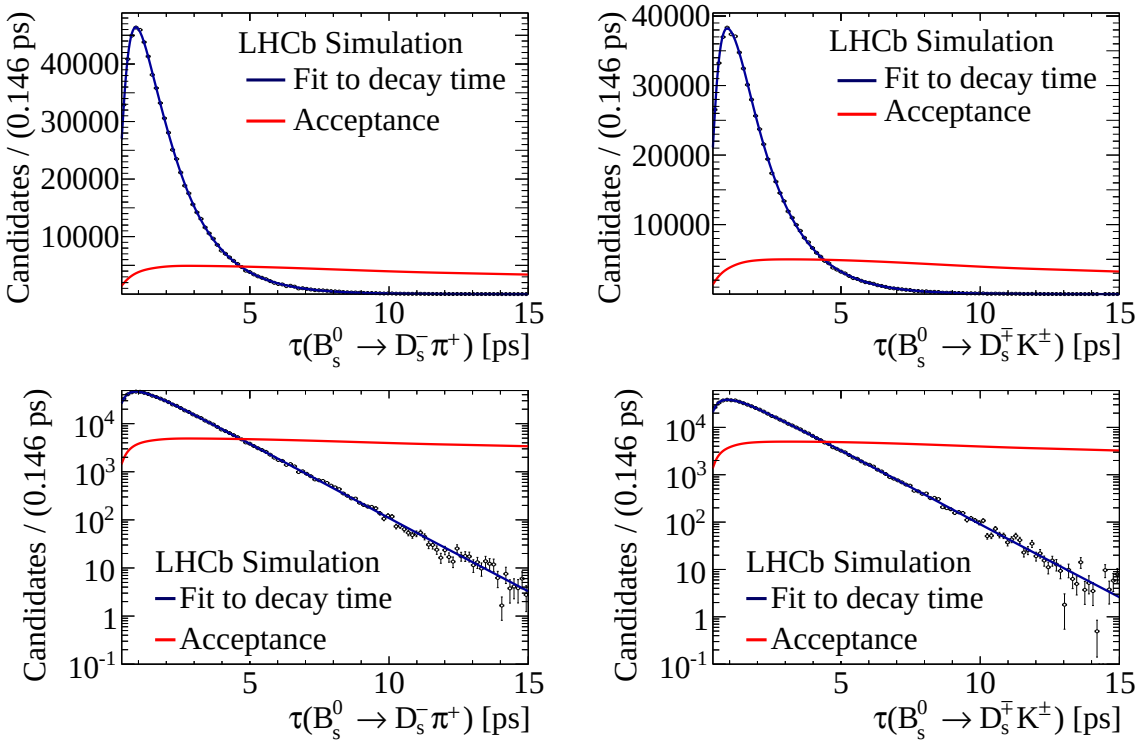

Figure 8.6: Decay-time acceptances for $B_s^0 \to D_s^- \pi^+$ (left) and $B_s^0 \to D_s^\mp K^\pm$ (right) simulation samples. At the bottom the fits are presented in logarithm scale. The acceptance function is shown in red, while the total fit curve given by Eq. 8.7 is blue.

Chapter 9

Decay-time fit using *sFit* approach

In order to measure the *CP* parameters given by Eq. 2.53, the decay-time fit, so-called *sFit*, is performed to both $B_s^0 \to D_s^{\mp} h^{\pm}$ modes. The *sFit* relies on the inputs described in all previous sections. In this chapter a general strategy of the fit is given, which is followed by a presentation of the obtained results as well as a description of how the fit is validated.

9.1 Production and detection asymmetries

In addition to detector effects such as finite tagging performance, decay-time resolution and acceptance, performing an unbiased time-dependent measurement requires detection and production asymmetries to be correctly modelled.

The *pp* interactions which occur at LHC are not charge symmetric, therefore a production asymmetry has to be considered as a source of systematic effect. The production asymmetry a^{prod} can be defined as:

$$a^{prod} = \frac{N_{\overline{P}}}{N_P} - 1, \tag{9.1}$$

where $N_{\overline{P}}$ (N_P) is the number of produced $\overline{P}$ (P) particles, for example $\overline{P} = \overline{B}_s^0$ and $P = B_s^0$. In case of beauty mesons, the production asymmetry is expected to be small with statistical error of 1% or less [106]. Thus, the signal asymmetry a^{prod} is fixed to zero and is validated in systematic studies discussed further in Sec. 10.1.

Since all final state particles in the $B_s^0 \to D_s^{\mp} h^{\pm}$ decay are kaon and pions their detection asymmetry has to be studied. This effect is defined as the asymmetry in the reconstruction efficiency of charge-conjugated states

$$a^{detector} = \frac{\epsilon_{\overline{f}}}{\epsilon_f} - 1, \tag{9.2}$$

where $\epsilon_{\overline{f}}$ (ϵ_f) is reconstruction efficiency of antiparticles $\overline{f}$ (particles f). The main sources of asymmetry occur due to: misalignment or inefficiencies in some subdetector modules, software biases or different interactions with the detector for charge-conjugate final states. The detection asymmetry in LHCb experiment have been studied in Ref. [107] and found to be $(+1.17 \pm 0.12)\%$ for the track reconstruction, $(-0.34 \pm 0.07)\%$ for L0 Hadron TOS, and $(+0.10 \pm 0.02)\%$ for non-TOS candidates. Since the signal sample is roughly half-half TOS and non-TOS, the numbers are added up to obtain a final detection asymmetry of $a^{detector} = (+1 \pm 0.5)\%$. The uncertainty of the asymmetry is inflated to account for

different kinematics between the modes. In the data fits the detection asymmetry is then constrained to this value and uncertainty and studied further as a systematic effect in Sec. 10.1.

9.2 General strategy

The signal time PDF is described using a `RooBDecay` PDF for both the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ channels. The fit is an unbinned maximum likelihood one. Theoretical decay rates given by Eq. 2.47 and Eq. 2.48 (and similarly for charge conjugate final states) have to be extended to take into account all detector effects such as: finite tagging performance and its asymmetry, decay-time resolution and acceptance as well as production and detection asymmetries. The set of parameters is given by:

$$p = \{C, D_f, D_{\bar{f}}, S_f, S_{\bar{f}}, \Gamma_s, \Delta\Gamma_s, \Delta m_s, v_i, p_0^{k,c}, p_1^{k,c}, \epsilon_c^{tag}, a_c^{tag}, a^{detector}, a^{prod}\} \tag{9.3}$$

with the five *CP* parameters C, D_f, $D_{\bar{f}}$, S_f, $S_{\bar{f}}$, the decay width Γ_s, the decay width difference $\Delta\Gamma_s$, the oscillation frequency Δm_s and the six decay-time acceptance parameters v_i, $i = 1, ..., 6$, the six tagging parameters $p_0^{k,c}$ and the six $p_1^{k,c}$ separate for initial state $k = \{B_s^0, \overline{B}_s^0\}$ and each tagging category $c = \{OS, SSK, OS + SSK\}$, the tagging efficiencies ϵ_c^{tag} and their asymmetries a_c^{tag} for each tagging category $c = \{OS, SSK, OS + SSK\}$, the detector asymmetry $a^{detector}$ and production asymmetry a^{prod}. The observables in the fit are: the candidate decay-time t, the combined tagging decision q_c, the charge of the bachelor particle q_f, the per-candidate time uncertainty σ_t, and the predicted combined mistag value ω_c. The *sWeights* are given as per-candidate weights to the fit. Since the predicted mistag is independent of the mass of the signal candidate (see Fig. 6.1), the predicted mistag is used as a per-candidate weight and "Punzi terms" [108] are ignored.

The decay-time fit uses both the OS cut based and the SSK neural net taggers. The tagging calibration parameters and uncertainties of the combination can be found in Sec.7.5. They are constrained in the fit to their central values and uncertainties using Gaussian constraints. The *sFit* uses per-candidate decay-time uncertainties with the signal $sWeighted$ template shown in Fig 8.3. The scale factor is fixed in the fit and varied as part of the systematic studies. Parameters which are fixed in the fit are

$$\Gamma_s = (0.661 \pm 0.007)\,\mathrm{ps}^{-1}\,, \tag{9.4}$$

$$\Delta\Gamma_s = (0.106 \pm 0.013)\,\mathrm{ps}^{-1}\,, \tag{9.5}$$

$$\rho(\Gamma_s, \Delta\Gamma_s) = -0.39. \tag{9.6}$$

They are taken from current LHCb measurements [109].

In this fit, the backgrounds are subtracted using the per-candidate *sWeights* previously extracted from the multidimensional fit described in Ch. 6. The *sWeights* are computed based on ten data samples (two magnet polarities and five D_s^- final states). The signal *sWeights* are shown in Fig. 9.1 for both the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ decays. High values correspond to signal-like candidates, while low values denote background-like candidates.

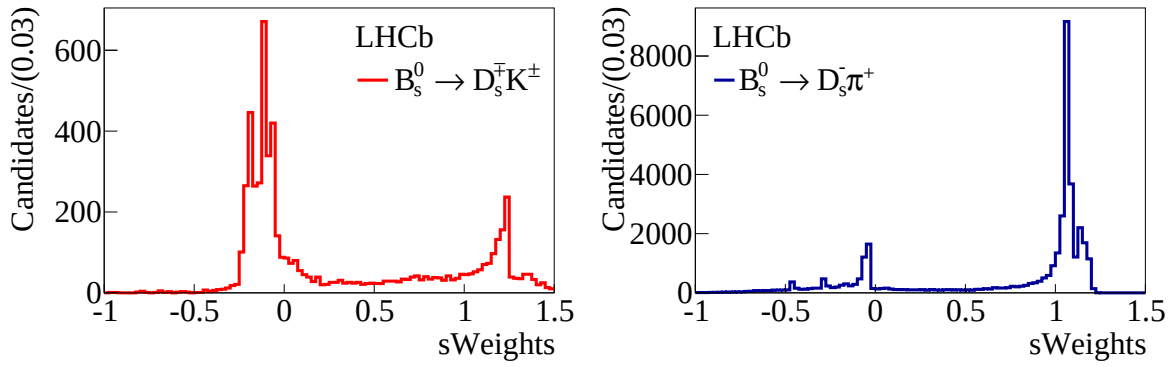

Figure 9.1: Signal *sWeights* distribution for the $B_s^0 \to D_s^{\mp} K^{\pm}$ (left) and $B_s^0 \to D_s^- \pi^+$ (right) decays.

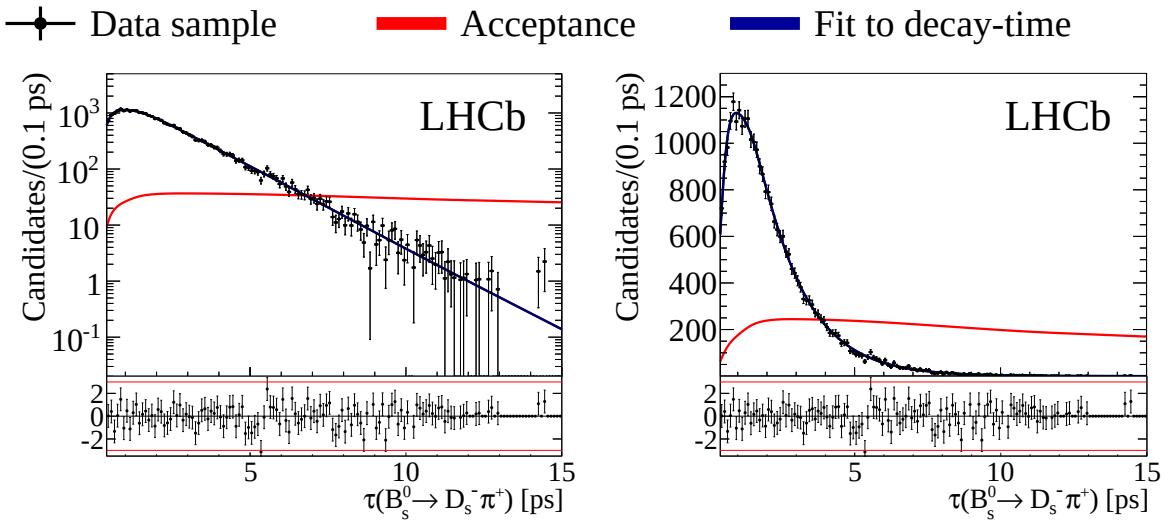

Figure 9.2: Decay-time *sFit* to the $B_s^0 \to D_s^- \pi^+$ data sample with logarithm (left) and linear (right) scale on the vertical axis. Different contributions to the fit are shown as coloured lines as described in the legend placed in the top.

9.3 Decay-time fit to $B_s^0 \to D_s^- \pi^+$ data

In addition to measuring the acceptance parameters, the correct measurement of Δm_s from $B_s^0 \to D_s^- \pi^+$ serves as a valuable control for the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay.

The combined signal sample for the $B_s^0 \to D_s^- \pi^+$ time fit contains both magnet polarities and all D_s^- final state samples. The signal $B_s^0 \to D_s^- \pi^+$ decay-time PDF is fitted to this sample, where Γ_s and $\Delta\Gamma_s$ are fixed to values from Eq. 9.4 and Eq. 9.5, respectively. In addition, all tagging parameters and asymmetries are fixed to their central values. Since $B_s^0 \to D_s^- \pi^+$ is flavour specific the *CP* parameters are fixed to $C = 1$, $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$. Thus the free parameters in the fit are Δm_s and the six parameters v_i, $i = 1, ..., 6$ describing the time acceptance function. The data distribution and the fit result are presented in Fig. 9.2. The fitted values are listed in Tab. 9.1, and the correlation matrix given in Tab. H.1 (App. H) for the acceptance fit. The fitted Δm_s agrees well with the existing LHCb measurement $\Delta m_s = 17.768\,\text{ps}^{-1}$ [4].

Table 9.1: Result of the *sFit* performed to the $B_s^0 \to D_s^- \pi^+$ data sample.

Parameter	Fitted value
Δm_s	$17.772 \pm 0.0215\,\mathrm{ps}^{-1}$
acceptance function	
v_1	0.459 ± 0.032
v_2	0.690 ± 0.052
v_3	0.885 ± 0.065
v_4	1.130 ± 0.081
v_5	1.223 ± 0.076
v_6	1.228 ± 0.121

9.4 Decay-time fit to $B_s^0 \to D_s^\mp K^\pm$ data

The signal $B_s^0 \to D_s^\mp K^\pm$ PDF is fitted to the combined data sample. In addition to the description from Sec. 9.2, the decay-time fit is executed with Δm_s fixed to the LHCb measurement value $\Delta m_s = (17.768 \pm 0.024)\,\mathrm{ps}^{-1}$ [4] and the acceptance parameters fixed to the values found in the $B_s^0 \to D_s^- \pi^+$ fit multiplied by the correction given in Tab. 8.1. The free parameters in the fit are the *CP* observables, which are limited to the fit range $[-4, 4]$, corresponding to about $\pm 10\sigma$ based on the expected uncertainties seen in the pseudo experiment studies. The fitted values are listed in Tab. 9.2, and the correlations between *CP* parameters are collected in Tab. 9.3. Finally the resulting PDF is shown in Fig. 9.3.

Table 9.2: Result of the *sFit* performed to the $B_s^0 \to D_s^\mp K^\pm$ data sample. All parameters other than the *CP* observables are constrained in the fit.

Parameter	Fitted value		
C	0.52 ± 0.25		
D_f	0.29 ± 0.42		
$D_{\bar{f}}$	0.14 ± 0.41		
S_f	-0.90 ± 0.31		
$S_{\bar{f}}$	-0.36 ± 0.34		
$a^{detector}$	0.010 ± 0.005		
	OS	SSK	OS+SSK
a_c^{tag}	-0.0028 ± 0.0016	0.0019 ± 0.0010	-0.0024 ± 0.0016
ϵ_c^{tag}	0.203 ± 0.002	0.291 ± 0.002	0.184 ± 0.002
$p_0^{B_s^0,c}$	0.366 ± 0.004	0.425 ± 0.007	0.339 ± 0.005
$p_0^{\overline{B}_s^0,c}$	0.377 ± 0.004	0.405 ± 0.011	0.339 ± 0.005
$p_1^{B_s^0,c}$	0.951 ± 0.040	0.999 ± 0.140	0.970 ± 0.036
$p_1^{\overline{B}_s^0,c}$	1.051 ± 0.040	0.992 ± 0.141	1.027 ± 0.037

Table 9.3: The correlation matrix of the $B_s^0 \to D_s^{\mp} K^{\pm}$ *CP* parameters for the *sFit*. Only the correlations between the *CP* parameters are shown, as the tagging efficiencies all have negligible correlations with the *CP* parameters anyway. As can be seen the only large correlation is the 50% between D_f and $D_{\bar{f}}$, which is understandable because they are both sensitive to the slowly varying hyperbolic terms induced by $\Delta\Gamma_s$.

Parameter	C	D_f	$D_{\bar{f}}$	S_f	$S_{\bar{f}}$
C	1.000	−0.071	−0.097	0.117	0.042
D_f	−0.071	1.000	0.500	0.044	−0.003
$D_{\bar{f}}$	−0.097	0.500	1.000	0.013	−0.005
S_f	0.117	0.044	0.013	1.000	−0.007
$S_{\bar{f}}$	0.042	−0.003	−0.005	−0.007	1.000

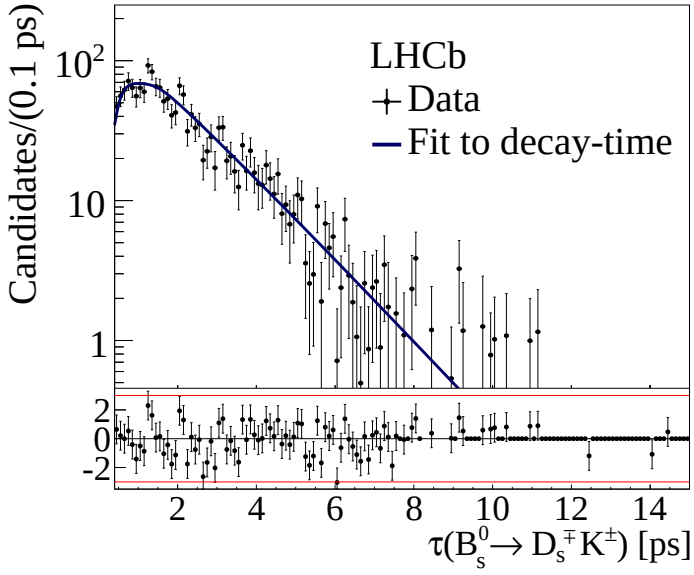

Figure 9.3: Decay-time *sFit* to the $B_s^0 \to D_s^{\mp} K^{\pm}$ data sample.

9.4.1 Folded asymmetries

Folded asymmetries refer to the procedure of plotting the difference between $B_s^0 \to D_s^+ K^-$ candidates which have been tagged as B_s^0 and $B_s^0 \to D_s^+ K^-$ candidates which have been tagged as $\overline{B}_s^0$ as a function of decauy-time, where the decay-time axis is "folded" by taking the modulo of the observed decay-time of every candidate with the oscillation frequency $2\pi/\Delta m_s$. The same plot is made for $B_s^0 \to D_s^- K^+$ candidates. The plotted asymmetry function is drawn in Fig. 9.4 using the central values of the *CP* observables, and is normalised using the expected dilution due to mistag and decay-time resolution.

9.5 Fit validation

Large ensembles of pseudo experiments are used to both validate the fit procedure and to evaluate systematic uncertainties, as introduced in Sec. 6.6. A standalone simulation generator creates pseudo experiment data sets, which are then fit. Unless specified otherwise, $r_{D_sK} = 0.37$, the CKM angle $\gamma = 70°$ and a strong phase of $\delta = 30°$ are used. Pseudo

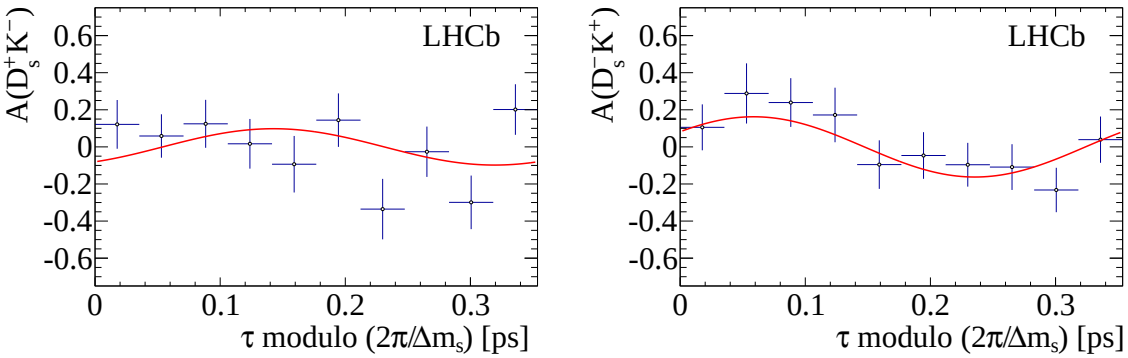

Figure 9.4: Folded asymmetry plots for the $B_s^0 \to D_s^{\mp} K^{\pm}$ sample.

experiments are generated with realistic *CP*-violation and oscillations for both signal and backgrounds, as well as with realistic production and detection asymmetries.

