

Measurement of time-dependent CP violation in $B^0 \rightarrow D^\mp\pi^\mp$ decays and optimisation of Flavour Tagging algorithms at LHCb

Thèse n. XXXX
présentée le XX xxxx 2018
à la Faculté des Sciences de Base
Laboratoire de physique des hautes énergies
Programme doctoral en Physique
École Polytechnique Fédérale de Lausanne



pour l'obtention du grade de Docteur ès Sciences
par

Vincenzo Battista

acceptée sur proposition du jury:
Prof. Olivier Schneider, directeur de thèse

Lausanne, EPFL, 2018

The ability to quote is a serviceable substitute for wit
W. S. Maugham

Abstract

Keywords: B physics, LHCb, LHC.

Résumé

Mots clefs : Physique du B , LHCb, LHC.

Sommario

Parole chiave: fisica del mesone B , LHCb, LHC.

Acknowledgements

Lausanne, XX XXXX 2018

V. B.

Contents

Abstract (English/Français/Italian)	v
Acknowledgements	xi
1 The Standard Model of particle physics and B physics	1
1.1 The Standard Model of particle physics	2
1.2 The Cabibbo-Kobayashi-Maskawa matrix	4
1.3 Physics of neutral B mesons	7
1.3.1 Oscillation of neutral mesons	7
1.3.2 Decay of neutral mesons	9
1.3.3 CP -violation in neutral meson systems	11
2 The Large Hadron Collider and the LHCb experiment	15
2.1 The Large Hadron Collider	16
2.2 The LHCb experiment	18
2.2.1 The tracking system	19
2.3 Particle Identification (PID)	22
2.3.1 The trigger system	25
2.3.2 Event reconstruction, simulation and software	27
2.3.3 Data collected by LHCb	28
3 Flavour Tagging algorithms and optimisation of the opposite-side electron tagger	31
3.1 Introduction to Flavour Tagging	32
3.1.1 Flavour tagging algorithms	32
3.2 Flavour tagging strategy for the $B^0 \rightarrow D^\pm \pi^\mp$ time-dependent analysis	38
3.2.1 Calibration of the Opposite Side Tagger Combination	40
3.2.2 Calibration of the Same Side Tagger Combination	43
4 Measurement of CP violation in $B^0 \rightarrow D^\mp \pi^\pm$ decays	47
4.1 Introduction	48
4.2 Data sample and selection	51
4.2.1 Stripping and trigger requirements	51
4.2.2 Preselection and samples definition	52

Contents

4.2.3	Vetoos for physics backgrounds	53
4.2.4	Wrongly associated PVs	56
4.2.5	Development of an MVA classifier	58
4.2.6	BDT selection optimisation	60
4.2.7	Multiple candidates	60
4.2.8	Selection performance	61
4.3	Simulation and expected sample composition	63
4.3.1	PIDK correction	63
4.3.2	Surviving physics backgrounds	64
4.4	Fits to the B^0 invariant mass	66
4.4.1	Probability density functions	67
4.4.2	Fit to data	68
4.4.3	<i>sWeight</i> calculation	72
4.4.4	Fit split by subsample	74
4.5	Time Resolution	78
4.5.1	Companion track momentum reweighting	78
4.5.2	Resolution determination from decay time error parameterisation .	79
4.6	Time-dependent selection efficiency	79
4.7	Decay time fit to data	82
4.7.1	Fit validation	85
4.8	Systematics	95
4.8.1	Systematic uncertainties from Gaussian constraints	95
4.8.2	Systematic uncertainties estimated with pseudoexperiments	96
Appendix		101
4.9	Opposite side tagging studies	102
4.9.1	Mass fit of $B^+ \rightarrow D^0\pi^+$	102
4.9.2	Reweighting of $B^+ \rightarrow D^0\pi^+$ to $B^0 \rightarrow D^\pm\pi^\mp$	103
4.9.3	GOF tests for OS calibration on $B^+ \rightarrow D^0\pi^+$ data	106
4.10	Selection studies	108
4.10.1	BDT input features	108
4.10.2	Multiple candidates	108
4.11	Particle Identification plots	112
4.12	Invariant mass fit studies	115
4.13	PDFs definitions	116
4.14	PDF for the decay-time fit	118
4.15	Correlation matrix of decay-time fit to $B^0 \rightarrow D^\mp\pi^\pm$	120
4.16	Decay-time fits in $B^0 \rightarrow D^\mp\pi^\pm$ data subsamples	121
4.17	Decay-time fit validation with bootstrapping	129
4.18	Test of the decay-time fit via a toy tagger	133
4.19	Inputs for $B^0 \rightarrow D^\pm\pi^\mp$ simulation	135
References	136

1 The Standard Model of particle physics and B physics

1.1 The Standard Model of particle physics

The *Standard Model* (SM) of particle physics [1, 2, 3] is a *non-abelian, Yang-Mills quantum field theory* based on the $SU(3) \times SU(2) \times U(1)$ gauge symmetry group. This model provides a coherent, unified and experimentally-established picture of electromagnetic, weak and strong interactions, as well as a description of the known elementary particles (quarks, leptons, gauge bosons and Higgs boson, Fig. 1.1).