9.5.1 Nominal pseudo experiments studies

The *sFit* results on the nominal pseudo experiments are unbiased and indicate slightly overestimated statistical uncertainties, in particular for $S_{\bar{f}}$ which appears to be significantly overestimated. This may be related to the way in which RooFit [75] corrects uncertainties for the *sWeights*. The results are collected in Tab. 9.4 and shown in Fig. 9.5.

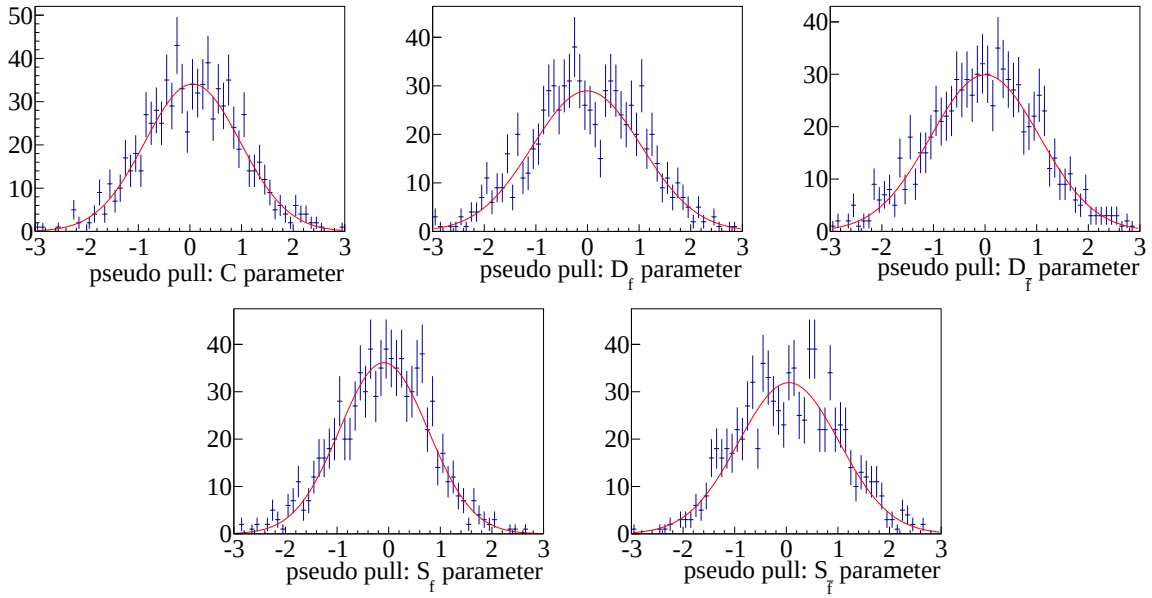

Figure 9.5: Pseudo pulls from pseudo experiments corresponding to *sFit* to the $B_s^0 \to D_s^{\mp} K^{\pm}$ sample. Clockwise from top left to right bottom: C, D_f, $D_{\bar{f}}$, S_f, $S_{\bar{f}}$ parameters.

9.5.2 Decay-time *sFit* cross-checks

To validate the obtained results several cross-checks are performed. Firstly, the data sample is split into subsamples according to following conditions: magnet polarities, trigger decision and BDTG response. In addition, the fits are performed to a sample with a tighter value of BDTG equal to 0.6. Next, the simulation samples are fitted. This is followed by

Table 9.4: Results of pseudo experiments corresponding to *sFit* to the $B_s^0 \to D_s^{\mp} K^{\pm}$ sample.

Parameter	μ of pseudo pull	σ of pseudo pull	average uncertainty
C	0.052 ± 0.034	0.925 ± 0.028	0.249 ± 0.001
D_f	-0.010 ± 0.040	1.056 ± 0.030	0.407 ± 0.001
$D_{\bar{f}}$	0.001 ± 0.039	1.054 ± 0.032	0.415 ± 0.001
S_f	-0.095 ± 0.032	0.860 ± 0.026	0.345 ± 0.001
$S_{\bar{f}}$	0.052 ± 0.037	0.976 ± 0.027	0.350 ± 0.001

the decay-time fit using different acceptance variations, without/with the kinematic fit described in Sec. 5.5, and with different type of resolution models. Finally, the results are verified using an alternative method of obtaining PIDK PDFs.

Splits

For each split subsample, all PDFs are redone under the appropriate conditions. The multidimensional part of the fit converged well in all cases. The $B_s^0 \to D_s^- \pi^+$ sample is split 40%(60%) for magnet up(down) and 42%(58%) for TOS(!TOS), while for $B_s^0 \to D_s^{\mp} K^{\pm}$ the percentages are 40%(60%) and 47%(53%), respectively. This is understandable because $B_s^0 \to D_s^{\mp} K^{\pm}$ contains more background, which is more likely to be !TOS.

Two samples: $BDTG \in (0.3, 0.9)$ ($BDTG > 0.9$) are selected, with a size 66% (34%) for $B_s^0 \to D_s^- \pi^+$ and 79%(21%) for $B_s^0 \to D_s^{\mp} K^{\pm}$, respectively. Again, the different signal-to-background ratios help to explain the smaller fraction of $B_s^0 \to D_s^{\mp} K^{\pm}$ in the more signal-like BDTG response bin.

The fitted Δm_s values for $B_s^0 \to D_s^- \pi^+$ splits are collected in Tab. 9.5. For the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay-time fitter the difference between the results obtained from magnet up (MU) and down (MD) samples, between TOS and !TOS as well as between $BDTG \in (0.3, 0.9)$ and $BDTG > 0.9$ are presented. Moreover the nominal result is compared with the obtained for $BDTG > 0.6$. All results are shown in Tab. 9.6. In addition for each split weighted comparisons with nominal result are performed, which are collected in Tab. 9.7.

The maximal difference for splits is 2.2σ in the BDTG split for C. However, the difference in $S_{\bar{f}}$ between the nominal result and $BDTG > 0.6$ fits is 0.345 ± 0.134. This cut removes about 10% of signal and specific backgrounds, and 60% of combinatorial background, from the data fit. To better understand the size of this deviation, pseudo experiment studies are used by removing an appropriate number of candidates from the nominal pseudo experiments and measuring by how much the pseudo experiments (which are essentially unbiased before and after the removal of these candidates) shifted. An equivalent or larger shift is found in approximately 2% of samples. This effect was studied further by tightening the decay-time cut on the B_s^0 from 0.4 ps to 0.5 ps, and it reduces to around a 2σ effect, with a difference of 0.236 ± 0.115 for $S_{\bar{f}}$, and smaller shifts for other parameters. In addition, the migration of per-candidate *sWeights* between the nominal result and > 0.6 fits is checked. In total about 40 candidates in the $BDTG > 0.6$ sample change from being signal to being background-like, and about 20 migrate in the other direction. Almost all these candidates have *sWeights* very close to zero. The influence of varying either fixed parameters or the acceptance was checked. In both cases none is found. Finally it is concluded that observed effect is a fluctuation, for which a relevant systematic is assigned.

Table 9.5: Fitted values of Δm_s for cross-check samples.

Cross check	Δm_s [ps^{-1}]
Nominal result	17.772 ± 0.0215
Magnet down	17.783 ± 0.0276
Magnet up	17.757 ± 0.0335
TOS	17.751 ± 0.0273
$!TOS$	17.810 ± 0.0345
$BDTG > 0.6$	17.771 ± 0.0218
$BDTG \in (0.3.0.9)$	17.813 ± 0.0442
$BDTG > 0.9$	17.759 ± 0.0247

Table 9.6: Difference of the fitted *CP* parameters for cross-check samples for the *sFit*. Note that the statistical uncertainty for $D_f/D_{\bar{f}}$ splits needs to be inflated by roughly 10% in order to take into account the uncorrelated uncertainties on the decay-time acceptance obtained from the relevant $B_s^0 \to D_s^- \pi^+$ fits.

Parameter	$MD - MU$	$!TOS - TOS$	$BDTG \in (0.3.0.9)$ $-BDTG > 0.9$	$Nominal$ $-BDTG > 0.6$
C	0.145 ± 0.575	0.582 ± 0.529	1.210 ± 0.544	-0.004 ± 0.078
D_f	0.311 ± 0.858	0.845 ± 0.838	-1.295 ± 0.989	-0.173 ± 0.192
$D_{\bar{f}}$	1.233 ± 0.872	1.640 ± 0.833	-1.278 ± 0.950	-0.280 ± 0.198
S_f	-0.234 ± 0.623	0.757 ± 0.638	-0.039 ± 0.610	-0.160 ± 0.100
$S_{\bar{f}}$	-0.072 ± 0.714	-0.642 ± 0.677	-0.039 ± 0.712	0.345 ± 0.134

Fitting full simulation sample

The impact of neglecting correlations between variables is checked by fitting to the official LHCb simulation samples. A cocktail of signal simulation $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ under the $B_s^0 \to D_s^{\mp} K^{\pm}$ hypothesis is created. Both modes have an applied PIDK cut on the bachelor particle. The size of the signal $B_s^0 \to D_s^{\mp} K^{\pm}$ sample is 5 or 20 times bigger than for the nominal fit, while $B_s^0 \to D_s^- \pi^+$ is mixed in with a statistic 7.5 times larger than for the nominal result to compensate for a lack of other specific backgrounds. Additionally, combinatorial background is generated using the nominal generator (due to a lack of sufficient statistics for the combinatorial in the official simulation samples) with a statistic 5 times bigger than the nominal one. For $B_s^0 \to D_s^{\mp} K^{\pm}$, the simulation samples were generated without *CP* violation, thus $C = 1$, $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$. Note that the sensitivities of the full simulation sample should not be used to draw conclusions about the sensitivity of the data sample, as the full simulation sample is generated with a very different value of $\Delta\Gamma_s$ and in addition has a different tagging power due to known data-simulation differences.

The cross-check sample are:

- signal $B_s^0 \to D_s^{\mp} K^{\pm}$ simulation 1: with statistic 5 times bigger than nominal,
- signal $B_s^0 \to D_s^{\mp} K^{\pm}$ simulation 2: with statistic 20 times bigger than nominal,
- signal $B_s^0 \to D_s^{\mp} K^{\pm}$ with $B_s^0 \to D_s^- \pi^+$,

Table 9.7: Difference of the fitted *CP* parameters between the average of cross check samples and nominal fit for the *sFit*. Note that the statistical uncertainty for $D_f/D_{\bar{f}}$ splits needs to be inflated by roughly 10% in order to take into account the uncorrelated uncertainties on the decay-time acceptance obtained from the relevant $B_s^0 \to D_s^- \pi^+$ fits.

Parameter	$MD + MU$ $-Nominal$	$!TOS + TOS$ $-Nominal$	$BDTG \in (0.3.0.9) + BDTG > 0.9$ $-Nominal$
C	0.014 ± 0.059	0.057 ± 0.021	-0.031 ± 0.013
D_f	0.072 ± 0.059	0.163 ± 0.099	0.153 ± 0.241
$D_{\bar{f}}$	0.119 ± 0.083	0.254 ± 0.087	0.190 ± 0.224
S_f	-0.001 ± 0.074	0.027 ± 0.052	-0.043 ± 0.132
$S_{\bar{f}}$	-0.010 ± 0.021	-0.015 ± 0.087	0.006 ± 0.076

- signal $B_s^0 \to D_s^{\mp} K^{\pm}$ with Combinatorial,
- signal $B_s^0 \to D_s^{\mp} K^{\pm}$ with Combinatorial and $B_s^0 \to D_s^- \pi^+$.

The results of the performed fits are collected in Tab. 9.8 and 9.9.

In addition the influence of the $BDTG$ cut on the signal simulation sample is checked. For this purpose the signal $B_s^0 \to D_s^{\mp} K^{\pm}$ sample is fitted with different $BDTG$ cuts:

- $BDTG \in (0.3, 0.6)$,
- $BDTG > 0.6$,
- $BDTG \in (0.3, 0.9)$,
- $BDTG > 0.9$.

The size of the sample does not depend on the $BDTG$ cut and is five times bigger than the nominal one. The lack of any observed bias supports the conclusion about seeing a fluctuation in the data set.

Table 9.8: *sFit* results to the full simulation samples, with five times more signal than in the data. The simulation samples were generated with $C = 1$, $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$.

Parameter	Signal 1 $(5 \times size)$	Signal with $B_s^0 \to D_s^- \pi^+$	Signal with Combinatorial	All contributions
C	1.020 ± 0.060	1.049 ± 0.065	1.017 ± 0.064	1.045 ± 0.067
D_f	-0.003 ± 0.268	0.058 ± 0.285	-0.065 ± 0.290	0.038 ± 0.301
$D_{\bar{f}}$	0.058 ± 0.268	0.135 ± 0.286	-0.062 ± 0.293	0.055 ± 0.303
S_f	-0.052 ± 0.093	0.043 ± 0.101	-0.048 ± 0.102	0.025 ± 0.107
$S_{\bar{f}}$	0.172 ± 0.091	0.182 ± 0.095	0.102 ± 0.100	0.141 ± 0.102

Acceptance variations

As described in Sec. 8.2 the position of knots is chosen in an ad-hoc manner. To verify this choice the difference between data results obtaining using different acceptance functions

Table 9.9: *sFit* results to the full simulation sample (twenty times the data sample) generated with $C = 1$, $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$.

Parameter	Signal 2 ($20 \times size$)
C	1.021 ± 0.031
D_f	0.003 ± 0.137
$D_{\bar{f}}$	-0.076 ± 0.138
S_f	0.081 ± 0.046
$S_{\bar{f}}$	0.030 ± 0.046

Table 9.10: *sFit* results to the signal simulation samples with a different cut on the $BDTG$ response. The simulation samples were generated with $C = 1$, $D_f = D_{\bar{f}} = S_f = S_{\bar{f}} = 0$.

Parameter	$BDTG \in (0.3, 0.6)$	$BDTG > 0.6$	$BDTG \in (0.3, 0.9)$	$BDTG > 0.9$
C	0.976 ± 0.071	0.988 ± 0.063	1.023 ± 0.066	0.992 ± 0.061
D_f	0.224 ± 0.472	-0.111 ± 0.278	-0.048 ± 0.360	-0.069 ± 0.257
$D_{\bar{f}}$	0.166 ± 0.491	-0.039 ± 0.271	0.015 ± 0.355	-0.004 ± 0.254
S_f	0.043 ± 0.104	-0.006 ± 0.092	0.203 ± 0.095	-0.122 ± 0.089
$S_{\bar{f}}$	0.095 ± 0.100	0.112 ± 0.096	0.072 ± 0.102	-0.158 ± 0.091

are studied. Firstly, the fit with knots which were placed at 0.25, 0.375, 0.5, 0.75, 1.0, 1.5, 2.0, 2.5, 3.0, 7.5, 12 ps is performed. The difference with respect to nominal result is shown in Tab. 9.11 for *sFit*. In addition, the influence of the position the first two knots in the acceptance function is verfified. For that purpose the decay-time fit is performed with several variations of knot positions:

- acceptance 1: 0.25, 0.6, 1.0, 2.0, 3.0, 12.0 ps,
- acceptance 2: 0.4, 0.6, 1.0, 2.0, 3.0, 12.0 ps,
- acceptance 3: 0.45, 0.75, 1.0, 2.0, 3.0,12.0 ps,
- acceptance 4: 0.5, 0.75, 1.0, 2.0, 3.0, 12.0 ps,
- acceptance 5: 0.5, 1.0, 1.5, 2.0, 3.0, 12.0 ps.

The differences of fitted *CP* parameters with respect to nominal results are shown in Tab. 9.12.

Verification of using kinematic fit

As mentioned in Sec. 5.5 and 8.1 the kinematic fit is used for computing the B_s^0 decay-time. The difference of fitted *CP* parameters for fitting with and without using kinematic fit is shown in Tab. 9.13.

Verification of decay-time resolution.

To validate whether the size of the shifts in data is similar to what is observed in pseudo experiments the decay-time fit is performed with a scale factor $S = 1.25$ and $S = 1.45$. In

Table 9.11: Difference of the fitted *CP* parameters for the nominal acceptance and the acceptance with double knots.

Parameter	Difference
C	-0.001 ± 0.015
D_f	-0.014 ± 0.034
$D_{\bar{f}}$	-0.014 ± 0.035
S_f	0.000 ± 0.011
$S_{\bar{f}}$	0.000 ± 0.019

Table 9.12: Difference of the fitted *CP* parameters with respect to the nominal result for several variations of acceptances.

Par.	Acceptance 1	Acceptance 2	Acceptance 3	Acceptance 4	Acceptance 5
C	0.000 ± 0.008	0.000 ± 0.010	0.000 ± 0.011	0.000 ± 0.012	0.001 ± 0.016
D_f	0.008 ± 0.026	-0.005 ± 0.022	-0.005 ± 0.021	-0.005 ± 0.021	0.018 ± 0.040
$D_{\bar{f}}$	0.007 ± 0.026	-0.005 ± 0.022	-0.005 ± 0.022	-0.005 ± 0.021	0.018 ± 0.041
S_f	0.000 ± 0.014	0.000 ± 0.016	0.000 ± 0.020	0.000 ± 0.022	0.000 ± 0.012
$S_{\bar{f}}$	0.000 ± 0.012	0.000 ± 0.014	0.000 ± 0.015	0.000 ± 0.015	-0.001 ± 0.022

addition, the influence of fitting with an average decay-time resolution modelled by a triple Gaussian (see App. G) is validated. Differences between nominal and fitted results are collected in Tab. 9.14. All shifts are compatible with expectations from pseudo experiments studies.

Verification of reweighting PIDK **PDFs.**

As described in Sec. 6.2 the PIDK PDFs are obtained by weighting the PID calibration samples to match the kinematics of signal or backgrounds. In the nominal fits two variables $\ln(p_T)$ and ln(nTracks) were used. As a cross-check fits are performed when $\ln(p_T)$ is replaced by $\ln(p)$. Difference of fitted *CP* parameters between nominal results and cross check fits are collected in Tab. 9.15.

Table 9.13: Difference of the fitted *CP* parameters for fitting with and without using the kinematic fit.

Parameter	Difference
C	-0.021 ± 0.033
D_f	-0.062 ± 0.058
$D_{\bar{f}}$	-0.025 ± 0.048
S_f	0.013 ± 0.032
$S_{\bar{f}}$	0.004 ± 0.042

Table 9.14: Difference of the fitted *CP* parameters between the nominal results and the different settings of decay-time resolution.

Parameter	Triple Gaussian	scale factor $S = 1.25$	scale factor $S = 1.45$
C	0.035 ± 0.054	0.027 ± 0.075	-0.020 ± 0.066
D_f	0.009 ± 0.016	0.000 ± 0.011	-0.002 ± 0.014
$D_{\bar{f}}$	0.021 ± 0.025	0.007 ± 0.019	0.003 ± 0.017
S_f	-0.045 ± 0.068	0.029 ± 0.087	-0.012 ± 0.081
$S_{\bar{f}}$	-0.048 ± 0.027	-0.019 ± 0.100	0.016 ± 0.088

Table 9.15: Difference of the fitted *CP* parameters between the nominal results and different reweighting PIDK PDFs.

Parameter	Difference
C	0.024 ± 0.056
D_f	0.023 ± 0.092
$D_{\bar{f}}$	0.042 ± 0.094
S_f	0.006 ± 0.035
$S_{\bar{f}}$	0.014 ± 0.077

Chapter 10

Studies of systematic uncertainties

Potential systematic uncertainties on the *CP* parameters estimated in the previous chapter are accounted for either by means of Gaussian constraints in the fit, or by means of test fits to nominal pseudo experiments. Several sources of systematics effects have to be taken into account, which are discussed in the next subsections. The results of these studies are summarised in Tab. 10.2.

10.1 Production and detection asymmetries

To validate the correct description of the production asymmetries, the pseudo experiments are generated with asymmetries and fitted back with different values. Specifically the production asymmetries of 1% for B_s^0 and B^0 mesons, and 3% for Λ_b^0 baryons are generated. Pseudo experiments are then fitted back with signal production asymmetries shifted by $\pm 3\%$. No change in the results is observed and hence the signal production asymmetry is set to zero in the nominal fit. No systematic uncertainty is assigned.

The constrained value of detector asymmetry $a^{detector} = 1\% \pm 0.5\%$ is validated by generating pseudo experiments in which the backgrounds are given realistic detection asymmetries (from 1% to 5% depending on the mode). No bias is observed in the obtained results.

10.2 Uncertainties due to fixed background yields

Potential systematic errors due to fixed background yields are evaluated by generating pseudo experiments with one value for these yields, and then fitting back with the yields fixed to double what was in the generator. No significant bias in the results is observed, thus no systematic uncertainty is assigned.

10.3 Tagging systematics

A prominent source of systematics is the limited knowledge of the tagging calibration parameters p_0 and p_1, introduced in Ch.7. Their uncertainty is incorporated into the nominal fits by means of Gaussian constraints, so that the corresponding systematics is already contained in the statistical errors reported by the fits. However, the effect of the constraints is negligible: the fits without the constraints return essentially the same

result. Since the calibration parameters of OS and SSK are independent, the resulting total systematic is reduced.

10.4 Uncorrelated systematics estimated from pseudo experiments

The uncorrelated systematic effects are due to the fixed value of Δm_s and the decay-time resolution scale factor. For these studies test fits are performed to the nominal pseudo experiment samples, in which the parameter in question (called μ) is varied up and down by its uncertainties. The systematic error is then given by:

$$\sigma_{\text{syst}} = \sqrt{\langle\mu\rangle^2 + \sigma^2_{\text{shifts}}}\,, \tag{10.1}$$

where $\langle\mu\rangle$ is the average shift observed in the test fits computed as $\mu = (\Delta_{\text{up}} - \Delta_{\text{down}})/2$, and σ_{shifts} is the width of the distribution. The resulting errors are given in Tab. 10.2 in the systematics summary section.

The systematic correlations of the *CP* parameters are computed based on the covariances of the dataset of the test fit results,

$$V_{ij} = \frac{1}{N-1}\sum_{k=1}^{N}(\mu_{ik} - \langle\mu_i\rangle)(\mu_{jk} - \langle\mu_j\rangle)\,, \tag{10.2}$$

where (i, j) run over the *CP* parameters $(C, S_f, S_{\bar{f}}, D_f, D_{\bar{f}})$, and k runs over the pseudo experiments. However, since a contribution to the error from a non-zero shift (the $\langle\mu\rangle$ term in Eq. (10.1)) is considered, the covariance has to be corrected accordingly,

$$V'_{ij} = \frac{V_{ij}}{\sqrt{V_{ii}V_{jj}}}\sigma_{\text{syst},i}\sigma_{\text{syst},j}\,. \tag{10.3}$$

To compute the total systematic correlations given in the summary section 10.7, all the covariances (from Eq. (10.3)) due to the different systematic sources are summed.

10.5 Correlated systematics: decay-time acceptance, Γ_s and $\Delta\Gamma_s$

The decay-time acceptance is determined from a fit to $B^0_s \to D^-_s\pi^+$ data, and is therefore correlated to Γ_s, which is in turn correlated to $\Delta\Gamma_s$ (see Eq. 9.6). For this reason these sources of systematic uncertainties are considered simultaneously. In addition, the time acceptance is corrected by the $B^0_s \to D^\mp_s K^\pm / B^0_s \to D^-_s\pi^+$ ratio as obtained from simulation, which is another source of correlated systematics and is considered in a similar way, as described at the end of this section.

The decay-time acceptance is described by a spline which has 6 parameters, v_i, $i = 1, ..., 6$, that are correlated, and fixed in the nominal fits. The associated systematic is assessed in the following way. The *CP* parameters are a function of the fixed parameters, whose exact form is unknown to us:

$$\vec{f}_{CP} \equiv \vec{f}_{CP}(\vec{v})\,, \tag{10.4}$$

where $\vec{f}_{CP} = (C, S_f, S_{\bar{f}}, D_f, D_{\bar{f}})$ and $\vec{v} = (v_1, \ldots, v_6, \Gamma_s, \Delta\Gamma_s)$. Then the systematic covariance of the CP parameters is given by linear error propagation,

$$V[\vec{f}_{CP}] = B^T \, V[\vec{v}] \, B \,, \tag{10.5}$$

where B is the Jacobian matrix of first derivatives,

$$[B]_{ij} = \frac{\partial f_i}{\partial v_j} \,. \tag{10.6}$$

To estimate the form of B test fits are used, in which each of the parameters v_i gets varied up and down by its error, as taken from Tab. 8.1 (to propagate errors from the $B_s^0 \to D_s^{\mp} K^{\pm} / B_s^0 \to D_s^- \pi^+$ ratio on simulation), from Tab. 9.1 (errors from the fit to $B_s^0 \to D_s^- \pi^+$ data), and from Eq. 9.6 (errors on Γ_s, $\Delta\Gamma_s$). Then the partial derivatives are estimated as

$$\frac{\partial f_i}{\partial v_j} \approx \frac{f_i(v_j + \sigma_{v_j}) - f_i(v_j - \sigma_{v_j})}{2\sigma_{v_j}} \,, \tag{10.7}$$

where the $f_i(v_j \pm \sigma_{v_j})$ are the shifts observed in the test fits. Eq. 10.5 is evaluted for each pseudo experiment, and the resulting covariance matrices averaged.

The acceptance parameters are determined from the fit to $B_s^0 \to D_s^- \pi^+$ data, where Γ_s controls the expected exponential slope. The acceptance will parametrise any difference between the observed and the expected slope, thus the acceptance systematic is strongly anti-correlated with the systematic uncertainty due to the fixed value of Γ_s. A subtlety are the exact values of the correlations of Γ_s and $\Delta\Gamma_s$ to the acceptance parameters. They are chosen in the following way. The correlations of Γ_s to the v_i are taken to be uniform $\rho = -0.7$. This is the smallest value which can be chosen requiring the correlation matrix $C[\vec{v}]$ to be positive-semidefinite, when $\Delta\Gamma_s$ is taken to be uncorrelated to the other parameters. Then, the correlation between Γ_s and $\Delta\Gamma_s$ is set to its nominal value (Eq. 9.6). The correlations of $\Delta\Gamma_s$ to the acceptance parameters is chosen to be uniform and as positive as possible without spoiling the positive-semidefiniteness of $C[\vec{v}]$. The resulting value is $\rho(\Delta\Gamma_s, v_i) = 0.4$, so that the correlation matrix is:

$$C[\vec{v}] = \left[\begin{array}{c|cc} C[v_1, \ldots, v_6] & -0.7 & 0.4 \\ \hline -0.7 & 1 & -0.39 \\ 0.4 & -0.39 & 1 \end{array}\right] . \tag{10.8}$$

The second source of decay-time acceptance systematic uncertainties is due to the $B_s^0 \to D_s^{\mp} K^{\pm} / B_s^0 \to D_s^- \pi^+$ ratios obtained from simulation. The above method is repeated for those (but without Γ_s and $\Delta\Gamma_s$ as they don't affect the simulation ratios), taking the central values of the ratios from Tab. 8.1 and their correlations from Tab. H.4. The resulting contributions are shown in Tabs. 10.1.

The effect of doubling the number of spline knots was checked, but negligible impact is found.

10.6 Sample splits

The sample splits introduced as cross checks in Sec. 9.5.2, e.g. for magnet polarity, showed only one discrepancy which is difficult to explain: that in the split by the BDTG response, into BDTG $\in (0.3, 0.9)$ and BDTG > 0.9 (Tab. 9.6). Both independent subsamples are in

Table 10.1: Decay-time acceptance systematic uncertainties in *sFit*, relative to the expected statistical errors from the toy study.

Parameter	C	S_f	$S_{\bar{f}}$	D_f	$D_{\bar{f}}$
due to $B_s^0 \to D_s^- \pi^+$ data fit (inc. Γ_s and $\Delta\Gamma_s$)	0.037	0.034	0.033	0.384	0.396
due to $B_s^0 \to D_s^\mp K^\pm / B_s^0 \to D_s^- \pi^+$ simulation ratios	0.021	0.019	0.019	0.186	0.186
total acceptance systematic uncertainty	0.043	0.039	0.038	0.427	0.437

Table 10.2: Total *sFit* systematic errors, relative to the statistical error. †The daggered contributions (Γ_s,$\Delta\Gamma_s$) are given separately only for comparison (see text).

Parameter	C	S_f	$S_{\bar{f}}$	D_f	$D_{\bar{f}}$
Δm_s	0.062	0.104	0.100	0.013	0.013
scale factor	0.104	0.092	0.096	0.004	0.004
$\Gamma_s^\dagger$	0.007	0.007	0.007	0.261	0.286
$\Delta\Gamma_s^\dagger$	0.043	0.039	0.038	0.384	0.385
acceptance, Γ_s, $\Delta\Gamma_s$	0.043	0.039	0.038	0.427	0.437
sample splits	0.124	0.072	0.071	0.000	0.000
total	0.179	0.161	0.160	0.427	0.437

acceptable agreement, but doubts are not completely ruled out. Therefore a systematic error is assigned based on the difference of the weighted average of the two subsamples, and the nominal result, see Tab. 9.7. A systematic relative to the expected statistical uncertainty of 12.4% for C, and 7.2% for S_f and $S_{\bar{f}}$ is assigned. These are relative to the expected statistical uncertainty. This represents a sizable systematic for C, S_f, and $S_{\bar{f}}$. A systematic for D_f and $D_{\bar{f}}$ is not assigned, as there the agreement with the nominal result is good. This systematic is included under the assumption that it is uncorrelated with any other considered source. The resulting numbers are given in Tabs. 10.2.

10.7 Systematics Summary

Total covariance matrix, taking into account the above mentioned systematic uncertainties is obtained by adding the covariance matrices from all sources. The resulting systematic errors are expressed relative to the expected statistical errors found in the pseudo experiment studies, see Tab. 9.4. Table 10.2 gives the errors. In these tables the separate contributions from Γ_s and $\Delta\Gamma_s$ are included to convey a feel of their relative importance. For this they are treated as uncorrelated systematics, as described in Sect. 10.4. There is benefit from accounting for their correlations, as their squared sum is larger than the combined contribution for the acceptance, Γ_s, and $\Delta\Gamma_s$. Table 10.3 shows the resulting systematic correlations.

Table 10.3: Total *sFit* systematic uncertainties correlations.

Parameter	C	S_f	$S_{\bar{f}}$	D_f	$D_{\bar{f}}$
C	1.00	-0.04	0.04	-0.18	-0.18
S_f	-0.04	1.00	-0.05	0.17	0.17
$S_{\bar{f}}$	0.04	-0.05	1.00	-0.16	-0.16
D_f	-0.18	0.17	-0.16	1.00	0.95
$D_{\bar{f}}$	-0.18	0.17	-0.16	0.95	1.00

Chapter 11

Determination of the CKM angle γ

The *CP* parameters obtained in Sec. 9.4 are used to determine the CKM angle γ from Eq. 2.53. In this chapter sensitivity studies on pseudo experiments are presented, which are followed by the CKM angle γ extraction. As shown in Fig. 4.2 the determination of the CKM angle γ can be performed using two different decay-time fits. In this chapter, the nominal *sFit* results are compared with an alternative approach called *cFit*.

11.1 Sensitivity on the CKM angle γ

From the results for the time-dependent *CP*-violation observables of the fits to the nominal pseudo experiments (see Sec. 9.5.1), the expected sensitivity on γ can be extracted. In these studies the GammaCombo package is used, which is LHCb software used to perform the CKM angle γ combination mentioned in Sec. 2.3.3. The strategy is to maximise the following likelihood

$$\mathcal{L}(\vec{\alpha}) = \exp\left(-\frac{1}{2}\left(\vec{A}(\vec{\alpha}) - \vec{A}_{\text{obs}}\right)^T V^{-1}\left(\vec{A}(\vec{\alpha}) - \vec{A}_{\text{obs}}\right)\right) \quad , \tag{11.1}$$

where $\vec{\alpha} = (\gamma, \phi_s, r_{D_sK}, \delta)^T$ is the vector of the physics parameters, $\vec{A}$ is the vector of observables expressed through Eq. 2.53, $\vec{A}_{\text{obs}}$ is the vector of the measured *CP*-violation observables and V is the experimental (statistical and systematic) covariance matrix. Confidence intervals are computed by evaluating the test statistics $\Delta\chi^2 = \chi^2(\vec{\alpha}'_{\text{min}}) - \chi^2(\vec{\alpha}_{\text{min}})$, where $\chi^2(\vec{\alpha}) = -2\ln\mathcal{L}(\vec{\alpha})$, in a frequentest way following Ref. [110]. Here, $\vec{\alpha}_{\text{min}}$ denotes the global maximum of Eq. 11.1, and $\vec{\alpha}_{\text{min}}$ is the conditional maximum when the parameter of interest is fixed.

In order to get reasonable results, the fit results where the initial fit (*cFit*, *sFit* or the multidimensional fit) doesn't converge are excluded. In addition, to prevent the CKM angle γ extraction from failing, the following requirements are applied: the relative errors of the *CP* observables must be greater than 1%, the absolute errors of the observables must be less than 10, the resulting one-sided errors of γ must be smaller than 300°, and they are not allowed to be too asymmetric: $(\sigma_\gamma^+ - \sigma_\gamma^-)/(\sigma_\gamma^+ + \sigma_\gamma^-) < 0.8$. Furthermore, if the confidence level calculation returns an interval limit which is one of the scan boundaries ($\gamma = 0°$ or $\gamma = 360°$), a second scan around the best minimum in a range of $\pm 85°$ is performed. If both scans return interval limits which are scan boundaries, then this specific result of the pseudo experiment is neglected.

The results of the performed pseudo experiment studies is shown in Fig. 11.1. The Gaussian fits to the pull distributions yield $\mu = (-0.13 \pm 0.04) \times 10^{-2}$, $\sigma = 1.17 \pm 0.03$ for

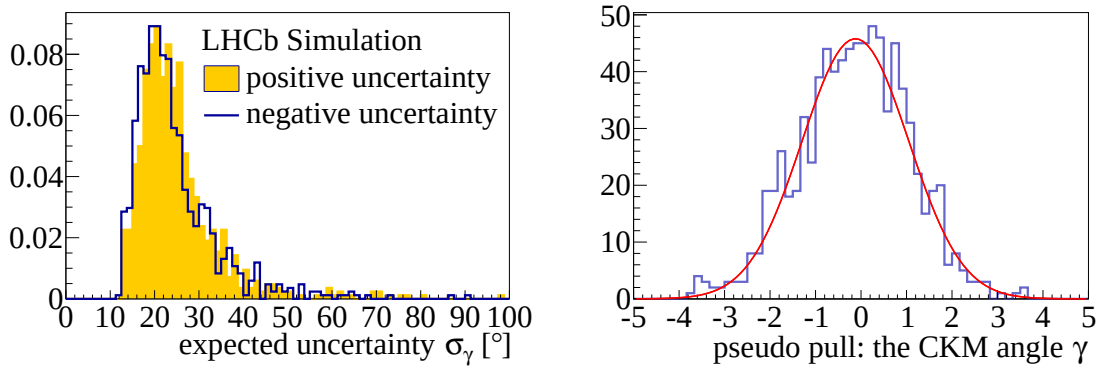

Figure 11.1: Left: the expected positive and negative uncertainties on the CKM angle γ. Right: the pseudo pulls evaluated for the CKM angle γ using the *CP* parameters obtained from the nominal pseudo experiments with γ generated at 70°.

the mean and width. The bias is within the sub-percent level of the γ value and error. The widths of the pulls is explicable, because the method is expected to undercover uncertainty. Overall, a sensitivity on the CKM angle γ of about $\sigma_\gamma \approx 24°$ (statistical uncertainty only) is expected, which is calculated as the mean of the two expectation values from the positive and negative uncertainty distributions.

11.2 Interpretation

The measurement of the *CP*-sensitive parameters is interpreted in terms of $\gamma - 2\beta_s$ and subsequently γ. The value β_s is constrained to the measurement from $B_s^0 \to J/\psi\, hh$ decays, $\phi_s = (0.01 \pm 0.07(\text{stat}) \pm 0.01(\text{syst}))\,\text{rad}$ [109]. Assuming no penguin pollution and no beyond SM contribution in these decays, $\phi_s = -2\beta_s$. The resulting confidence intervals are given in Tab. 11.1. Figures 11.2 and 11.3 show the p-value curves [111] for γ, δ, and r_{D_sK}, and Fig. 11.4 shows two-dimensional contours of the profile likelihood $\mathcal{L}(\vec{\alpha}'_{\text{min}})$.

Table 11.1: The confidence intervals for the CKM angle γ, strong phase δ and amplitudes ratio r_{D_sK} at 68% C.L. All errors (stat. and syst.) are included. The intervals for the angles are expressed modulo 180°.

Parameter	*sFit*	*cFit*
γ	$(113^{+30}_{-44})°$	$(115^{+28}_{-43})°$
δ	$(1^{+21}_{-21})°$	$(3^{+19}_{-20})°$
r_{D_sK}	$0.48^{+0.18}_{-0.16}$	$0.53^{+0.17}_{-0.16}$

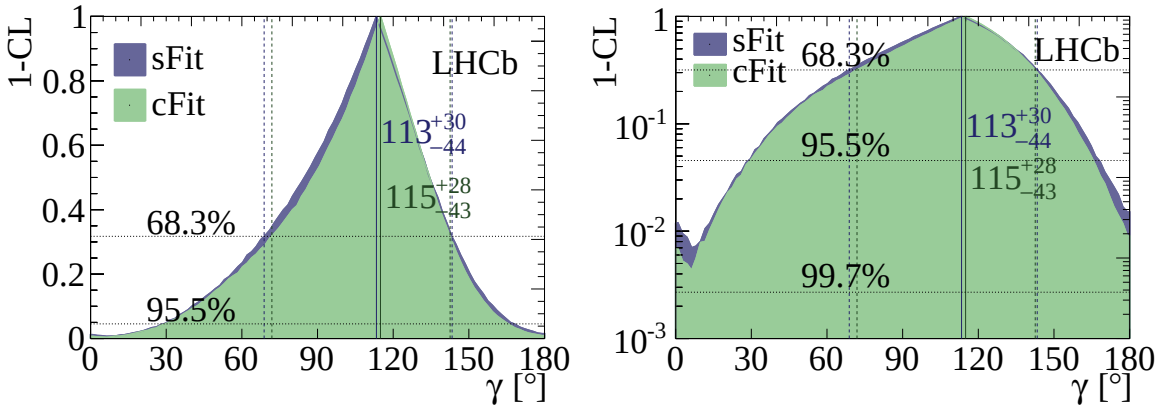

Figure 11.2: The $1-\mathrm{CL}$ contours for the CKM angle γ, in linear (left) and logarithmic scale (right). Green: *cFit*, blue: *sFit*. All errors, statistical and systematic, are included.

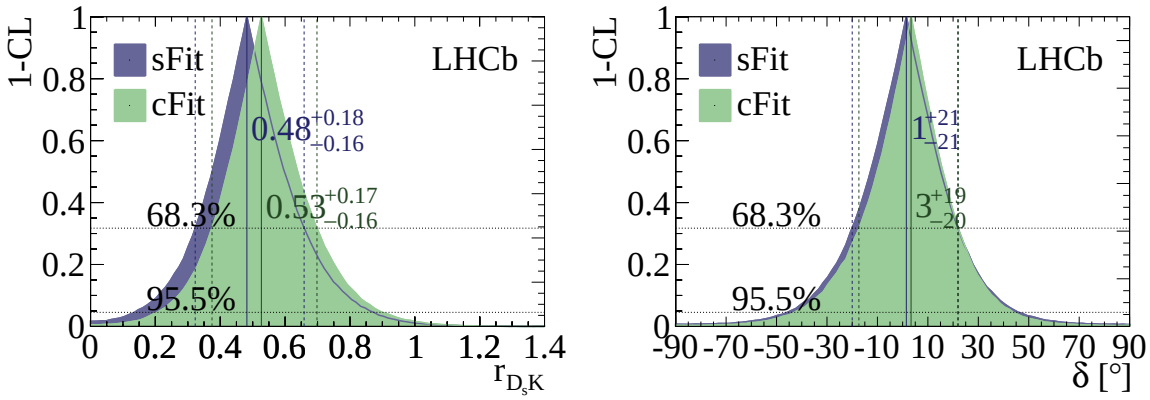

Figure 11.3: The $1-\mathrm{CL}$ contours for r_{D_sK} (left) and δ (right) parameters. Green: *cFit*, blue: *sFit*. All errors, statistical and systematic, are included.