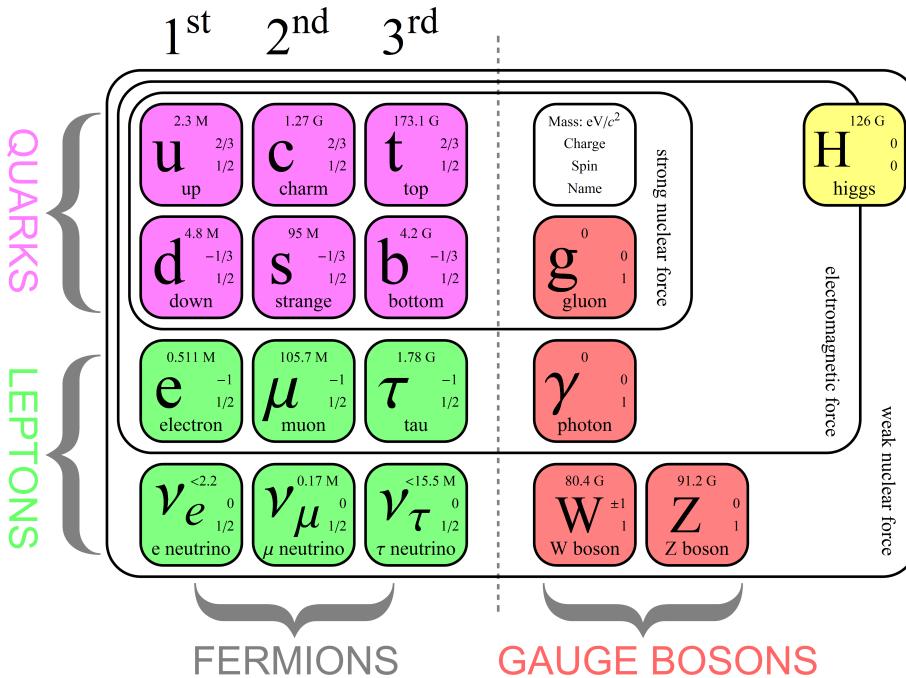


Figure 1.1 – Elementary particles described by the SM [4].

All particles are either *fermions* or *bosons*. The fermions (leptons, quarks) have half-integer spin and follow the Fermi-Dirac statistics [5, 6], whereas Bosons (gauge bosons, Higgs boson) have integer spin and follow the Bose-Einstein statistics [7].

Leptons (spin- $\frac{1}{2}$) include three charged ¹, massive particles (electron e^- , muon μ^- and tau τ^-), which interact via the electromagnetic and weak interactions, and three neutral, (nearly) massless particles, called *neutrinos* (ν_e , ν_μ and ν_τ), which only experience weak interactions.

Six different types of quarks (spin- $\frac{1}{2}$) exist: the *up-type* quarks up (u), charm (c) and top/truth (t), having charge $\frac{2}{3}$, and the *down-type* quarks down (d), strange (s), and bottom/beauty (b), which have charge $-\frac{1}{3}$. They can interact via electromagnetic, weak and strong interactions, and they are all massive.

¹Electric charge is always intended in units of the fundamental charge of the electron

- 20 The fundamental interactions are *mediated* by gauge bosons (spin-1). The photon (γ) is
21 responsible for the electromagnetic interaction, whereas Z^0 and W^\pm bosons are mediators
22 for the weak interaction. These two forces are considered to be different manifestations
23 of a single *electroweak* interaction, which is responsible for all electric and magnetic
24 phenomena as well as some radioactive decays. The strong interaction among quarks is
25 mediated by the gluons g . Photon and gluons are massless, whereas the weak force gauge
26 bosons have a non-zero mass.
- 27 For all the particles mentioned so far, an *antiparticle* exist. Each antiparticle has the
28 same mass of the corresponding particle, but opposite quantum numbers (electric charge,
29 lepton number etc...).
- 30 Quarks do not exist in a free state: they can only be bound inside *hadrons* via the
31 *confinement* mechanism, a feature of the strong interaction. A hadron can be composed
32 by a quark-antiquark pair (*meson*), or by three quarks or antiquarks (*baryons*). Examples
33 of mesons include the B^0 ($\bar{b}d$) and D^+ ($c\bar{d}$) mesons, whereas proton (uud) and neutron
34 (udd) are examples of baryons. Recently, there have been evidences for more complex
35 states (tetraquarks [8], pentaquarks [9]).
- 36 The non-zero mass of leptons, quarks and weak force gauge bosons would require a
37 gauge symmetry breaking term in the SM Lagrangian density. The *Brout-Englert-Higgs*
38 *mechanism* [10, 11, 12] introduces a scalar (spin-0) field, called Higgs field, and a potential
39 which allows the Higgs field to have a non-zero vacuum expectation value. This implies
40 that the Gauge symmetry is broken *dynamically*, and that the masses of the particles
41 arise from the resulting interaction with the Higgs field. The quantum of the Higgs field
42 is known as Higgs boson, the last SM particle to be discovered experimentally [13, 14].
- 43 All massive particles experience the fourth fundamental interaction, the gravitational
44 force, but this is described by another field theory, the General Relativity (GR), currently
45 not unified with the SM.
- 46 Any experimental signature that is not described by the SM would be a hint for *new*
47 *physics* (NP). Altought the SM is known to be an incomplete theory because of different
48 unsolved problems, such as dark matter, *naturalness*, matter-antimatter asymmetry, lack
49 of SM-GR unification etc..., no evidence for NP has been found so far.