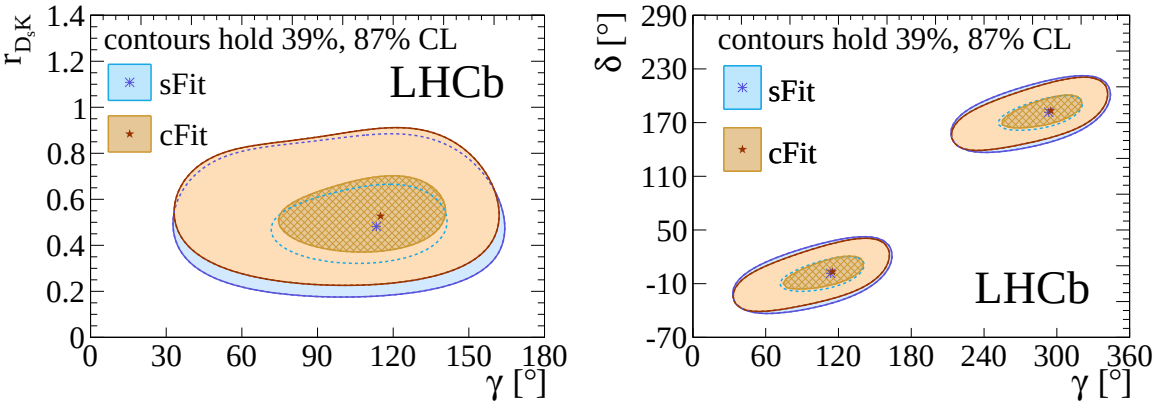

Figure 11.4: The profile likelihood contours of the CKM angle γ vs. r_{D_sK} parameter, and the CKM angle γ vs. δ parameter. The contours are the $n\sigma$ profile likelihood contours, where $\Delta\chi^2 = n^2$ with $n = 1, 2$. The markers denote the best-fit values. Orange: *cFit*, blue: *sFit*. All errors, statistical and systematic, are included.

Chapter 12

Conclusion

The excellent performance of the LHCb detector has made it possible to perform the world-first measurement of the CKM angle γ using $B_s^0 \to D_s^{\mp} K^{\pm}$ decays, one of the key benchmark measurement in the roadmap of the LHCb experiment in 2009 [112].

The study has been performed based on a dataset corresponding to 1.0 fb^{-1} recorded in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ in 2011. The *CP*-violation observables are found to be:

$$\begin{aligned} C &= 0.52 \pm 0.25 \pm 0.04\,, \\ D_f &= 0.29 \pm 0.42 \pm 0.17\,, \\ D_{\bar{f}} &= 0.14 \pm 0.41 \pm 0.18\,, \\ S_f &= -0.90 \pm 0.31 \pm 0.06\,, \\ S_{\bar{f}} &= -0.36 \pm 0.34 \pm 0.06\,, \end{aligned}$$

where the first (second) uncertainties are statistical (systematic).

Preliminary results of *CP* parameters were already published and the author of thesis is one of the contact authors of the conference contribution given in [113]. Unfortunately the statistical uncertainties were too pronounced to allow for a relevant measurement of the CKM angle γ. Since then, the development of the multidimensional fit improved the statistical uncertainty by about 25-35 % depending on the observable. This kind of approach has been used for the first time in the LHCb experiment. In addition, it is also the first time in the LHCb experiment that the *sPlot* technique has been successfully applied in an environment with so much background resulting in a perfect agreement with standard *cFit* method.

The *CP* observables are used to perform the first measurement of the CKM angle γ, the strong phase δ and amplitudes ratio r_{D_sK} in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. The resulting confidence intervals are, at 68% CL,

$$\begin{aligned} \gamma &= (113^{+30}_{-44})^{\circ}\,, \\ \delta &= (1^{+21}_{-21})^{\circ}\,, \\ r_{D_sK} &= 0.48^{+0.18}_{-0.16}\,, \end{aligned}$$

where the error contains both statistical and systematic uncertainties.

Time-dependent measurements using $B^0 \to D^{(*)\mp}\pi^{\pm}$ decays were performed by both the BaBar [47, 48] and the Belle [49, 50] collaborations. The study shown in this thesis, is the first measurement of the CKM angle γ in a time dependent analysis of tree decays, that excludes the *CP*-conservation hypothesis at more than $\sim 2.5\sigma$. In addition, as predicted,

the amplitudes ratio for these decay r_{D_sK} is an order of magnitude larger than that for $B^0 \to D^{(*)\mp}\pi^{\pm}$ $r_{D^{(*)}\pi} \approx 0.02$ [51], which makes a measurement using the $B_s^0 \to D_s^{\mp}K^{\pm}$ decay more sensitive to *CP* violation.

The value of strong phase δ is found to be close to 0, in agreement with theoretical predictions based on factorisation [46].

The CKM angle γ result agrees within one standard deviation with the world-average: $\gamma = (73.2^{+6.3}_{-7.0})^\circ$ [23]. In addition, the results have been included in the nominal LHCb combination of parameter in question resulting in $\gamma = (74.6^{+8.4}_{-9.2})^\circ$ [30].

Since the presented measurement is one of the most challenging analysis performed by the LHCb collaboration, it was not possible to study the full data collected in 2010-2012. The additional 2.0 fb^{-1} recorded in *pp* collisions at $\sqrt{s} = 8\,\text{TeV}$ in 2012 are currently being analysed and is not expected that measurement will be systematically limited in the near future. Assuming 5 and 50 fb^{-1} recorded data for the LHCb experiment in 2015-2018 and Upgrade runs, the predicted statistical sensitivity for measurements are 11° and 2.0° [114], respectively.

A possible extension of this analysis is a time-dependent analysis using the $B_s^0 \to D_s^{*\mp}K^{\pm}$ decay, whose branching fraction has been published [115] by the LHCb collaboration. The author of this thesis contributed in an essential way to the mentioned analysis. The obtained yield is promising and further developments such as using the multidimensional fit can make it possible to perform a standalone measurement of the CKM angle γ in the near future.

The results of thesis have been published in the Journal of High Energy Physics [116]. In addition, they have been presented by author of thesis at 37th International Conference on High Energy Physics [117, 118], where they have been highlighted as one of the most valuable new results in flavour physics.

Appendices

A Primary vertex reconstruction for 2015 data

The primary vertex (PV) algorithm begins with a set of input tracks. In 2010-2012 data taking two algorithms were in use: online and offline. Since the time budget in HLT1 was limited, the online reconstruction was based on VELO unfitted tracks, which didn't have an accurate covariance matrix. The most precise results were obtained by the second, offline, reconstruction based on Kalman fitted [71] tracks. The online (offline) algorithm was executed in Moore (Brunel) (see Fig. 3.15).

The algorithms have to be adapted for the new running conditions which begin in 2015. Thanks to developments in the trigger, the time budget is increased and it is possible to perform a Kalman fit in HLT1. Therefore the aim of these developments is to make the online algorithm as similar as possible to the offline approach. For this purpose the new 2015 online algorithm uses Kalman fitted VELO tracks.

The agreement between algorithms can be quantified by the decorrelation variable given by:

$$\frac{A_{OFF} - A_{ON}}{\sigma_{OFF}}, \tag{1}$$

where $A = \{x, y, z\}$ denotes coordinates, $A_{OFF}(A_{ON})$ is position of PVs found by the offline (online) algorithm and σ_{OFF} is the position uncertainty for the offline approach. The obtained distribution is parametrized by a triple Gaussian:

$$PDF_{tG}(\mu, \sigma_1, \sigma_2, \sigma_3, f_1, f_2; X) = f_1 G(\mu, \sigma_1; X) + f_2 G(\mu, \sigma_2; X) + (1 - f_1 - f_2) G(\mu, \sigma_3; X), \tag{2}$$

where $G(\mu, \sigma_i; X)$ denotes a single Gaussian distribution with common mean μ and width $\sigma = \sigma_i$ and X is an observable. The average width $\bar{\sigma}$ is defined as:

$$\overline{\sigma}^{tG} = \sqrt{f_1\sigma_1^2 + (1 - f_1) f_2 \sigma_2^2 + (1 - f_1)(1 - f_2)\sigma_3^2}. \tag{3}$$

Finally, the difference d between approaches is described by:

$$d = \sqrt{\overline{\sigma}^{tG} + \sigma_N} - 1, \tag{4}$$

where σ_N denotes the nominal value equals to 1. The initial difference between algorithms is 64[%] (58[%]) for the x(z) coordinate. Since the performance for y axis is similar to x, it is not quoted. The agreement is improved by using Kalman fitted Velo tracks for the so-called online 2015 algorithm, where the results are 21[%] and 16[%] for the x and z coordinates, respectively. This difference can be reduced even more by including Long tracks with a transverse momentum above 500 MeV/c. However, the final choice about including Long tracks can be made only after taking first 13 TeV data. As a baseline the new, improved algorithm uses only Kalman fitted Velo tracks and it is considered in further studies as the nominal online approach for 2015 data. The summary of decorrelation variables is collected in Tab. A.1.

A.1 Performance

The quality of the primary vertex algorithm can be studied in terms of its resolution and pseudo pulls. Again, the results for the y coordinate are not quoted. All studies are based on a $b\bar{b}$-inclusive simulation sample in nominal 2015 conditions.

Table A.1: The differences between the offline and online primary vertex algorithms.

Parameter	online 2010-2012 unfitted Velo tracks	online 2015 Kalman fitted Velo tracks	online 2015 Kalman fitted Velo + Long $p_T > 500$ MeV/c
$\overline{\sigma}_x^{tG}$	1.294	0.679	0.414
μ_x	0.007	-0.002	-0.001
d [%]	64	21	8.4
$\overline{\sigma}_z^{tG}$	1.123	0.599	0.478
μ_z	-0.002	0.004	0.0003
d [%]	58	16	10

Global resolution

The global resolution is determined from the distribution of the difference between the fitted and generated position of the primary vertex, ΔA, where A denotes a given coordinate. The resolution is obtained by fitting a triple Gaussian given by Eq. 2 and taking its averaged width as the parameter in question. The results are collected in Tab. A.2 and an example fit for the online 2015 algorithm is presented in Fig. A.1. The resolution for the z axis is compatible within uncertainty for all considered approaches. At the same time, the results obtained for the x axis is worse by (2-3)% with respect to offline algorithm. However, taking into account size of the global resolution for the z coordinate, it is concluded that results are in good agreement.

Table A.2: The global resolution for the online 2015 and offline primary vertex reconstructions.

Parameter [mm]	online 2015	offline
Δx	0.0205 ± 0.0001	0.0198 ± 0.0001
Δz	0.1432 ± 0.0006	0.1433 ± 0.0006

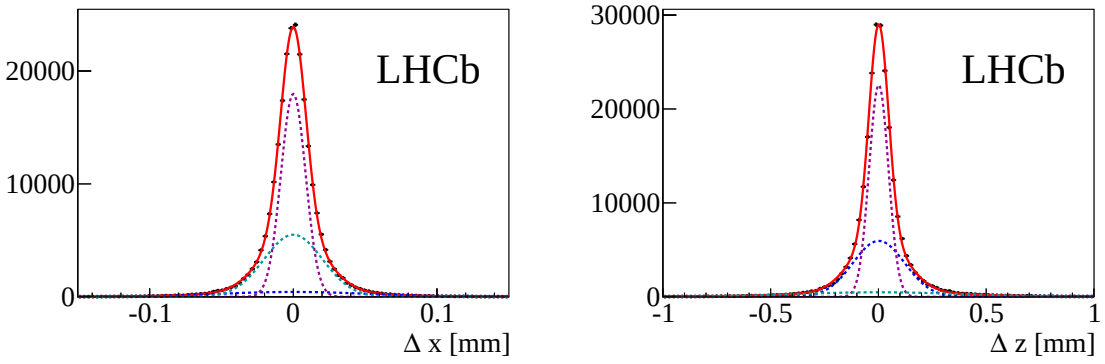

Figure A.1: The resolution obtained in the $b\bar{b}$-inclusive sample for the online 2015 algorithm. Left: x coordinate, right: z coordinate. The solid line corresponds to a triple Gaussian fit, while the dashed lines denote its respective components.

It has to be noted that for all VELO tracks used in the online 2015 algorithm, the transverse momentum is set to an average value of 400 MeV/c. This approximated

magnitude is used to perform the Kalman fit and obtain the correction for multiple scattering, thus its influence in the global resolution is studied. Figure A.2 shows the global resolution as a function of average p_T and the minimum is found to be at 300 MeV/c.

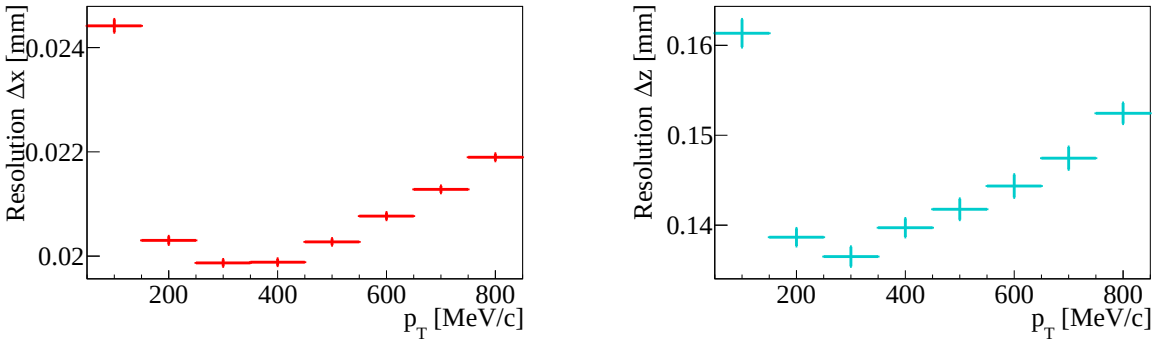

Figure A.2: Global resolution as a function of an average track's transverse momentum obtained based on the $b\bar{b}$-inclusive simulation sample for the online 2015 algorithm. Left: x coordinate, right: z coordinate.

Pseudo pulls

The pseudo pulls for the primary vertex are given by the global resolution divided by reconstructed position uncertainty. This distribution is parametrised by a double Gaussian:

$$PDF_{dG}(\mu, \sigma_1, \sigma_2, f_1; P) = f_1 G(\mu, \sigma_1; P) + (f_1 - 1) G(\mu, \sigma_2; P), \tag{5}$$

where $G(\mu, \sigma_i; t)$ denotes a single Gaussian distribution with common mean μ and width $\sigma = \sigma_i$. The P denotes a pseudo pull observable. The average width $\overline{\sigma}^{dG}$ is defined as:

$$\overline{\sigma}^{dG} = \sqrt{f_1 \sigma_1^2 + (1 - f_1)\sigma_2^2}. \tag{6}$$

An average width $\overline{\sigma}^{dG} = 1$ and $\mu = 0$ are signatures of unbiased results. The obtained results are collected in Tab. A.3, while the fits are shown for the online 2015 algorithm in Fig. A.3. The mean μ_x for both online 2015 and offline approaches is consistent with zero within uncertainty. A small bias is observed for z coordinate, where the results for the online 2015 algorithm are less shifted. In addition, the average widths for the online 2015 describe uncertainties accurately, while the offline uncertainties are underestimated.

Table A.3: The pseudo pulls for the online 2015 and offline primary vertex algorithms.

Parameter	online 2015	offline
μ_x	0.001 ± 0.002	-0.001 ± 0.002
$\overline{\sigma}_x^{dG}$	1.04± 0.02	1.17 ± 0.04
μ_z	0.030± 0.002	0.038 ± 0.002
$\overline{\sigma}_z^{dG}$	1.06± 0.02	1.13 ± 0.02

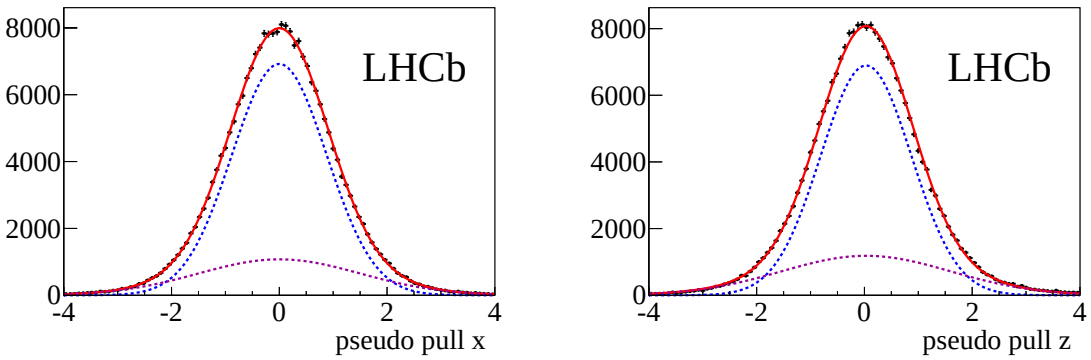

Figure A.3: The pseudo pull obtained in the $b\bar{b}$-inclusive sample for the online 2015 algorithm. Left: x coordinate, right: z coordinate. The solid line corresponds to a double Gaussian fit, while the dashed lines denote its respective components.

Track's impact parameter

Another quantity which has to be considered is the track's impact parameter which, as described in Sec. 5.1, is defined as the minimum perpendicular distance between the reconstructed track and the primary vertex. The magnitude of the impact parameter is calculated for all tracks from the best container. Figure A.4 shows the obtained results which are split according to the transverse momentum of the particle associated to the track: below and above 1 GeV/c. This represents 85% and 15% of tracks in the container, respectively. In the first case, below 1 GeV/c, the distributions are in good agreement for various algorithms. A discrepancy is seen for tracks above 1 GeV/c, where the impact parameter calculated with respect to online 2015 primary vertices is shifted upwards. Most likely, it can be explained by the influence of setting an average p_T to all VELO tracks described in Sec. A.1, where the set value is much less than 1 GeV/c and thus the position of the reconstructed primary vertex is slightly shifted. It is worth noting that this observed discrepancy is removed by adding Long tracks with $p_T > 500$ MeV/c. To study this effect the Landau distribution [99] is fitted with mean μ_L and width σ_L, the results are listed in Tab. A.4. As can be seen the Landau's widths are in good agreement, whereas the mean is shifted by 9% upwards. Since the mean can be interpreted as the expected resolution, this implies 9% worse resolution for 15% of tracks.

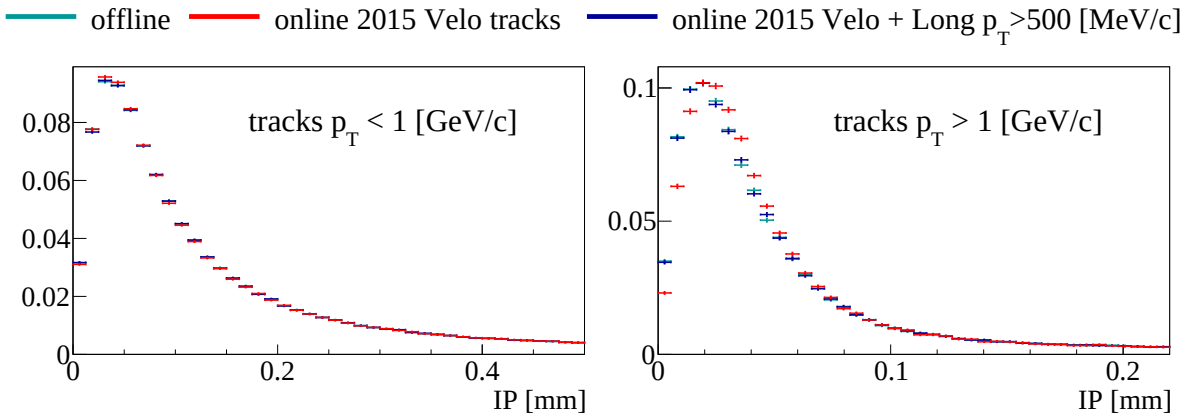

Figure A.4: The impact parameter calculated with respect to the primary vertex. Left: tracks with $p_T < 1$ GeV/c, right: $p_T > 1$ GeV/c.

Table A.4: Result of fitting a Landau distribution to the impact parameter calculated with respect to the online 2015 and offline primary vertices for track with $p_T > 1$ GeV/c.

Parameter	online 2015	offline
σ_L	0.0097 ± 0.0004	0.0096 ± 0.0004
μ_L	0.0228 ± 0.0006	0.0208 ± 0.0006

Influence for the decay-time resolution

A possible influence on the decay-time resolution is studied based on the signal $B_s^0 \to \phi^0\phi^0$ simulation sample, generated with the nominal 2015 conditions. The time resolution is defined by:

$$\Delta t = t_{rec} - t_{true}, \tag{7}$$

where t_{rec} is the reconstructed decay-time and t_{true} corresponds to the true decay-time generated in the simulation sample. The resulting distribution is parametrised by a triple Gaussian given by Eq. 2. The results are collected in Tab. A.5 and presented in Fig. A.5. The average of the decay-time resolution is in good agreement. No influence of changing the primary vertex reconstruction is found.