⁵⁰ **1.2 The Cabibbo-Kobayashi-Maskawa matrix**

The lagrangian density describing the interactions between quarks and W^\pm (*charged current interaction*) can be written as follows:

$$\mathcal{L}_{cc} = \frac{g}{\sqrt{2}} \begin{pmatrix} \bar{u} & \bar{c} & \bar{t} \end{pmatrix} V_{CKM} \gamma^\mu \frac{(1 - \gamma^5)}{2} \begin{pmatrix} d \\ s \\ b \end{pmatrix} W_\mu^+ + h.c., \quad (1.1)$$

where g is a coupling constant, γ^μ are Dirac matrices and V_{CKM} , known as Cabibbo-Kobayashi-Maskawa (CKM) matrix [15, 16], couples the *flavour* eigenstates d , s and b to the *mass* (or *physical*) eigenstates d' , s' and b' :

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (1.2)$$

- ⁵¹ The CKM matrix is unitary ($V_{CKM}^\dagger V_{CKM} = 1$), so it can be written in terms of four
⁵² independent parameters, namely three angles and a *weak phase* δ . The latter is the
⁵³ source of all *CP-violating* phenomena in the SM, i.e. asymmetries between particles and
⁵⁴ anti-particles; in fact, the *complexity* of V_{CKM} implies that the SM lagrangian density is
⁵⁵ non *CP*-invariant, in agreement with the experimentally observed *CP*-violation.

A first, standard parameterisation of the CKM matrix [17] gives:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (1.3)$$

- ⁵⁶ where $s_{ij} = \sin \theta_{ij}$ and $c_{ij} = \cos \theta_{ij}$.

Another, useful parameterisation is given by *Wolfenstein* [18] and points out the order of magnitude of each matrix element. By defining the following quantities:

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \quad s_{23} = A\lambda^2 = A \left| \frac{V_{cb}}{V_{us}} \right|, \quad (1.4)$$

$$s_{13}e^{i\delta} = V_{ub}^* = A\lambda^3(\rho + i\eta), \quad (1.5)$$

the matrix V_{CKM} can be rewritten as a series expansion in powers of λ , given that λ is a small number:

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (1.6)$$

- 57 From 1.6, one can see that quark transitions within the same family (e.g. $u \rightarrow d$) are more
 58 probable, whereas transitions between different families (e.g. $b \rightarrow c$) are more suppressed.
 59 CP -violation is a consequence of $\eta \neq 0$.

The unitarity condition $V_{\text{CKM}}^\dagger V_{\text{CKM}} = 1$ can be rewritten in terms of six scalar equations. Two of them are particularly relevant for the b -hadrons phenomenology:

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

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0. \quad (1.8)$$

These two equations can be graphically represented as *triangles* in the $(\bar{\rho}, \bar{\eta})$ complex plane. Having defined the following angles:

$$\alpha = \phi_2 = \arg \left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right], \quad \beta = \phi_1 = \arg \left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right], \quad (1.9)$$

$$\gamma = \phi_3 = \arg \left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right], \quad \beta_s = \chi = \arg \left[-\frac{V_{cb}V_{cs}^*}{V_{tb}V_{ts}^*} \right]. \quad (1.10)$$

- 60 the triangles given by Eq. 1.7 and Eq. 1.8 can be depicted as shown in Fig. 1.2. The first
 61 triangle, defined by Eq. 1.7, is known as *Unitarity Triangle* (UT) and its elements can
 62 be measured from analyses of B^0 and B^\pm decays. The other triangle (Eq. 1.8) can be
 63 studied from decays of B_s^0 mesons. These measurements allow to test the CKM paradigm
 64 for CP -violation in the SM. Any deviation from unitarity would be a signature of new
 65 physics beyond the SM.