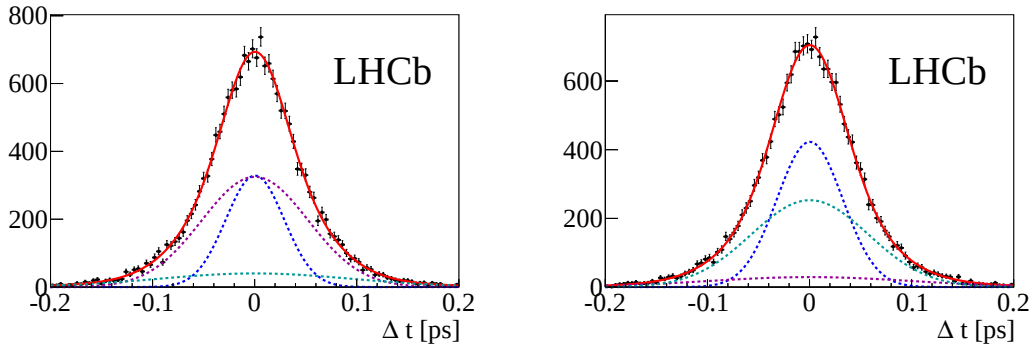

Figure A.5: Triple Gaussian time resolution fitted to the $B_s^0 \to \phi^0\phi^0$ simulation sample. The solid curve corresponds to the full fit, while the dashed curves are components. Left: online 2015 reconstruction, Right: offline reconstruction.

Table A.5: The fit parameters for the decay-time resolution.

Parameters	online 2015	offline
μ [ps]	0.00031 ± 0.00035	0.00033 ± 0.00035
f_1	0.23 ± 0.18	0.31 ± 0.10
f_2	0.64 ± 0.09	0.82 ± 0.13
σ_1 [ps]	0.028 ± 0.005	0.028 ± 0.003
σ_2 [ps]	0.044 ± 0.008	0.052 ± 0.006
σ_3 [ps]	0.077 ± 0.007	0.095 ± 0.025
$\bar{\sigma}^{tG}$ [ps]	0.0528 ± 0.007	0.0536 ± 0.009

A.2 Optimisation

The reconstruction algorithm is optimised in order to maximise the efficiency and minimise the fake rate. As a cross-check studies have been performed for the $c\bar{c}$-inclusive simulation sample. The obtained results are in good agreement, thus only studies using the $b\bar{b}$-inclusive sample are presented. The primary vertex algorithm used in these studies corresponds to the online 2015 approach i.e. the reconstruction using VELO tracks fitted by Kalman fitter.

The reconstruction efficiency is defined as the ratio of the number of reconstructed PVs to the number of reconstructible generated PVs. The latter is a pp interaction vertex which contains at least a minimal number of reconstructed VELO tracks. In the presented studies the minimal number of reconstructed tracks ($\min_{Tracks}$) varies between 3-5 and is used in the optimisation procedure.

Reconstructed PVs can be split into two categories: vertices which are matched by any generated PV (so-called true PV) and the remaining, not matched vertices (so-called false PV). The PV is matched when the distance to any reconstructible PV is $|\Delta z| < 5\sigma_z$, where σ_z denotes the estimated position uncertainty along z axis. The fake rate is defined as the ratio of the number of false PVs and all reconstructible generated PVs.

The optimisation is considered in a two dimensional space:

- minimum number of tracks in vertex: $\min_{Tracks} \in [3, 4, 5]$,
- maximum χ^2 to accept track: $\max_{\chi^2} \in [0.1, 40.0]$.

In the past (2010-2012 data taking) the nominal setting were: $\min_{Tracks} = 5$ and $\max_{\chi^2} = 9.0$. The efficiency and fake rate as a function of the above variables are shown in Fig. A.6. For $\min_{Tracks} = 4$ the efficiency increases by about 3% with respect to the nominal conditions, which can be considered as the new optimal value. At the same time, the fake rate rises twice. To prevent the growing fake rate, a radial distance cut is used as described later. An additional 1.5% in efficiency can be gained for $\min_{Tracks} = 3$, however in this condition the fake rate increases three times with respect to the nominal conditions, thus $\min_{Tracks} = 3$ is excluded from further studies. More detailed optimisation is considered in the narrower $\max_{\chi^2}$ range $[6.0, 12.0]$ and for $\min_{Tracks} \in [4, 5]$.

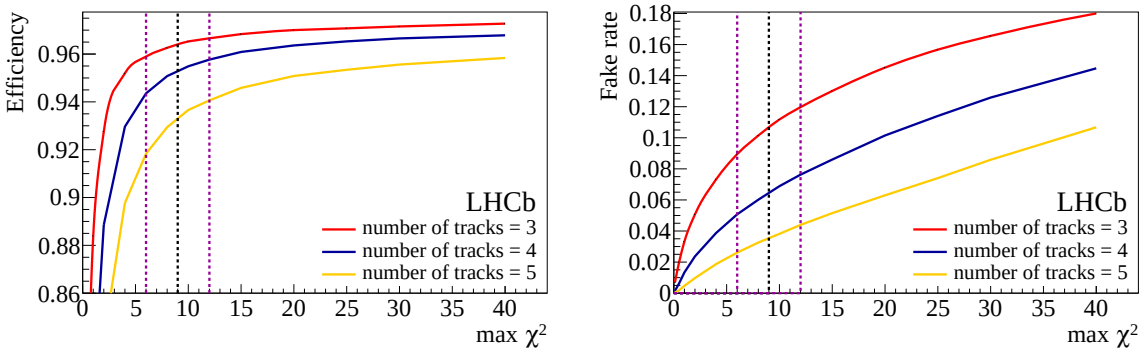

Figure A.6: Efficiency (left) and fake rate (right) as a function of minimal track multiplicty and maximal χ^2 to accept the tack. The default value: $\max_{\chi^2} = 9.0$ is denoted as a black dashed line. The red, blue and orange curves correspond to $\min_{Tracks} \in [3, 4, 5]$, respectively. The violet vertical lines correspond to the most interesting range which is considered in the further studies.

Radial distance optimisation

In order to reduce the fake rate several requirements were considered:

- radial distance of PV: $r = \sqrt{x^2 + y^2}$, where (x,y) is PV's position,
- minimal radial distance of track: $\min RD_{Track} = \min(\sqrt{x^2 + y^2})$ where (x,y) is the track position the closest to the beamline for tracks coming from the PV,
- maximal radial distance of track: $\max RD_{Track} = \max(\sqrt{x^2 + y^2})$ where (x,y) is the track position the closest to the beamline for tracks coming from the PV.

Figure A.7 shows the efficiency versus fake rate rejection, so-called ROC curve, for radial variables. As can be seen, the radial distance of the PV is the most powerful variable to reduce the fake rate, thus its optimisation is performed.

The radial distance distribution for false and true PVs is shown in Fig. A.8. The distributions for primary vertices reconstructed using $\min_{Tracks} = 4$ and $\max_{\chi^2} \in [6, 9, 12]$ are in good agreement. In addition, the radial distance for most of the true PVs is below 0.2 mm, while for false PVs it is extended until 8 mm.

The optimisation of radial distance maximises the purity multiplied by the true PVs efficiency, where the purity is defined as a ratio of the number of true PVs to all reconstructed PVs found in the event. The function of merit is maximised for $\min_{Tracks} = 4$ and $\max_{\chi^2} \in [6, 9, 12]$. The optimal working point is found independently on the maximal χ^2 to accept the track, with the optimal value $r < 0.2$ mm. Figure A.9 shows the optimisation function with all components, as an example, for $\min_{Tracks} = 4$ and $\max_{\chi^2} = 9$.

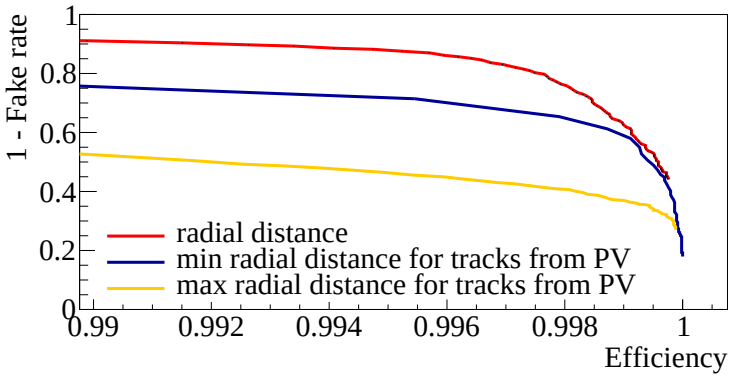

Figure A.7: Efficiency versus fake rate rejection (ROC curve) for radial variables.

Final optimisation

Figure A.10 shows the influence of applying the radial distance cut $r < 0.2$ to the reconstruction efficiency and fake rate for the following algorithm settings: $\min_{Tracks} = [4, 5]$ and $\max_{\chi^2} \in [6, 12]$. The radial cut is 99.5% efficient, while the fake rate is reduced by 85%.

In addition, the reconstruction using $\max_{\chi^2} = 12$ allows more tracks to form a primary vertex and thus provides better resolution for PVs with low track multiplicity. Table A.6 presents the resolution and for two reconstructions: $\max_{\chi^2} = 9$ and $\max_{\chi^2} = 12$. The additional tracks don't bias the vertex position. Therefore the optimal working point is set to be $\max_{\chi^2} = 12$.

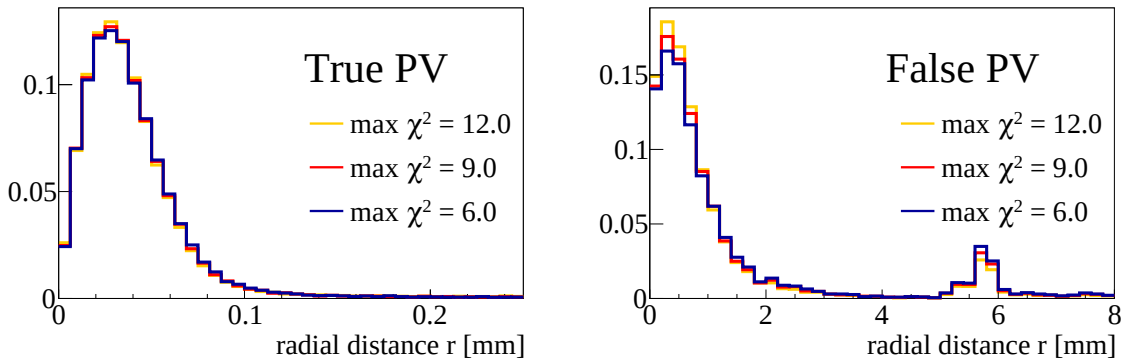

Figure A.8: Radial distance distributions for true PVs (left) and fake PVs (right) for the reconstruction with $\min_{Tracks} = 4$.

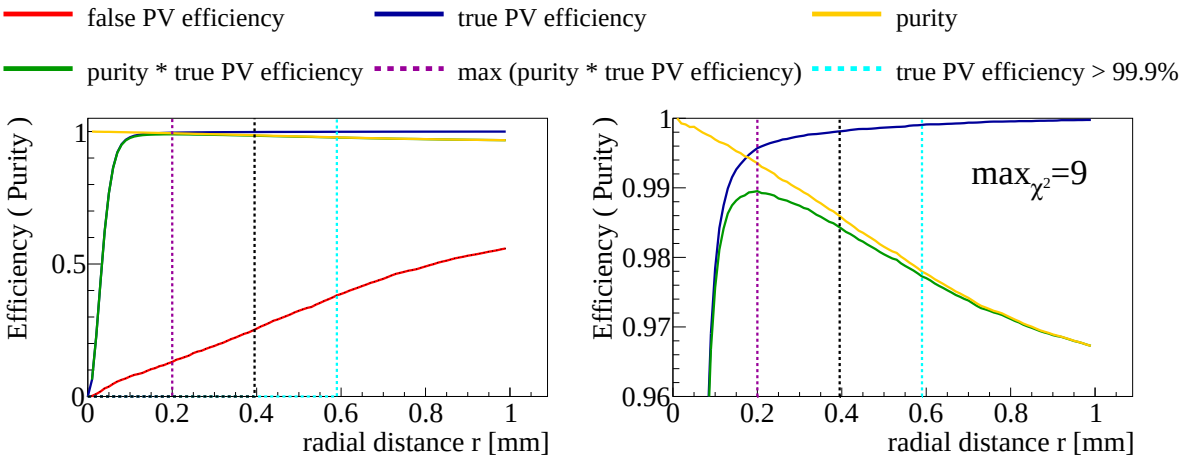

Figure A.9: Optimization of the radial distance cut for $\min_{Tracks} = 4$ and $\max_{\chi^2} = 9$. Left: full range, right: zoomed range above efficiency equal to 0.96. Different contributions are shown as coloured lines as described in the legend placed in the top.

Comparison

The true Primary Vertices can be split into several subgroups depending on their properties such as:

- distance to the closest reconstructible primary vertices:
  - Isolated - with distance to closest reconstructible PV $|\Delta z| > 10$ mm,
  - Non-Isolated - with distance to closest reconstructible PV $|\Delta z| < 10$ mm,
- multiplicity of tracks:
  - 1st - with the greatest track multiplicity,
  - 2nd - with 2nd greatest track multiplicity,
  - 3rd - with 3rd greatest track multiplicity,
- depending on daughter particles:
  - beauty - with beauty particle coming from it,

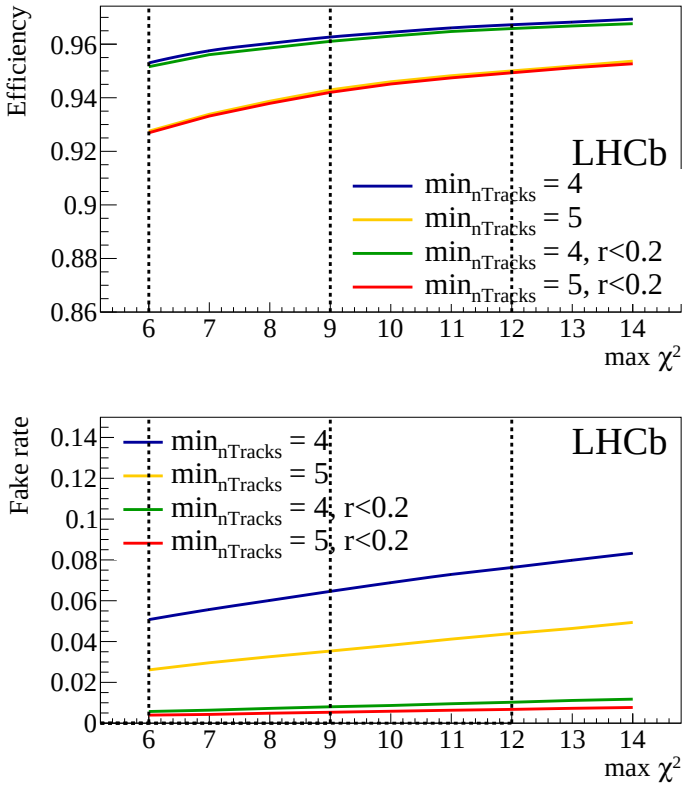

Figure A.10: Efficiency and fake rate as a function of minimal track multiplicty and maximal χ^2 to accept the tack before and after applying the radial distance cut $r < 0.2$.

Table A.6: Global resolution for primary vertex reconstructed with $\min_{Tracks} = 4$ and $\max_{\chi^2} = [9, 12]$.

Parameter [mm]	$max_{\chi^2} = 9$	$max_{\chi^2} = 12$
μ_x	2.88×10^{-5}	0.05×10^{-5}
Δx	0.0199	0.0195
μ_z	0.0016	0.0018
Δz	0.0140	0.0136

– charm - with charm particle coming from it,

– other.

In the same way, false primary vertices can be categorised depending on the misidentification into:

- charm - when the reconstructed primary vertex is a charm decay vertex,
- beauty - when the reconstructed primary vertex is a beauty decay vertex,
- other.

Table A.7 presents the efficiency comparison for all primary vertex categories between nominal ($\min_{Tracks} = 5$, $\max_{\chi^2} = 9.0$) and optimal ($\min_{Tracks} = 4$, $\max_{\chi^2} = 12.0$,$r < 0.2$) conditions for the online 2015 algorithm. Similar results for the fake rate are collected

Table A.7: Comparison of reconstruction efficiency for nominal ($\min_{Tracks} = 5$, $\max_{\chi^2} = 9.0$) and optimal ($\min_{Tracks} = 4$, $\max_{\chi^2} = 12.0$,$r < 0.2$) conditions for the online 2015 algorithm split by true PVs categories.

	Nominal	Optimal
true PV category	Efficiency [%]	
All	93.3	96.6
Isolated	93.3	98.0
Non-Isolated	79.8	87.0
1st	99.6	99.9
2nd	92.3	96.4
3rd	83.1	91.1
charm	98.1	99.4
beauty	98.6	99.1
other	86.5	93.1

Table A.8: Comparison of fake rate for nominal ($\min_{Tracks} = 5$, $\max_{\chi^2} = 9.0$) and optimal ($\min_{Tracks} = 4$, $\max_{\chi^2} = 12.0$,$r < 0.2$) conditions for the online 2015 algorithm split by false PV categories.

	Nominal	Optimal
false PV category	Fake rate [%]	
All	3.5	1.0
charm	0.1	0.1
beauty	2.5	0.8
other	0.9	0.1

in Tab. A.8. A substantial gain is visible for each category of true PVs, in particular for Non-Isolated (8%), 3rd (8%) and other (7%). The global efficiency increased by 3.3%. At the same time the fake rate reduces from 3.5% to 1.0%, with the most substantial reduction for beauty primary vertex misidentification, from 2.8% to 0.8%.

The reconstruction efficiency depends on the track multiplicity associated to the vertex, which varies with the data samples used. Moreover, a non-negligible effect occurs due to different numbers of *pp* interactions. In the case of multiple *pp* interactions the efficiency of close vertices is degraded; the lower multiplicity PV are affected due to track migration from the low multiplicity primary vertices to high multiplicity one. The efficiency of all true PVs as a function of track multiplicity is shown in Fig. A.11 for both the optimal and nominal conditions of the online 2015 algorithm. The PV reconstruction efficiency is high for large track multiplicity in the vertex. A substantial gain with respect to nominal conditions is observed for low track multiplicity.

A.3 Conclusion

Detailed studies of the primary vertex reconstruction have been performed. Thanks to using Kalman fitted VELO tracks as input the performance of the new algorithm is more similar to offline approach. In addition, in order to maximise the reconstruction efficiency

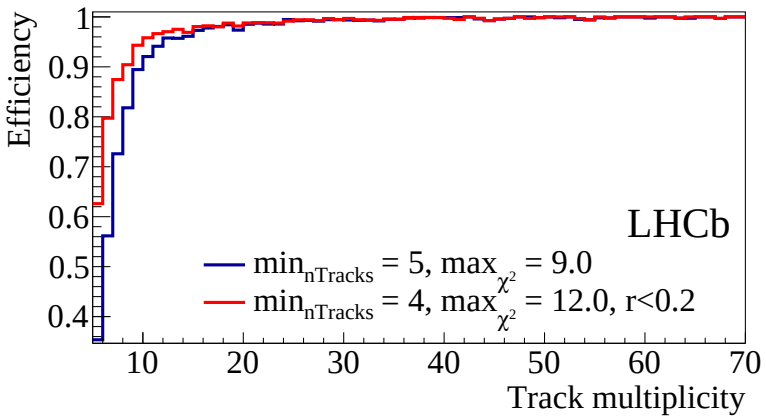

Figure A.11: Reconstruction efficiency as a function of track multiplicity for nominal ($\min_{Tracks} = 5$, $\max_{\chi^2} = 9.0$) and optimal ($\min_{Tracks} = 4$, $\max_{\chi^2} = 12.0$, $r < 0.2$ conditions and all reconstructed PVs.

and minimise the fake rate online 2015 algorithm has been optimised. The new working point is found to be $\min_{Tracks} = 4$, $\max_{\chi^2} = 12.0$, $r < 0.2$ with a 3% gain for efficiency and ∼70% fake rate reduction.

B `B2DXFitters` package

The `B2DXFitters` is a public package, which is a part of the Urania software project (see Fig. 3.15). It is based on the RooFit [75] framework and contains a core written in C++ followed by python scripts. The package provides all necessary scripts for performing the time-dependent measurement of beauty meson decays and is successfully used by other time-dependent analyses which consider similar decay chains.

As presented in Fig. B.1 the main fitting procedure is split in three stages. In addition to the classes included in the RooFit framework it uses:

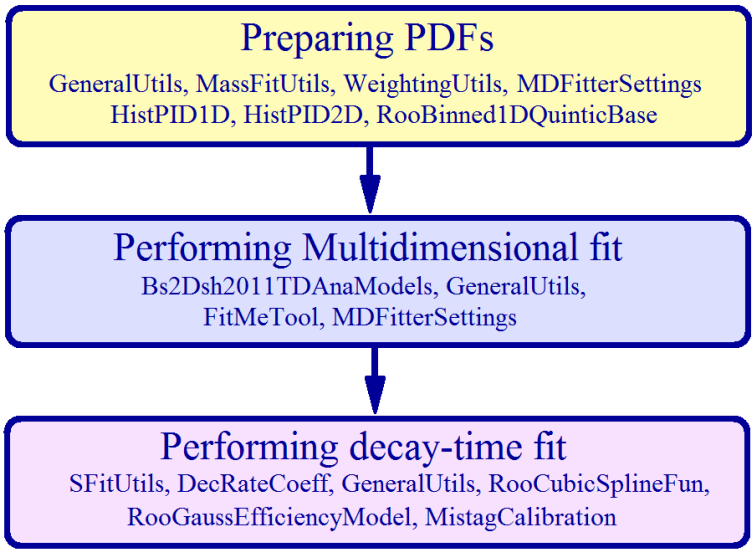

Figure B.1: A flowchart of the main fitting procedure in `B2DXFitter` package.

- **preparing PDFs**,
  which selects candidates under conditions set in `MDFitterSetting` and saves obtained data in the external file together with all necessary PDFs. The main classes are: `GeneralUtils` provides input/output tools, `MassFitUtils` is responsible for selecting candidates from data and simulation, `WeightingUtils` provides PIDK PDFs, `HistPID1D` weights PDFs by misidentification/efficiency of the PIDK requirements, `HistPID2D` applies two dimensional weighting for simulation samples to look like data and `RooBinned1D-QuinticBase` creates PDFs without freedom paramaters.

- **performing multidimensional fit**,
  which executes the multidimensional fit described in Chapter 6 and obtain *sWeights*. The main classes are: `Bs2Dsh2011TDAnaModels` contains all signal and background description and creates total PDFs, `GeneralUtils` provides input/output tools, `FitMeTool` is responsible for fit settings and obtaining *sWeights*, `MDFitterSetting` set data and PDFs configurations.

- **performing decay-time fit**,
  carries out the *sFit*. The main classes are: `SFitUtils` provides the *sWeighted* and combined data set as well as creates all necessary mistag rate and decay-time uncertainties PDFs, `DecRateCoeff` obtains the coefficients which are in front of the *CP* parameters in decay rates given by Eq. 7.9, all detector effects such

as detector/production asymmetries are taken into account, `GeneralUtils` provides input/output tools, `RooCubicSplineFun` creates the decay-time acceptance using splines described in Sec. 8.2, `RooGaussEfficiencyModel` obtains per-candidate decay-time resolution presented in Sec. 8.1, `MistagCalibration` provides calibrated mistag rate. The decay-time PDFs is taken as standard RooFit class: `RooBDecay`.

In addition to the main fitting procedure the package provides all the necessary scripts for visualisation of the obtained results. Moreover, a standalone pseudo experiment generator is a part of the package, which together with the fitting scripts allow to perform full studies. A standalone code produces samples in the same format as the data. It is a unique approach, which helps to find any incorrectly modelled effects in the main fit procedure. Furthermore, the `B2DXFitter` package supports: obtaining correlations between observables (see Fig. 6.1), determination of the expected background yields together with fitting $B_s^0 \to D_s^- \pi^+$ under $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ mass hypothesis (described in Sec. 6.3), performing fits to signal simulation samples and combinatorial background samples taken from the sideband, the calibration mistag rate using $B_s^0 \to D_s^- \pi^+ sFit$, performing fits to decay-time acceptance in simulation samples (see Fig. 8.6) and obtaining the triple Gaussian decay-time resolution model (see App. G).

C Additional information about selection.

Detailed information about the requirements applied in all stages of the trigger selection is listed in Tab. C.1 for the `1TrackAllL0` line, Tab. C.2 for the 2-, 3-, or 4-body `TopoBBDT` lines and Tab. C.3 for the inclusive $\phi^0 \to K^+K^-$ trigger line. The requirements used in the preselection are shown in Tab. C.4.

Table C.1: Selection criteria applied to candidates in `1TrackAllL0` at the HLT1 trigger stage. Ref. [67]

Applied to	Description	Requirement
Velo tracks	IP	> 100 μm
	Number of hits	> 9
	Number of missed hits	< 3
Other tracks	IPχ^2	> 16
	Number of hits	> 16
	p_T	$> 1700\,\text{MeV}/c$
	p	$> 10\,\text{GeV}/c$
	track χ^2/ndf	< 2.5

Table C.2: Selection applied to candidates in the 2-, 3-, or 4-body `TopoBBDT` at the HLT2 trigger stage. Indication $> 3, 4, 4\,\text{GeV}/c$ means that requirements > 3, > 4 and > 4 are applied to 2-body, 3-body and 4-body combination, respectively. Ref. [119]

Applied to	Description	Requirement
Input track	IP χ^2	> 4
	track χ^2	< 3
	p	$> 5\,\text{GeV}/c$
2-,3-, or 4-body combination	$\sum \lvert p_T \rvert$	$> 3, 4, 4\,\text{GeV}/c$
	DOCA	< 0.2 mm
	FDχ^2 wrt. PV	> 100
	p_T	$> 500\,\text{MeV}/c$
	BDT	$> 0.4, 0.4, 0.3$
	invariant mass	$< 7\,\text{GeV}/c^2$

Table C.3: Selection applied to candidates in the inclusive $\phi^0 \to K^+K^-$ trigger line at the HLT2 trigger stage. Ref. [119]

Applied to	Description	Requirement
$K^\pm$	IP χ^2	> 6
	track χ^2	< 5
	p_T	$> 800\,\text{MeV}/c$
	PIDK	> 0
$\phi^0(1020)$	$m(K^+K^-)$	in $[1000, 1040]\,\text{MeV}/c^2$
	p_T	$> 1800\,\text{MeV}/c$
	Vertex χ^2	< 20
	$DOCA$	< 0.2 mm

Table C.4: The preselection requirements.

Applied to	Description	Requirement
All tracks	track χ^2/ndf	< 4
	IPχ^2	> 4
	p	$> 1\,\text{GeV}/c$
	p_T	$> 0.1\,\text{GeV}/c$
D_s^0	$\sum p_T$	$> 1.8\,\text{GeV}/c$
	DOCA	< 0.5 mm
	Vertex χ^2/ndf	< 10
	FDχ^2 wrt. PV	> 36
	DIRA wrt. PV	> 0
	At least one child with:	
	track χ^2/ndf	< 10
	p	$> 5\,\text{GeV}/c$
	p_T	$> 0.5\,\text{GeV}/c$
bachelor	track χ^2/ndf	< 10
	p	$> 5\,\text{GeV}/c$
	p_T	$> 0.5\,\text{GeV}/c$
B_s^0	$\sum p_T$	$> 5\,\text{GeV}/c$
	Vertex χ^2/ndf	< 10
	FDχ^2 wrt. PV	> 36
	IPχ^2	< 25
	decay-time	> 0.2 ps
	DIRA wrt. PV	> 0.999
	BDT	> 0.5

D Fit to the $B^0 \to D^-\pi^+$ data sample

This appendix contains additional information about fits to the control mode: $B^0 \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$. The selection of the $B^0 \to D^-\pi^+$ sample is listed in Tab. D.1. The fits to the combinatorial background from $m(D^-\pi^+)$ invariant mass sideband: $m(D^-\pi^+) \in [5450, 7000]\,\text{MeV}/c^2$, but with full $m(K^+\pi^-\pi^-)$ range: $m(K^+\pi^-\pi^-) \in [1830, 1920]\,\text{MeV}/c^2$ are shown in Fig. D.1. The PDFs for the backgrounds used in control sample A (without PIDK cut on the bachelor) are presented in Fig. D.2. Tables D.2 and D.3 contain the results of fitting a double Crystal Ball function to $m(D^-\pi^+)$ and $m(K^+\pi^-\pi^-)$ invariant mass, respectively, based on simulation $B^0 \to D^-\pi^+$ samples. Finally, the numerical results of fits to control sample A and B are listed in Tab. D.4 and Tab. D.5, respectively.

Table D.1: Offline selection for $B^0 \to D^-\pi^+$ candidates. The only difference between control sample B and control sample A is due to the PIDK cut to the bachelor pion.

Description	Requirement
BDTG response	> 0.3
$m(D^-\pi^+)$ mass	$[5000, 6000]\,\text{MeV}/c^2$
$m(K^+\pi^-\pi^-)$ mass	$[1830, 1910]\,\text{MeV}/c^2$
D^- decay-time (wrt. B^0)	$> 0\,\text{ps}$
D^- FDχ^2 (wrt. B^0)	> 9
PIDK	< 0 (sample B) or none (sample A)
$m(\pi^-K^+)$	$> 840\,\text{MeV}/c^2$
$B^0 \to D^0K^+\pi^-$, $D^0 \to K^+\pi^-$ veto	
$m(D^-\pi^+)$ with π misID as K	not in $[1850, 1890]\,\text{MeV}/c^2$
semileptonic backgrounds veto	
PIDμ	< 2
$B^0 \to D^{*-}(D^0\pi^-)\pi^+$, $D^0 \to K^+\pi^-\pi^0$ veto	
$(m(K^+\pi^-\pi^-) - m(\pi^-K^+))$	> 200
Λ_c^+ veto:	
p veto for pions, or	PIDp < 0, or
$m(K^+\pi^-\pi^-)$ under $m(pK^-\pi^-)$ hypothesis	not in $[2255, 2315]\,\text{MeV}/c^2$
D_s^- veto:	
kaon veto for pions, or	PIDK < 0, or
$m(K^+\pi^-\pi^-)$ under $m(K^+K^-\pi^-)$ hypothesis	not in $[1950, 2300]\,\text{MeV}/c^2$

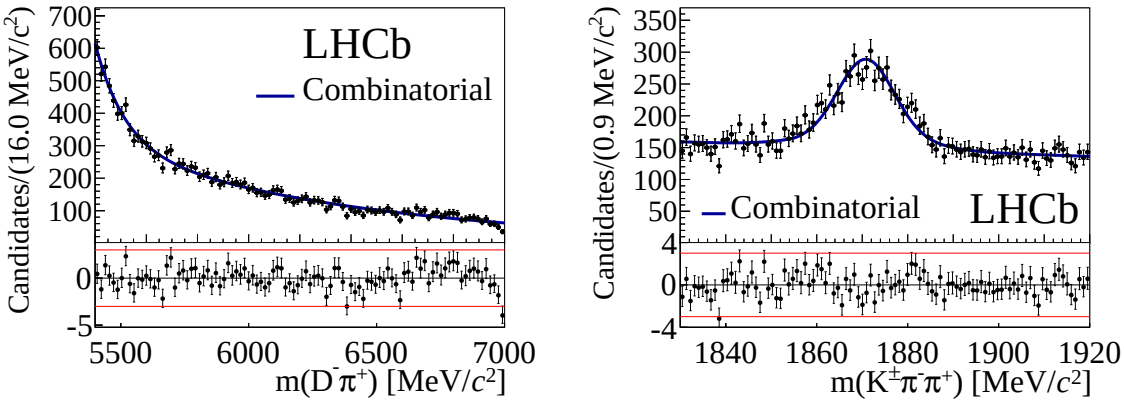

Figure D.1: The results of the fit to the combinatorial background in the $B^0 \to D^-\pi^+$ sample. The solid, blue curve corresponds to the result of the fit. Left: $m(D^-\pi^+)$ invariant mass. Right: $m(K^+\pi^-\pi^-)$ invariant mass.

Table D.2: Parameters for the sum of the double Crystal Ball function describing the signal $m(D^-\pi^+)$ invariant mass for the $B^0 \to D^-\pi^+$ decay, obtained from signal simulation samples.

Parameter	Fitted value	
	Control sample A	Control sample B
μ_{B^0} [MeV/c^2]	5279.3 ± 0.1	5280.2 ± 0.1
$\sigma^1_{B^0}$ [MeV/c^2]	27.5 ± 0.7	8.6 ± 0.7
$\sigma^2_{B^0}$ [MeV/c^2]	13.1 ± 0.13	14.5 ± 0.1
$\alpha^1_{B^0}$	1.0 ± 0.1	0.9 ± 0.1
$\alpha^2_{B^0}$	-4.8 ± 1.2	-2.1 ± 0.1
$n^1_{B^0}$	2.3 ± 0.3	1.8 ± 0.1
$n^2_{B^0}$	18.0 ± 14.3	5.4 ± 1.1
f_{B^0}	0.17 ± 0.01	0.25 ± 0.02

Table D.3: Parameters for the sum of the double Crystal Ball function describing the signal $m(K^+\pi^-\pi^-)$ invariant mass for the $B^0 \to D^-\pi^+$ decay, obtained from signal simulation samples.

Parameter	Fitted value	
	Control sample A	Control sample B
μ_{D^-} [MeV/c^2]	1869.8 ± 0.04	1869.8 ± 0.04
$\sigma^1_{D^-}$ [MeV/c^2]	9.8 ± 0.3	10.0 ± 1.5
$\sigma^2_{D^-}$ [MeV/c^2]	5.6 ± 0.1	5.6 ± 0.3
$\alpha^1_{D^-}$	1.8 ± 0.06	1.7 ± 0.1
$\alpha^2_{D^-}$	-2.8 ± 0.2	-3.0 ± 1.7
$n^1_{D^-}$	1.6 ± 0.1	1.5 ± 0.1
$n^2_{D^-}$	0.9 ± 0.3	0.6 ± 2.2
f_{D^-}	0.22 ± 0.02	0.41 ± 0.12

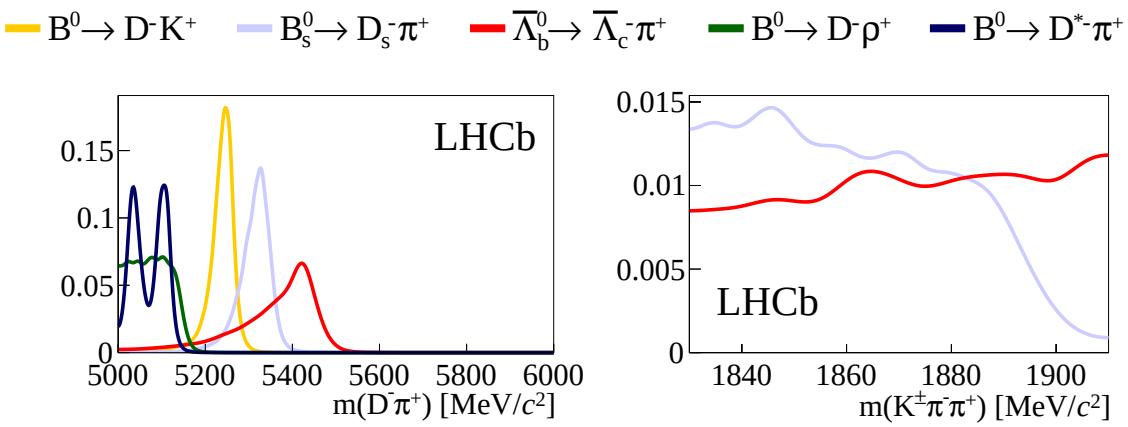

Figure D.2: The background PDFs for the control sample A used in the fit to the $B^0 \to D^-\pi^+$ sample, combined for both magnet polarities. Left: $m(D^-\pi^+)$ invariant mass. Right: $m(K^+\pi^-\pi^-)$ invariant mass. Different background contributions are shown as coloured lines as described in the legend above the plots.

Table D.4: Results for the control sample A fit to the $B^0 \to D^-\pi^+$ sample. The N_i are the yields of the signal and background contributions. Means μ_{B^0} and μ_{D^-} are the parameters of the double Crystal Ball used to describe the signal in $m(D^-\pi^+)$ and $m(K^+\pi^-\pi^-)$ invariant masses, respectively. The $s_{B^0(D^-)}$ are the slopes of combinatorial background. The fraction $fC_{B^0_s}$ is the fraction between exponentials in $m(D^-\pi^+)$ invariant mass, whereas f_{D^-} are the fractions between the exponential and signal PDF in $m(K^+\pi^-\pi^-)$ invariant mass.

Parameter	Fitted value		Parameter	Fitted value
	Magnet Down	Magnet Up		
$N_{B^0\to D^-\pi^+}$	77571 ± 334	53880 ± 274	$\sigma^1_{B^0}$ [MeV/c^2]	25.9 ± 0.38
$N_{B^0\to D^-\rho^+}$	37813 ± 553	26802 ± 456	$\sigma^2_{B^0}$ [MeV/c^2]	15.9 ± 0.07
$N_{B^0\to D^{*-}\pi^+}$	17242 ± 489	11632 ± 409	$\sigma^1_{D^-}$ [MeV/c^2]	11.5 ± 0.06
$N_{B^0_s\to D^-_s\pi^+}$	1021 ± 123	955 ± 106	$\sigma^2_{D^-}$ [MeV/c^2]	6.7 ± 0.03
$N_{\overline{\Lambda}^0_b\to\overline{\Lambda}^-_c\pi^+}$	308 ± 106	236 ± 87	$s^2_{B^0}$	-0.0042 ± 0.00011
N_{Comb}	20481 ± 302	14221 ± 239	s_{D^-}	-0.0014 ± 0.00032
μ_{B^0} [MeV/c^2]	5283.2 ± 0.10	5282.6 ± 0.08	fC_{B^0}	0.78 ± 0.01
μ_{D^-} [MeV/c^2]	1871.2 ± 0.03	1871.0 ± 0.03	fC_{D^-}	0.83 ± 0.01

Table D.5: Results for the control sample B fit to the $B^0 \to D^-\pi^+$ sample. The N_i are the yields of the signal and background contributions. Means μ_{B^0} and μ_{D^-} are the parameters of the double Crystal Ball used to describe the signal in $m(D^-\pi^+)$ and $m(K^+\pi^-\pi^-)$ invariant masses, respectively. The $s_{B^0(D^-)}$ are the slopes of combinatorial background. The fraction $fC_{B^0_s}$ is the fraction between exponentials in B^0 mass, whereas fC_{D^-} are the fractions between the exponential and signal PDF in $m(K^+\pi^-\pi^-)$ invariant mass.

Parameter	Fitted value	Parameter	Fitted value	Parameter	Fitted value
$N_{B^0\to D^-\pi^+}$	109420 ± 360	μ_{B^0} [MeV/c^2]	5283.8 ± 0.06	$s^1_{B^0}$	0.0 ± 0.0
$N_{B^0\to D^-\rho^+}$	54240 ± 632	μ_{D^-} [MeV/c^2]	1871.1 ± 0.02	$s^2_{B^0}$	-0.00418 ± 0.00016
$N_{B^0\to D^{*-}\pi^+}$	24395 ± 554	$\sigma^1_{B^0}$ [MeV/c^2]	10.94 ± 0.25	s_{D^-}	-0.00096 ± 0.00038
$N_{B^0_s\to D^-_s\pi^+}$	1294 ± 119	$\sigma^2_{B^0}$ [MeV/c^2]	17.66 ± 0.08	fC_{B^0}	0.186 ± 0.017
$N_{\overline{\Lambda}^0_b\to\overline{\Lambda}^-_c\pi^+}$	213 ± 298	$\sigma^1_{D^-}$ [MeV/c^2]	11.54 ± 0.07	fC_{D^-}	0.872 ± 0.009
N_{Comb}	23413 ± 401	$\sigma^2_{D^-}$ [MeV/c^2]	6.68 ± 0.03		

E Signal and background PDFs used in the fit to the $B_s^0 \to D_s^- \pi^+$ sample

E.1 Signal

The signal PDFs for $B_s^0 \to D_s^- \pi^+$ are shown in Fig. E.1. The results of the double Crystal Ball parametrisation are listed in Tab. E.1 separately for each D_s^- final state, for the $m(D_s^\mp h^\pm)$ invariant mass, while the results for the $m(h^- h^+ h^\pm)$ invariant mass are collected in Tab. E.2.