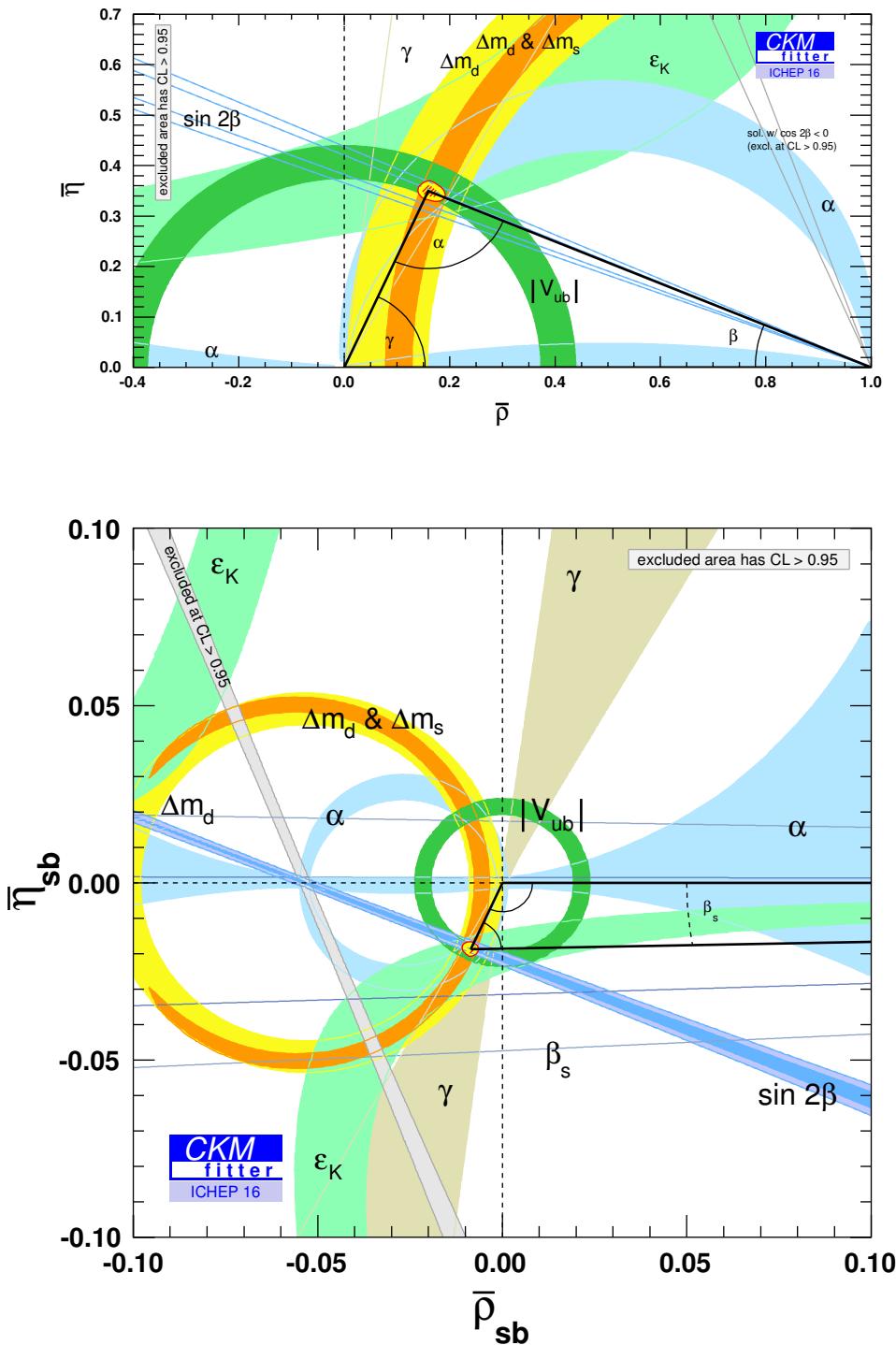


Figure 1.2 – Graphical representation of two of the six unitarity conditions of the CKM matrix, superimposed with the current experimental constraints [19].

66 1.3 Physics of neutral B mesons

67 The theory of neutral B meson oscillation, decays and CP -violation presented here is
 68 derived from Ref. [20] and [21].

69 1.3.1 Oscillation of neutral mesons

Neutral B meson states are characterised by the following quark content:

$$|B^0\rangle = |d\bar{b}\rangle, \quad |\bar{B}^0\rangle = |\bar{d}b\rangle, \quad (1.11)$$

$$|B_s^0\rangle = |s\bar{b}\rangle, \quad |\bar{B}_s^0\rangle = |\bar{s}b\rangle. \quad (1.12)$$

All neutral mesons will be denoted as P^0 or \bar{P}^0 hereafter. These states are eigenstates of the CP operator with eigenvalues \pm up to an arbitrary phase factor $e^{i\phi_{CP}}$. Since charged-currents do not conserve flavour quantum numbers (e.g. strangeness, beauty etc...), a neutral meson can transform itself into its own anti-meson, and viceversa. So, the time evolution of a neutral B meson system can be generally written as follows:

$$|\Psi(t)\rangle = a(t)|P^0\rangle + b(t)|\bar{P}^0\rangle + \sum_i c_i(t)|f_i\rangle, \quad (1.13)$$

70 where $|f_i\rangle$ are all the possible final state, and $c(0) = 0$ as initial condition.