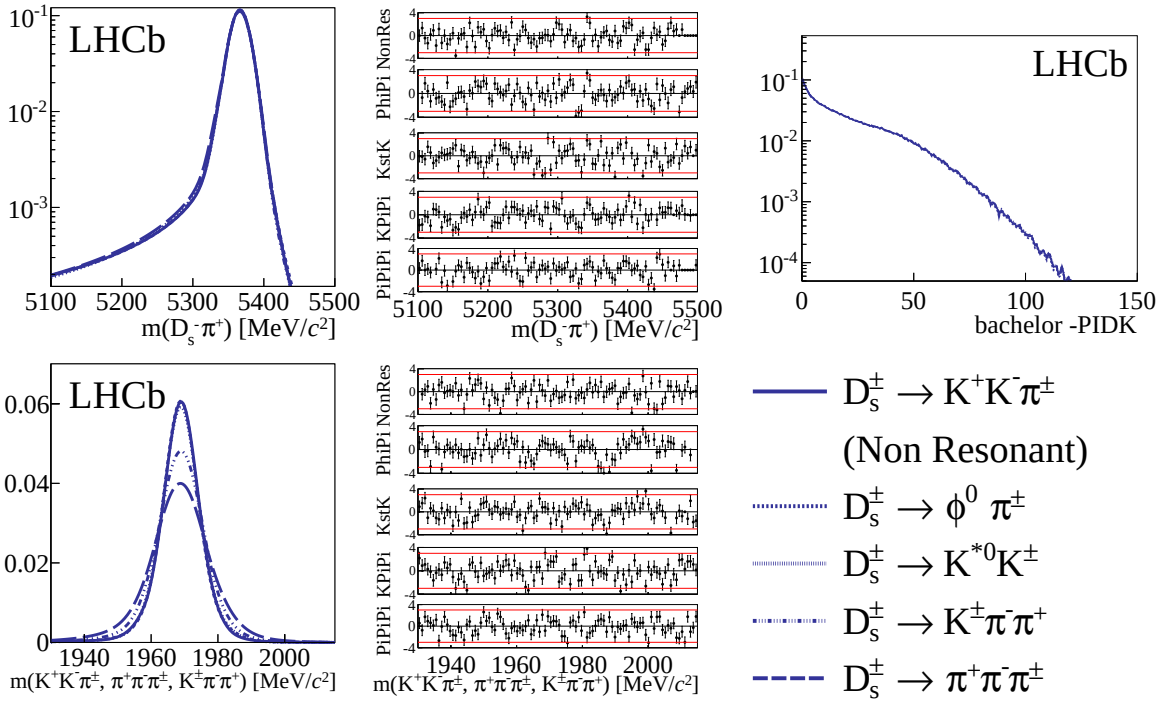

Figure E.1: Signal PDFs of $B_s^0 \to D_s^- \pi^+$ sample as evaluated from simulation on the joint magnet up and down samples, separately for each D_s^- final states. Different D_s^- final states are shown as blue lines as described in the legend placed in the bottom right. Top left: the $m(D_s^- \pi^+)$ invariant mass together with the resulting pull distributions, top right: the -PIDK variable, bottom: the $m(h^- h^+ h^\pm)$ invariant mass together with the resulting pull distributions.

Table E.1: Parameters for the sum of the double Crystal Ball function describing the signal $m(D_s^\mp h^\pm)$ invariant mass of $B_s^0 \to D_s^- \pi^+$ sample, obtained from signal simulation, separately for each D_s^- final state.

Parameter	$D_s^- \to (KK\pi)_{\text{nonres}}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$\mu_{B_s^0}$ [MeV/c^2]	5366.3 ± 0.04	5366.3 ± 0.02	5366.3 ± 0.02	5366.2 ± 0.04	5366.2 ± 0.05
$\sigma^1_{B_s^0}$ [MeV/c^2]	11.54 ± 0.13	16.60 ± 0.17	11.65 ± 0.14	11.43 ± 0.13	11.39 ± 0.10
$\sigma^2_{B_s^0}$ [MeV/c^2]	16.18 ± 0.09	11.49 ± 0.08	15.00 ± 0.09	16.87 ± 0.12	17.65 ± 0.10
$\alpha^1_{B_s^0}$	1.91	-2.09	1.70	1.91	2.09
$\alpha^2_{B_s^0}$	-2.04	1.89	-1.84	-2.26	-2.33
$n^1_{B_s^0}$	1.13 ± 0.02	5.27 ± 0.26	1.27 ± 0.02	1.16 ± 0.02	1.27 ± 0.02
$n^2_{B_s^0}$	6.14 ± 0.51	1.15 ± 0.02	9.66 ± 0.09	4.12 ± 0.33	4.02 ± 0.36
$f_{B_s^0}$	0.54 ± 0.04	0.44 ± 0.02	0.54 ± 0.03	0.54 ± 0.04	0.70 ± 0.02

Table E.2: Parameters for the sum of the double Crystal Ball function describing the signal $m(h^-h^+h^\pm)$ invariant mass of $B_s^0 \to D_s^-\pi^+$ sample, obtained from signal simulation, separately for each D_s^- final state.

Parameter	$D_s^- \to (KK\pi)_{\rm nonres}$	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$\mu_{D_s^-}$ [MeV/c^2]	1969.0 ± 0.02	1968.9 ± 0.01	1968.9 ± 0.01	1969.0 ± 0.04	1968.8 ± 0.05
$\sigma^1_{D_s^-}$ [MeV/c^2]	4.39 ± 0.08	8.72 ± 0.08	7.88 ± 0.08	6.77 ± 0.19	8.42 ± 0.12
$\sigma^2_{D_s^-}$ [MeV/c^2]	7.15 ± 0.15	4.62 ± 0.03	4.59 ± 0.05	6.49 ± 0.16	7.26 ± 0.18
$\alpha^1_{D_s^-}$	2.20 ± 0.12	1.80 ± 0.04	1.97 ± 0.05	0.92 ± 0.11	0.95 ± 0.11
$\alpha^2_{D_s^-}$	-2.02 ± 0.09	-3.21 ± 0.12	-2.77 ± 0.11	-1.28 ± 0.04	-1.04 ± 0.06
$n^1_{D_s^-}$	0.79 ± 0.15	2.67 ± 0.31	2.08 ± 0.24	9.28 ± 5.59	12.9 ± 11.1
$n^2_{D_s^-}$	5.56 ± 1.13	0.45 ± 0.14	1.08 ± 0.20	46.5 ± 9.30	70.0 ± 8.15
$f_{D_s^-}$	0.49 ± 0.03	0.35 ± 0.09	0.46 ± 0.02	0.36 ± 0.04	0.50 ± 0.06

E.2 Combinatorial background

The parameterizations of the combinatorial background in the $B_s^0 \to D_s^-\pi^+$ sample using the sideband region $m(D_s^\mp h^\pm) > 5800\,\text{MeV}/c^2$ region with the full D_s^- range: $m(h^-h^+h^\pm) = (1930, 2015)$ MeV/c^2 are shown in Fig. E.2.

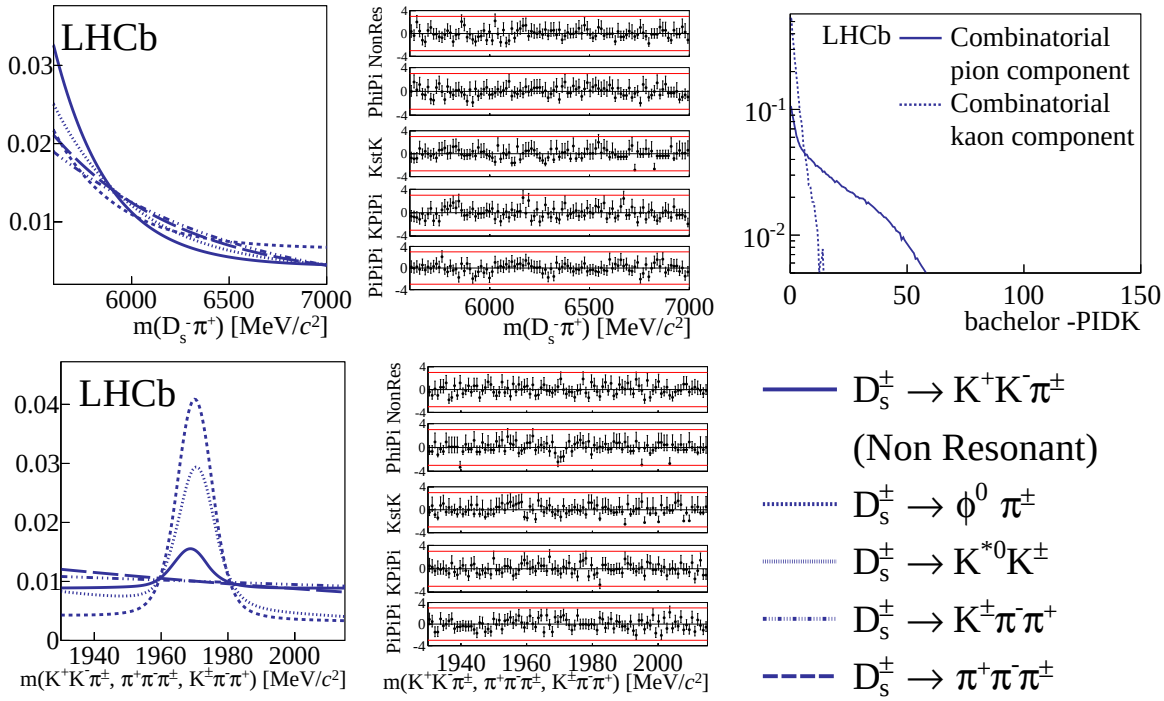

Figure E.2: Combinatorial background PDFs of $B_s^0 \to D_s^-\pi^+$ sample as evaluated from $m(D_s^\mp h^\pm)$ invariant mass sideband on the joint magnet up and down samples, separately for each D_s^- final states. For the $m(D_s^\mp h^\pm)$ and $m(h^-h^+h^\pm)$ invariant masses, different D_s^- final states are shown as blue lines as described in the legend in the legend placed in the bottom right. For the PIDK variable pion and kaon components are described in the legend placed on the plots. Top left: the $m(D_s^-\pi^+)$ invariant mass together with the resulting pull distributions, top right: the -PIDK variable, bottom: the $m(h^-h^+h^\pm)$ invariant mass together with the resulting pull distributions.

E.3 Other fully and partially reconstructed backgrounds

The PDFs for fully reconstructed backgrounds: $B^0 \to D^- \pi^+$, $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ as well as the partially reconstructed: $B_s^0 \to D_s^{*-} \pi^+$ are shown in Fig. E.3

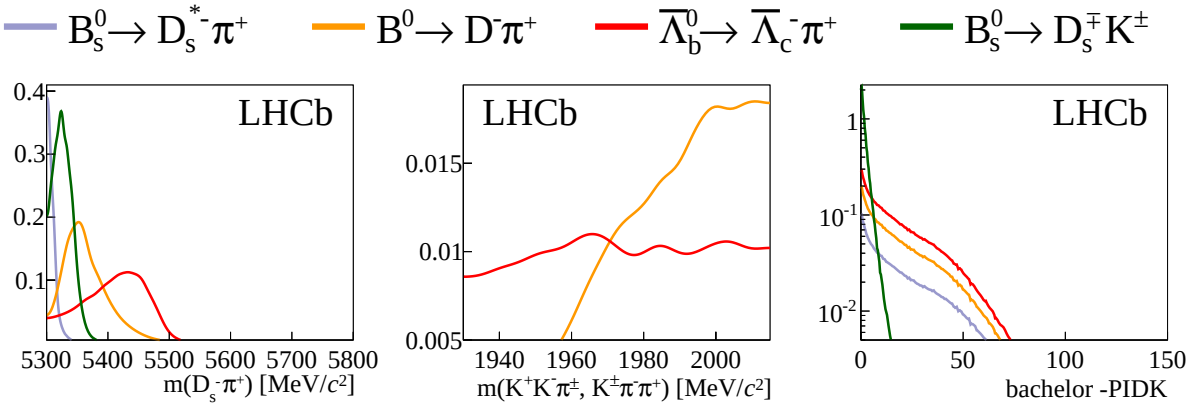

Figure E.3: The resulting PDFs taken from simulation for the remaining fully reconstructed backgrounds in the $B_s^0 \to D_s^- \pi^+$ sample. Different background contributions are shown as coloured lines as described in the legend above the plots. Left: $m(D_s^{\mp} h^{\pm})$ invariant mass, Middle: $m(h^- h^+ h^{\pm})$ invariant mass, Right: -PIDK variable.

F Multidimensional fit to $B_s^0 \to D_s^- \pi^+$

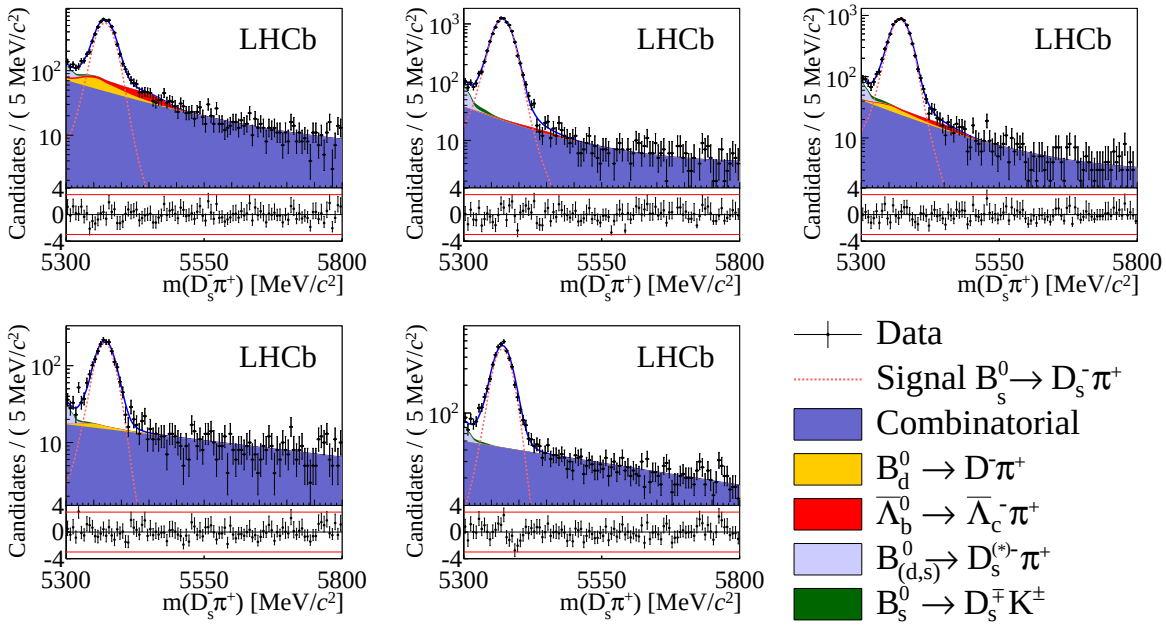

Figure F.1: Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of $m(D_s^{\mp} h^{\pm})$ invariant mass for D_s^- final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{\text{nonres}}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend placed in the bottom right.

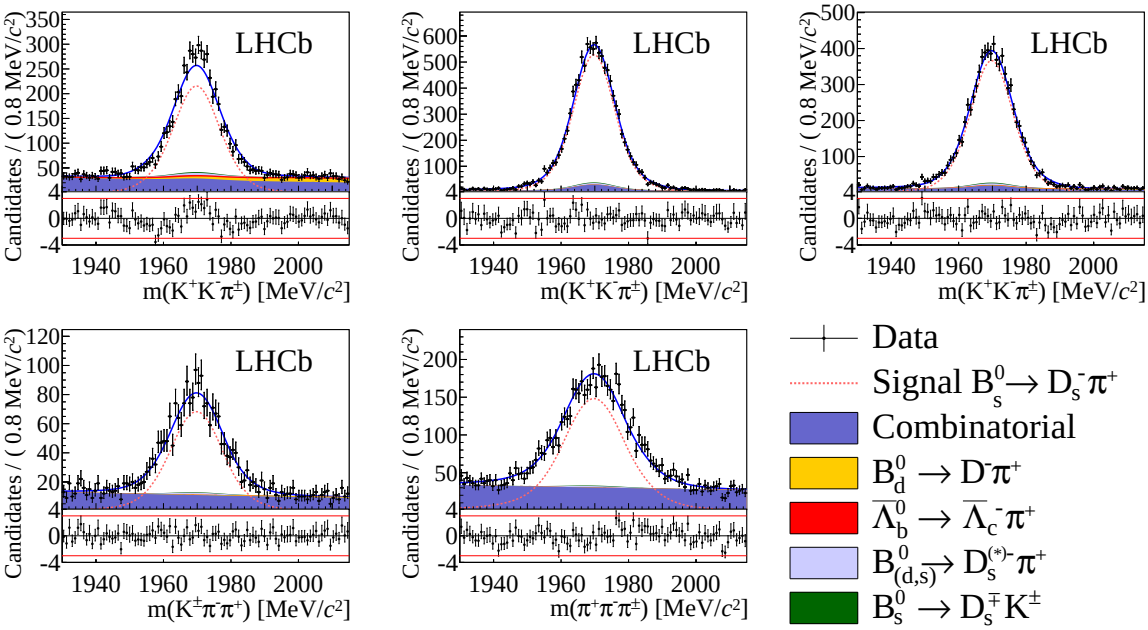

Figure F.2: Result of the simultaneous fit to the $B_s^0 \to D_s^- \pi^+$ candidates. Distributions of $m(h^-h^+h^\pm)$ invariant mass for D_s^- final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{\rm nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend placed in the bottom right.

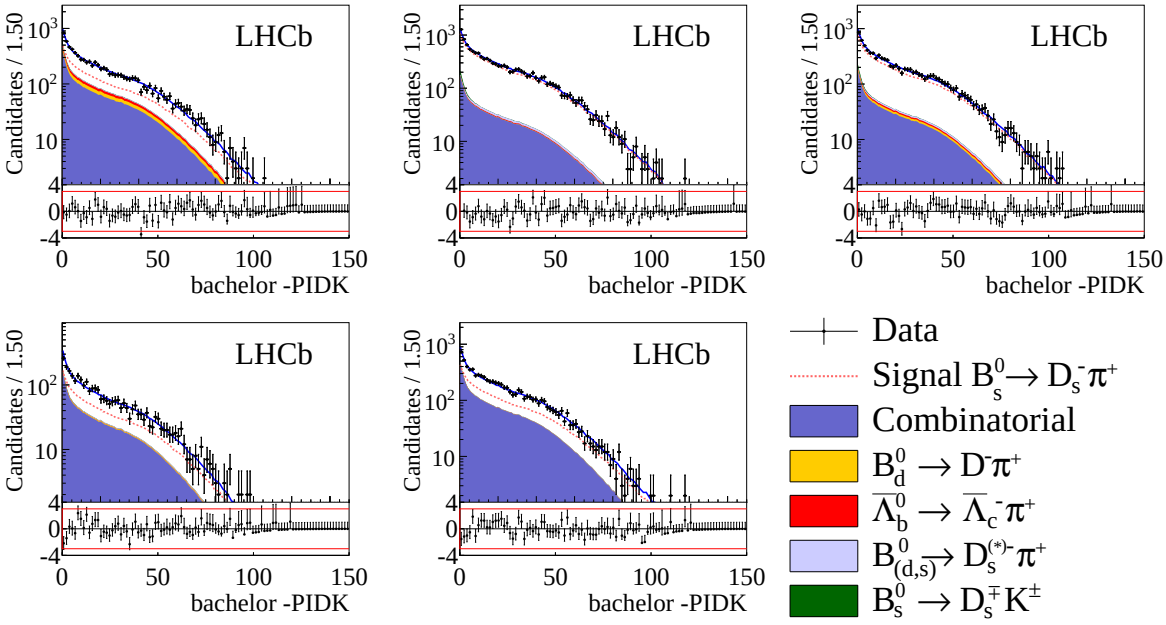

Figure F.3: Result of the simultaneous fit to the $B_s^0 \to D_s^\mp K^\pm$ candidates. Distributions of -PIDK for D_s^- final states with combined magnet polarities, from top left to bottom: $D_s^- \to (KK\pi)_{\rm nonres}$, $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. Different contributions to the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend placed in the bottom right.

G Decay-time resolution modelled by a triple Gaussian

The decay-time resolution is determined in the following way:

- the PDF of the estimated decay-time uncertainties $PDF(\sigma_t)$ are taken from $sWeighted$ $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ data, as shown in Fig. 8.3,
- the decay-time observable t is distributed as a Gaussian with mean $\mu = 0$ and width $\sigma = \sigma_t$: $PDF(\tau|\sigma_t) = G(\mu, \sigma_t; t)$,
- the conditional decay-time PDF is defined as: $PDF(t) = PDF(t|\sigma_t) \cdot PDF(\sigma_t)$,
- the decay-time resolution is parametrised by a triple Gaussian:

$$R(\sigma_1, \sigma_2, \sigma_3, f_1, f_2; t) = f_1 G(0, \sigma_1; t) + f_2 G(0, \sigma_2; t) + (1 - f_1 - f_2) G(0, \sigma_3; t), \quad (8)$$

  where $G(0, \sigma_i; t)$ denotes single Gaussian distribution with mean $\mu = 0$ and width $\sigma = \sigma_i$,
- the decay-time resolution is fitted to the conditional time distribution $PDF(t)$.

The obtained results are shown in Fig. G.1 and collected in Tab. G.1. The fitted parameters are in good agreement between the two modes. In the decay-time *sFit* to data, the widths of the resolution have to multiplied by the scale factor: $s_{\sigma_t} = 1.37 \pm 0.10$.

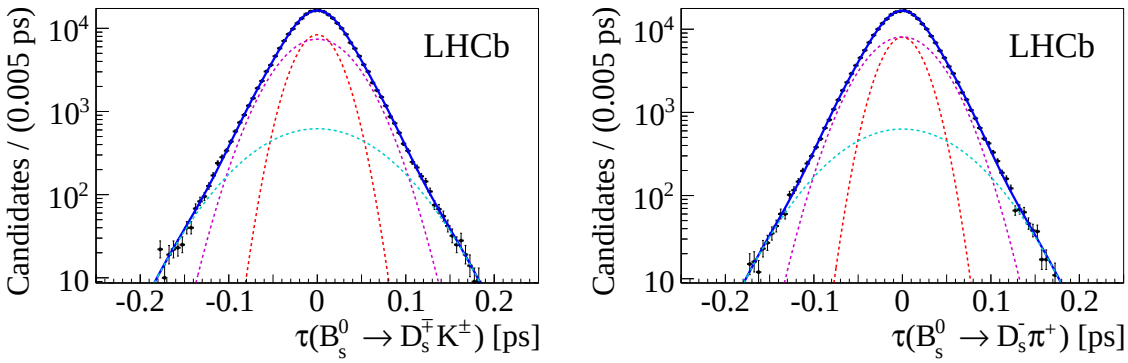

Figure G.1: The resolution model parametrises by the triple Gaussian for the $B_s^0 \to D_s^\mp K^\pm$ (left) and $B_s^0 \to D_s^- \pi^+$ (right) decays.

Table G.1: Result of fitting decay-time resolution modelled by the triple Gaussian to the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ samples.

Parameters	Fit to $B_s^0 \to D_s^- \pi^+$	Fit to $B_s^0 \to D_s^\mp K^\pm$
f_1	33.7 ± 3.4	36.7 ± 3.5
f_2	58.6 ± 2.5	55.4 ± 2.6
σ_1 [ps]	2.10 ± 0.06	2.19 ± 0.06
σ_2 [ps]	3.61 ± 0.09	3.73 ± 0.09
σ_3 [ps]	6.14 ± 0.17	6.30 ± 0.17

H Acceptance parameter correlations

Here are the correlation matrices of the six spline parameters of the decay-time acceptance parameterisation, discussed in the decay-time acceptance section (Sec. 8.2), and needed to evaluate the systematics in Sec. 10.5.

Table H.1: Correlation matrix of the $B_s^0 \to D_s^- \pi^+$ acceptance *sFit* on data, corresponding to the results given in Tab. 9.1.

Params.	v_1	v_2	v_3	v_4	v_5	v_6
v_1	1.000	0.868	0.779	0.892	0.894	0.840
v_2	0.868	1.000	0.615	0.870	0.822	0.797
v_3	0.779	0.615	1.000	0.695	0.857	0.725
v_4	0.892	0.870	0.695	1.000	0.840	0.878
v_5	0.894	0.822	0.857	0.840	1.000	0.748
v_6	0.840	0.797	0.725	0.878	0.748	1.000

Table H.2: Correlation matrix for parameters listed in Tab. 8.1 (the $B_s^0 \to D_s^- \pi^+$ simulation).

Params.	v_1	v_2	v_3	v_4	v_5	v_6
v_1	1.000	0.893	0.842	0.917	0.922	0.874
v_2	0.893	1.000	0.702	0.898	0.859	0.836
v_3	0.842	0.702	1.000	0.765	0.895	0.787
v_4	0.917	0.898	0.765	1.000	0.873	0.899
v_5	0.922	0.859	0.895	0.873	1.000	0.801
v_6	0.874	0.836	0.787	0.899	0.801	1.000

Table H.3: Correlation matrix for parameters listed in Tab. 8.1 (the $B_s^0 \to D_s^\mp K^\pm$ simulation).

Params.	v_1	v_2	v_3	v_4	v_5	v_6
v_1	1.000	0.893	0.835	0.916	0.919	0.873
v_2	0.893	1.000	0.695	0.898	0.857	0.836
v_3	0.835	0.695	1.000	0.760	0.891	0.782
v_4	0.916	0.898	0.760	1.000	0.873	0.900
v_5	0.919	0.857	0.891	0.873	1.000	0.800
v_6	0.873	0.836	0.782	0.900	0.800	1.000

Table H.4: Correlation matrix for parameters listed in Table 8.1 (ratios of acceptance parameters $B_s^0 \to D_s^{\mp} K^{\pm} / B_s^0 \to D_s^- \pi^+$). These correlations are computed by propagating the covariances of Tabs. H.2 and H.3 through the calculation of the ratio.

Params.	v_1	v_2	v_3	v_4	v_5	v_6
v_1	1.	0.72	0.71	0.8	0.79	0.76
v_2	0.72	1.	0.55	0.79	0.73	0.72
v_3	0.71	0.55	1.	0.77	0.9	0.78
v_4	0.8	0.79	0.77	1.	0.87	0.91
v_5	0.79	0.73	0.9	0.87	1.	0.8
v_6	0.76	0.72	0.78	0.91	0.8	1.

Bibliography

[1] A. Sakharov, *Violation of CP invariance, c asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[2] LHCb collaboration, R. Aaij *et al.*, *Measurement of the B^0–$\bar{B}^0$ oscillation frequency Δm_d with the decays $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi K^{*0}$*, Phys. Lett. **B719** (2013) 318, arXiv:1210.6750.

[3] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286.

[4] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency in the decay $B_s^0 \to D_s^+\pi^-$*, New J. Phys. **15** (2013) 053021, arXiv:1304.4741.

[5] LHCb, R. Aaij *et al.*, *Study of the kinematic dependences of Λ_b^0 production in pp collisions and a measurement of the $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ branching fraction*, JHEP **1408** (2014) 143, arXiv:1405.6842.

[6] LHCb collaboration, R. Aaij *et al.*, *Determination of the branching fractions of $B_s^0 \to D_s^{\mp}K^{\pm}$ and $B^0 \to D_s^-K^+$* , arXiv:1412.7654.

[7] D. Griffiths, *Introduction To Elementary Particles*, John Wiley and sons (1987), ISBN: 0-471-60386-4.

[8] F. Halzen and A. Martin, *Quarks and Leptons: and Introduction Course in Modern Particle Physics*, John Wiley and sons (1984), ISBN: 0-471-88741-2.

[9] G. Branco and J. Lavoura, L Silva, *CP Violation*, International Series of Monographs on Physics (1999), Clarendon Press, Oxford.

[10] M. Gell-Mann, *A schematic model of baryons and mesons*, Physics Letters **8** (1964), no. 3 214 .

[11] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343.

[12] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[13] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367.

[14] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[15] ATLAS collaboration, G. Aad *et al.*, *The ATLAS experiment at the CERN LHC*, JINST **3** (2008) S08003.

[16] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[17] E. Noether, *Invariante Variationsprobleme*, Nachr. d. Konig. Gesselsch. d. Wiss. zu Gottigen (1918) 235.

[18] C. Wu *et al.*, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413.

[19] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *Decay of the* K_2^0 *Meson*, Phys. Rev. Lett. **13** (1964) 138.

[20] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[21] M. Kobayashi and T. Maskawa, *CP Violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[22] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[23] The CKMfitter Group, J. Charles *et al.*, *updated results and plots available at: http://ckmfitter.in2p3.fr*, Eur. Phys. J. **C41** (2005) 1.

[24] The Belle collaboration, A. Abashian *et al.*, *The Belle detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **479** (2002), no. 1 117 , Detectors for Asymmetric B-factories.

[25] KEK: An Asymmetric Electron-Positron Collider for B Physics, web page: http://www-kekb.kek.jp/.

[26] BaBar, B. Aubert *et al.*, *The BaBar detector*, Nucl. Instrum. Meth. **A479** (2002) 1, `arXiv:hep-ex/0105044`.

[27] SLAC National Accelerator Laboratory, web page: https://www6.slac.stanford.edu/.

[28] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[29] CERN: the European Organization for Nuclear Research, web page: http://home.web.cern.ch/.

[30] LHCb Collaboration, R. Aaij *et al.*, *Improved constraints on* γ*: CKM2014 update*, LHCb-CONF-2014-04.

[31] LHCb collaboration, R. Aaij *et al.*, *First observation of* CP *violation in the decays of* B_s^0 *mesons*, Phys. Rev. Lett. **110** (2013) 221601, `arXiv:1304.6173`.

[32] LHCb Collaboration, R. Aaij *et al.*, *Observation of CP violation in* $B^\pm \to DK^\pm$ *decays*, Phys. Lett. B **712** (2012) 203, `arXiv:1203.3662`.

[33] M. Beneke, G. Buchalla, A. Lenz, and U. Nierste, *CP asymmetry in flavor specific B decays beyond leading logarithms*, Phys. Lett. **B576** (2003) 173, arXiv:hep-ph/0307344.

[34] R. Fleischer, *Flavor physics and CP violation*, arXiv:hep-ph/0405091 CERN-PH-TH-2004-085, arXiv:hep-ph/0405091.

[35] B. Bhattacharya, M. Imbeault, and D. London, *Extraction of the CP-violating phase γ using $B \to K\pi\pi$ and $B \to KK\bar{K}$ decays*, Phys. Lett. **B728** (2014) 206, arXiv:1303.0846.

[36] LHCb collaboration, *Evidence for CP violation in $B \to KK\pi$ and $B \to \pi\pi\pi$ decays*, LHCb-CONF-2012-028.

[37] R. H. Dalitz, *Decay of τ mesons of known charge*, Phys. Rev. **94** (1954) 1046.

[38] M. Gronau, *U-spin breaking in CP asymmetries in B decays*, Phys. Lett. **B727** (2013) 136, arXiv:1308.3448.

[39] LHCb collaboration, R. Aaij *et al.*, *Determination of γ and $-2\beta_s$ from charmless two-body decays of beauty mesons*, Phys. Lett. **B741** (2014) 1 LHCb-PAPER-2014-045, arXiv:1408.4368.

[40] M. Gronau and D. Wyler, *On determining a weak phase from CP asymmetries in charged B decays*, Phys. Lett. **B265** (1991) 172.

[41] M. Gronau and D. London, *How to determine all the angles of the unitarity triangle from $B^0_d \to DK_s$ and $B^0_s \to D\phi$*, Phys. Lett. **B253** (1991) 483.

[42] D. Atwood, I. Dunietz, and A. Soni, *Enhanced CP Violation with $B \to KD^0\left(\overline{D}^0\right)$ modes and extraction of the Cabibbo-Kobayashi-Maskawa angle γ*, Phys. Rev. Lett. **78** (1997), no. 17 3257.

[43] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, *Determining gamma using $B^\pm \to DK^\pm$ with multibody D decays*, Phys. Rev. **D68** (2003) 054018, arXiv:hep-ph/0303187.

[44] I. Dunietz and R. G. Sachs, *Asymmetry between inclusive charmed and anticharmed modes in B^0, $\overline{B}^0$ decay as a measure of CP violation*, Phys. Rev. **D37** (1988) 3186.

[45] R. Aleksan, I. Dunietz, and B. Kayser, *Determining the CP violating phase gamma*, Z. Phys. **C54** (1992) 653.

[46] R. Fleischer, *New strategies to obtain insights into CP violation through $B_{(s)} \to D^\pm_{(s)}K^\mp$, $D^{*\pm}_{(s)}K^\mp$, ... and $B_{(d)} \to D^\pm\pi^\mp$, $D^{*\pm}\pi^\mp$, ... decays*, Nucl. Phys. **B671** (2003) 459, arXiv:hep-ph/0304027.

[47] BaBar collaboration, B. Aubert *et al.*, *Measurement of time-dependent CP-violating asymmetries and constraints on $\sin(2\beta+\gamma)$ with partial reconstruction of $B \to D^{*\mp}\pi^\pm$ decays*, Phys. Rev. **D71** (2005) 112003, arXiv:hep-ex/0504035.

[48] BaBar collaboration, B. Aubert *et al.*, *Measurement of time-dependent CP asymmetries in $B^0 \to D^{(*)\pm}\pi^\mp$ and $B^0 \to D^\pm\rho^\mp$ decays*, Phys. Rev. **D73** (2006) 111101, arXiv:hep-ex/0602049.

[49] Belle collaboration, F. J. Ronga *et al.*, *Measurement of CP violation in $B^0 \to D^{*-}\pi^+$ and $B^0 \to D^-\pi^+$ decays*, Phys. Rev. D **73** (2006) 092003.

[50] Belle Collaboration, S. Bahinipati *et al.*, *Measurements of time-dependent CP asymmetries in $B \to D^{*\mp}\pi^\pm$ decays using a partial reconstruction technique*, Phys. Rev. **D84** (2011) 021101, arXiv:1102.0888.

[51] M. A. Baak, *Measurement of CKM angle gamma with charmed B^0 meson decays*, PhD thesis, Vrije Universiteit Amsterdam, 2007, SLAC-R-858.

[52] J. Brod and J. Zupan, *The ultimate theoretical error on gamma from $B \to DK$ decays*, JHEP **1401** (2014) 051, arXiv:1308.5663.

[53] L. Evans *et al.*, *The LHC machine*, JINST **3** (2008) S08001.

[54] Linear collider LINAC2, web page: http://home.web.cern.ch/about/accelerators/linear-accelerator-2.

[55] BOOSTER, web page: http://home.web.cern.ch/about/accelerators/proton-synchrotron-booster.

[56] Proton Synchrotron, web page: http://home.web.cern.ch/about/accelerators/proton-synchrotron.

[57] Super Proton Synchrotron, web page: http://home.web.cern.ch/about/accelerators/super-proton-synchrotron.

[58] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[59] Accelerators system at CERN, web page: http://www.nature.com/nature/journal/v448/n7151/box/nature06077_BX2.html.

[60] CDF collaboration, web page: http://www-cdf.fnal.gov/.

[61] LHCb Collaboration, *$b\bar{b}$ production angle plots*, web page: http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[62] LHCb collaboration, R. Aaij *et al.*, *Precision luminosity measurements at LHCb* , JINST **9** (2014) P12005.

[63] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.

[64] P. Cherenkov, *Visible emission of clean liquids by action of γ radiation*, Doklady Akademii Nauk SSSR **2** (1934) 451.

[65] A. Powell *et al.*, *Particle identification at LHCb*, PoS **ICHEP2010** (2010) 020, LHCb-PROC-2011-008.

[66] LHCb Collaboration, *The layout of the LHCb trigger*, http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html.

[67] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, arXiv:1211.3055.

[68] V. V. Gligorov, *A single track HLT1 trigger*, LHCb-PUB-2011-003.

[69] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861.

[70] M. Williams *et al.*, *The HLT2 Topological Lines*, LHCb-PUB-2011-002.

[71] R. Kalman, *A new approach to linear filtering and prediction problems*, Journal of Basic Engineering **82** (1960) 35.

[72] LHCb, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015), no. 02 P02007, arXiv:1408.1251.

[73] LHCb, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015), no. 07 1530022, arXiv:1412.6352.

[74] J. W. Tukey, *Approximate weights*, Ann. Math. Statist. **19** (1948) 91.

[75] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[76] The Gaudi project, web page: http://proj-gaudi.web.cern.ch/proj-gaudi/.

[77] M. Clemencic *et al.*, *The LHCb simulation application,* Gauss*: design, evolution and experience*, J. of Phys. : Conf. Ser. **331** (2011) 032023.

[78] I. Belyaev *et al.*, *Handling of the generation of primary events in* Gauss*, the LHCb simulation framework*, Nuclear Science Symposium Conference Record (NSS/MIC) **IEEE** (2010) 1155.

[79] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[80] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, arXiv:hep-ph/0506026.

[81] GEANT4 collaboration, S. Agostinelli *et al.*, *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[82] GEANT4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[83] The Boole project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/.

[84] The Moore project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/.

[85] The Brunel project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/.

[86] The DaVinci project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/.

[87] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[88] The Urania project, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/urania/releases/v2r4/.

[89] The B2DXFitter package, web page: http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/urania/packages/B2DXFitters.

[90] LHCb collaboration, R. Aaij *et al.*, *Measurements of the branching fractions of the decays $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$*, JHEP **06** (2012) 115, arXiv:1204.1237.

[91] LHCb collaboration, *Updated average f_s/f_d b-hadron production fraction ratio for* 7 TeV *pp collisions*, LHCb-CONF-2013-011.

[92] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[93] Y. Xie, *sFit: a method for background subtraction in maximum likelihood fit*, arXiv:0905.0724.

[94] A. Hocker *et al.*, *TMVA - toolkit for multivariate data analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[95] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[96] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. **A543** (2005) 577, arXiv:physics/0408124.

[97] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[98] K. Cranmer, *Kernel Estimation in High Energy Physics.*, Comput. Phys. Commun. **136** (2001) 198.

[99] L. Landau, *On the energy loss of fast particles by ionization*, J. Phys. (USSR) **8** (1944) 201.

[100] S. Poisson, *Probabilite des jugements en matiere criminelle et en matiere civile, precedees des regles generales du calcul des probabilities.*, Bachelier (1837) 206.

[101] C. Gauss, *Theoria motvs corporvm coelestivm : in sectionibvs conicis solem ambientivm.*, Hambvrgi : Sumtibus F. Perthes et I. H. Besser, (1809).

[102] M. Dorigo *et al.*, *b-flavour tagging in pp collisions*, Nuclear Physics B (2014), Proceedings ICHEP 2014.

[103] G. A. Krocker, *Development and calibration of a same side kaon tagging algorithm and measurement of the B_s^0-$\bar{B}_s^0$ oscillation frequency Δm_s at the LHCb experiment*, PhD thesis, Heildelberg U., Sep, 2013, CERN-THESIS-2013-213.

[104] LHCb collaboration, *Optimization and calibration of the same-side kaon tagging algorithm using hadronic B_s^0 decays in 2011 data*, LHCb-CONF-2012-033.

[105] M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation*, arXiv:1407.0748.

[106] R. W. Lambert and F. Muheim, *LHCb Hybrid Photon Detectors and Sensitivity to Flavour Specific Asymmetry in Neutral B-Meson Mixing*, PhD thesis, Edinburgh U., Edinburgh, 2009, Presented on 07 Jan 2009.

[107] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, JHEP **07** (2014) 041, arXiv:1405.2797.

[108] G. Punzi, *Comments on Likelihood fits with variable resolution*, arXiv:physics/0401045, presented at PhyStat2003, Stanford, CA, USA, September 2003.

[109] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays*, Phys. Rev. **D87** (2013) 112010, arXiv:1304.2600.

[110] LHCb collaboration, R. Aaij *et al.*, *A measurement of the CKM angle γ from a combination of $B^\pm \to Dh^\pm$ analyses*, Phys. Lett. **B726** (2013) 151, arXiv:1305.2050.

[111] R. Fisher, *Statistical Methods for Research Workers*, Oliver & Boyd (1957), ISBN 0-05-002170-2.

[112] LHCb, B. Adeva *et al.*, *Roadmap for selected key measurements of LHCb*, arXiv:0912.4179.

[113] LHCb collaboration, *Measurement of time-dependent CP-violation observables in $B_s^0 \to D_s^\mp K^\pm$*, LHCb-CONF-2012-029. Linked to LHCb-ANA-2012-068.

[114] LHCb, R. Aaij *et al.*, *Implications of LHCb measurements and future prospects*, Eur. Phys. J. **C73** (2013), no. 4 2373, arXiv:1208.3355.

[115] LHCb Collaboration, R. Aaij *et al.*, *First observation and branching fraction for the $B_s^\pm \to D_s^{*\mp}K^\pm$* , LHCb-PAPER-2015-008, submitted to JHEP.

[116] LHCb, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \to D_s^\mp K^\pm$ decays*, JHEP **1411** (2014) 060, arXiv:1407.6127.

[117] 37th International Conference on High Energy Physics, http://ichep2014.es//.

[118] A. Dziurda, *Measurement of the CKM angle γ with $B_s^0 \to D_s^\mp K^\pm$ decays*, Nuclear Physics B Proceedings (2014), submitted.

[119] V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, Tech. Rep. CERN-LHCb-PUB-2011-016, CERN, Geneva, Sep, 2011. LHCb-PUB-2011-016.