Since the typical time scales of weak interactions are much longer than strong interaction timescales, we can neglect all weak interactions among final states (*Weisskopf-Wigner approximation*). So, we can write the Schroedinger equation for $|\Psi(t)\rangle$ in terms of an effective, non-hermitian hamiltonian \mathcal{H} :

$$i\partial_t \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathcal{H} \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}. \quad (1.14)$$

The \mathcal{H} matrix can be rewritten as the sum of two hermitian matrices M and Γ :

$$\mathcal{H} = M - \frac{i}{2}\Gamma = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}. \quad (1.15)$$

Assuming CPT-invariance ($H_{11} = H_{22} = H_0$, $M_{11} = M_{22} = M_0$, $\Gamma_{11} = \Gamma_{22} = \Gamma_0$), the

eigenvalues of \mathcal{H} are:

$$\lambda_L = m_L - \frac{i}{2}\Gamma_L = H_0 + \sqrt{H_{12}H_{21}} = H_0 + \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}, \quad (1.16)$$

$$\lambda_H = m_H - \frac{i}{2}\Gamma_H = H_0 - \sqrt{H_{12}H_{21}} = H_0 - \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}. \quad (1.17)$$

where L ("light") and H ("heavy") refer to the value of the mass for each eigenstate. The corresponding eigenvectors are:

$$|P_L\rangle = p|P^0\rangle + q|\bar{P}^0\rangle, \quad |P_H\rangle = p|P^0\rangle - q|\bar{P}^0\rangle \quad (1.18)$$

where p and q satisfy $|p|^2 + |q|^2 = 1$ and are given by:

$$\frac{q}{p} = \sqrt{\frac{H_{21}}{H_{12}}} = \sqrt{\frac{M_{21} - \frac{i}{2}\Gamma_{21}}{M_{12} - \frac{i}{2}\Gamma_{12}}} = \sqrt{\frac{M_{21}^* - \frac{i}{2}\Gamma_{21}^*}{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}}. \quad (1.19)$$

It can be shown that, assuming CP -violation in weak interactions and choosing $\phi_{CP} = \pi$, we have:

$$M_{12} = |M_{12}|e^{i\phi_M}, \quad \Gamma_{12} = |\Gamma_{12}|e^{i(\phi_L - \pi)}, \quad (1.20)$$

where ϕ_M and ϕ_Γ are extra phases (*mixing* and *decay* phases respectively) appearing because of CP -violation:

$$M_{12} = M_{12}^*e^{2i\phi_M}, \quad \Gamma_{12} = \Gamma_{12}^*e^{2i\phi_\Gamma}. \quad (1.21)$$

For the neutral B meson system, the ratio $|\Gamma_{12}/M_{12}|$ is expected to be small in the SM; as a consequence, it can be shown that

$$\frac{q}{p} = -e^{-i\phi_M} \left[1 - \frac{1}{2} \left| \frac{\Gamma_{12}}{M_{12}} \right| \sin(\phi_M - \phi_\Gamma) + \mathcal{O} \left(\left| \frac{\Gamma_{12}}{M_{12}} \right|^2 \right) \right], \quad (1.22)$$

⁷¹ where $\phi_M - \phi_\Gamma = 0 + \mathcal{O}(m_c^2/m_b^2)$.

The difference and the average of mass and width between mass eigenstates can be defined as:

$$\Delta m = m_H - m_L = \Re(\lambda_H - \lambda_L), \quad m = \frac{m_L + m_H}{2} = \frac{\Re(\lambda_H + \lambda_L)}{2}, \quad (1.23)$$

$$\Delta\Gamma = \Gamma_L - \Gamma_H = -2\Im(\lambda_L - \lambda_H), \quad \Gamma = \frac{\Gamma_L + \Gamma_H}{2} = -\frac{\Im(\lambda_H + \lambda_L)}{4}. \quad (1.24)$$

- ⁷² The sign convention for $\Delta\Gamma$ is chosen to have a positive-definite experimental value for
⁷³ the B_s^0 system (for B^0 , experiments give a result compatible with zero, in agreement with
⁷⁴ the SM).

The time evolution of the states $|P^0(t)\rangle$ and $|\bar{P}^0(t)\rangle$ when they are initially produced as $|P^0(0)\rangle$ and $|\bar{P}^0(0)\rangle$ can be obtained from the effective hamiltonian:

$$|P^0(t)\rangle = g_+(t)|P^0(t)\rangle + \frac{q}{p}g_-(t)|\bar{P}^0(t)\rangle, \quad (1.25)$$

$$|\bar{P}^0(t)\rangle = g_+(t)|\bar{P}^0(t)\rangle + \frac{p}{q}g_-(t)|P^0(t)\rangle. \quad (1.26)$$

$$(1.27)$$

The functions $g_{\pm}(t)$ are built in terms of the eigenvalues:

$$g_{\pm} = \frac{1}{2} \left(e^{-im_H t - \frac{1}{2}\Gamma_H t} \pm e^{-im_L t - \frac{1}{2}\Gamma_L t} \right). \quad (1.28)$$

The probabilities that a state initially produced as $|P^0(0)\rangle$ or $|\bar{P}^0(0)\rangle$ becomes a $|P^0(t)\rangle$ or $|\bar{P}^0(t)\rangle$ are the following:

$$|\langle P^0(0)|P^0(t)|\rangle|^2 = |g_+(t)|^2 = \frac{e^{-\Gamma t}}{2} \left(\cosh \frac{\Delta\Gamma t}{2} + \cos \Delta m t \right), \quad (1.29)$$

$$|\langle \bar{P}^0(0)|P^0(t)|\rangle|^2 = \left| \frac{q}{p} \right|^2 |g_-(t)|^2 = \left| \frac{q}{p} \right|^2 \frac{e^{-\Gamma t}}{2} \left(\cosh \frac{\Delta\Gamma t}{2} - \cos \Delta m t \right), \quad (1.30)$$

$$|\langle P^0(0)|\bar{P}^0(t)|\rangle|^2 = \left| \frac{p}{q} \right|^2 |g_-(t)|^2 = \left| \frac{p}{q} \right|^2 \frac{e^{-\Gamma t}}{2} \left(\cosh \frac{\Delta\Gamma t}{2} - \cos \Delta m t \right), \quad (1.31)$$

$$|\langle \bar{P}^0(0)|\bar{P}^0(t)|\rangle|^2 = |g_+(t)|^2 = \frac{e^{-\Gamma t}}{2} \left(\cosh \frac{\Delta\Gamma t}{2} + \cos \Delta m t \right). \quad (1.32)$$

- ⁷⁵ The equations above describe the *oscillation* or *mixing* of two neutral mesons.

1.3.2 Decay of neutral mesons

The amplitude for the decay of a neutral meson into a final state f can be obtained from the effective hamiltonian \mathcal{H} :

$$A_f = \langle f | \mathcal{H} | P^0 \rangle, \quad \bar{A}_f = \langle f | \mathcal{H} | \bar{P}^0 \rangle, \quad (1.33)$$

$$A_{\bar{f}} = \langle \bar{f} | \mathcal{H} | P^0 \rangle, \quad \bar{A}_{\bar{f}} = \langle \bar{f} | \mathcal{H} | \bar{P}^0 \rangle. \quad (1.34)$$

After defining the following parameters:

$$\lambda_f = \frac{q}{p} \frac{\bar{A}_f}{A_f}, \quad \frac{1}{\bar{\lambda}_{\bar{f}}} = \lambda_{\bar{f}} = \frac{q}{p} \frac{\bar{A}_{\bar{f}}}{A_{\bar{f}}}, \quad (1.35)$$

is it possible to write the *decay rates* for neutral mesons decaying into f or \bar{f} :

$$\begin{aligned} \frac{d\Gamma(P^0 \rightarrow f)}{dt}(t) = N_f |A_f|^2 \frac{1 + |\lambda_f|^2}{2} e^{-\Gamma t} \\ \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_f \sinh\left(\frac{\Delta\Gamma t}{2}\right) + C_f \cos(\Delta m t) - S_f \sin(\Delta m t) \right], \end{aligned} \quad (1.36)$$

$$\begin{aligned} \frac{d\Gamma(\bar{P}^0 \rightarrow f)}{dt}(t) = N_f |A_f|^2 \left| \frac{p}{q} \right|^2 \frac{1 + |\lambda_f|^2}{2} e^{-\Gamma t} \\ \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_f \sinh\left(\frac{\Delta\Gamma t}{2}\right) - C_f \cos(\Delta m t) + S_f \sin(\Delta m t) \right], \end{aligned} \quad (1.37)$$

$$\begin{aligned} \frac{d\Gamma(P^0 \rightarrow \bar{f})}{dt}(t) = N_f |\bar{A}_{\bar{f}}|^2 \left| \frac{q}{p} \right|^2 \frac{1 + |\bar{\lambda}_{\bar{f}}|^2}{2} e^{-\Gamma t} \\ \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma t}{2}\right) - C_{\bar{f}} \cos(\Delta m t) + S_{\bar{f}} \sin(\Delta m t) \right], \end{aligned} \quad (1.38)$$

$$\begin{aligned} \frac{d\Gamma(\bar{P}^0 \rightarrow \bar{f})}{dt}(t) = N_f |\bar{A}_{\bar{f}}|^2 \frac{1 + |\bar{\lambda}_{\bar{f}}|^2}{2} e^{-\Gamma t} \\ \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma t}{2}\right) + C_{\bar{f}} \cos(\Delta m t) - S_{\bar{f}} \sin(\Delta m t) \right]. \end{aligned} \quad (1.39)$$

where N_f is a time-independent normalisation factor and

$$D_f = -\frac{2\Re\lambda_f}{1 + |\lambda_f|^2}, \quad C_f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, \quad S_f = \frac{2\Im\lambda_f}{1 + |\lambda_f|^2}, \quad (1.40)$$

$$D_{\bar{f}} = -\frac{2\Re\bar{\lambda}_{\bar{f}}}{1 + |\bar{\lambda}_{\bar{f}}|^2}, \quad C_{\bar{f}} = -\frac{1 - |\bar{\lambda}_{\bar{f}}|^2}{1 + |\bar{\lambda}_{\bar{f}}|^2}, \quad S_{\bar{f}} = -\frac{2\Im\bar{\lambda}_{\bar{f}}}{1 + |\bar{\lambda}_{\bar{f}}|^2}. \quad (1.41)$$

⁷⁷ are known as *CP-coefficients*.

⁷⁸ 1.3.3 CP -violation in neutral meson systems

⁷⁹ Three types of CP -violation can occur. They are briefly sketched in Fig. 1.3 and described
⁸⁰ in the following paragraphs.

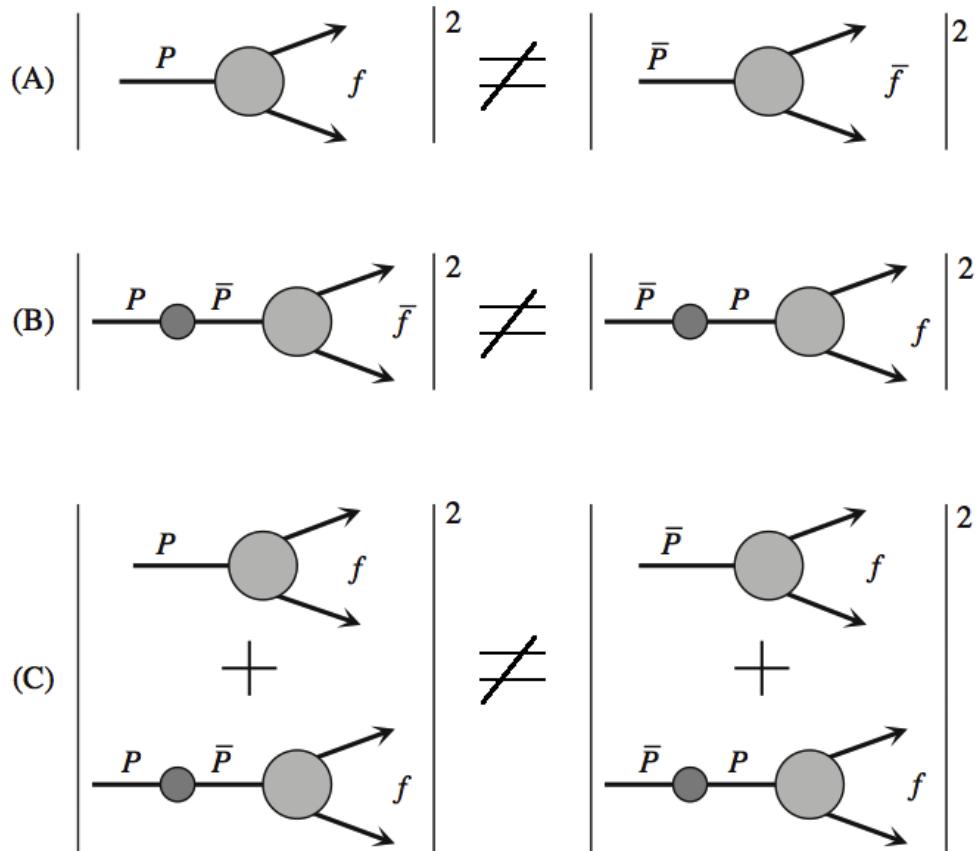


Figure 1.3 – Graphical representation of CP -violation in decay (A), mixing (B) and interference between mixing and decay (C) [21].

⁸¹ CP -violation in decays

CP -violation in decays, also known as *direct* CP -violation, happens when the decay amplitude for $P \rightarrow f$ is different from the CP -conjugated process $\bar{P} \rightarrow \bar{f}$:

$$\left| \frac{\bar{A}_{\bar{f}}}{A_f} \right| \neq 1. \quad (1.42)$$

This kind of CP -violation occurs if, for each decay, at least two amplitudes with different weak (ϕ_j) and strong (δ_j) phases contribute:

$$A_f = \sum_j |A_j| e^{i(\delta_j + \phi_j)}, \quad \bar{A}_{\bar{f}} = \sum_j |\bar{A}_j| e^{i(\delta_j - \phi_j)}. \quad (1.43)$$

In fact, the strong phases are invariant under CP -conjugation, whereas the weak phases get the opposite sign. The following asymmetry between final states can be measured to determine direct CP -violation experimentally for *charged mesons*, where mixing effects are absent:

$$\mathcal{A}_{f^\pm} = \frac{\Gamma(P^- \rightarrow f^-) - \Gamma(P^+ \rightarrow f^+)}{\Gamma(P^- \rightarrow f^-) + \Gamma(P^+ \rightarrow f^+)} = \frac{\left| \frac{\bar{A}_{\bar{f}}}{A_f} \right|^2 - 1}{\left| \frac{\bar{A}_{\bar{f}}}{A_f} \right|^2 + 1} \quad (1.44)$$

82 CP -violation in mixing

CP -violation in mixing, also called *indirect CP* violation, occurs when the oscillation rate for $\bar{P}^0 \rightarrow P^0$ is different from the CP -conjugated process $P^0 \rightarrow \bar{P}^0$. These two oscillation probabilities are given by Eq. 1.30 and Eq. 1.31 respectively. It turns out that they are identical unless

$$\left| \frac{q}{p} \right| \neq 1. \quad (1.45)$$

From Eq. 1.22, it can be seen that CP -violation in mixing occurs when the relative phase $\phi_M - \phi_\Gamma$ is different from any multiple of π . It is possible to measure the $|q/p|$ ratio by comparing the oscillation rates in semileptonic decays of neutral mesons $P^0 \rightarrow l^+ X$ and $\bar{P}^0 \rightarrow l^- X$, where no direct CP -violation occurs. The decays where oscillation occurred are identified by reconstructing "wrong sign" leptons:

$$\mathcal{A}_{SL} = \frac{\frac{d\Gamma(\bar{P}^0 \rightarrow l^+ X)}{dt} - \frac{d\Gamma(P^0 \rightarrow l^- X)}{dt}}{\frac{d\Gamma(\bar{P}^0 \rightarrow l^+ X)}{dt} + \frac{d\Gamma(P^0 \rightarrow l^- X)}{dt}} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \quad (1.46)$$

83 CP -violation in the interference between mixing and decay

This type of decay occurs when a neutral meson can decay directly to a given final state, $P^0 \rightarrow f$, or via a mixing, $P^0 \rightarrow \bar{P}^0 \rightarrow f$. This can happen only if the final state f is common to both P^0 and \bar{P}^0 . This type of CP -violation can occur also if other sources of CP -violation (mixing or decay) are absent. In general, the interference between mixing

and decay can be accessed by studying the following asymmetries:

$$\mathcal{A}_f(t) = \frac{\frac{d\Gamma(P^0 \rightarrow f)}{dt} - \frac{d\Gamma(\bar{P}^0 \rightarrow f)}{dt}}{\frac{d\Gamma(P^0 \rightarrow f)}{dt} + \frac{d\Gamma(\bar{P}^0 \rightarrow f)}{dt}} \quad (1.47)$$

$$\mathcal{A}_{\bar{f}}(t) = \frac{\frac{d\Gamma(P^0 \rightarrow \bar{f})}{dt} - \frac{d\Gamma(\bar{P}^0 \rightarrow \bar{f})}{dt}}{\frac{d\Gamma(P^0 \rightarrow \bar{f})}{dt} + \frac{d\Gamma(\bar{P}^0 \rightarrow \bar{f})}{dt}} \quad (1.48)$$

A relevant example is the case of neutral B mesons, where $|q/p| = 1$. Using Eq. 1.36, 1.37, 1.38 and 1.39, the asymmetries 1.47 and 1.48 take the following forms:

$$\mathcal{A}_f(t) = \frac{C_f \cos(\Delta mt) - S_f \sin(\Delta mt)}{\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_f \sinh\left(\frac{\Delta\Gamma t}{2}\right)}, \quad (1.49)$$

$$\mathcal{A}_{\bar{f}}(t) = \frac{-C_{\bar{f}} \cos(\Delta mt) + S_{\bar{f}} \sin(\Delta mt)}{\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma t}{2}\right)}. \quad (1.50)$$

⁸⁴ So, the CP -coefficients can be directly measured from a time-dependent analysis of the B decays.

2 The Large Hadron Collider and the ⁸⁷ LHCb experiment

88 2.1 The Large Hadron Collider

89 The *Large Hadron Collider* (LHC) is a circular collider with a circumference of 26.66 Km.
90 It's located at CERN, near Geneva, between Switzerland and France. The LHC is
91 designed to produce proton-proton (pp) collisions with a *luminosity* of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and
92 a center of mass energy of 14 TeV. In the first data taking period before the first long
93 shutdown, called Run 1 (2010-2012), the center of mass energy reached 7 TeV (2010-2011)
94 and 8 TeV (2012).

95 The proton bunches, produced from hydrogen gas and made up with 10^{11} protons, pass
96 through different intermediate accelerating stages (Fig. 2.1):

- 97 • LINAC 2 (50 MeV);
98 • Proton Synchrotron Booster (1.4 GeV);
99 • Proton Synchrotron (25 GeV);
100 • Super Proton Synchroton (450 GeV).

101 Finally, they are injected clockwise and counter-clock wise into LHC and accelerated to
102 their final energy. At LHC, in addition to LHCb, there are two general-purpose detectors
103 (ATLAS and CMS), a detector dedicated to quark matter and quark-gluon plasma physics
104 (ALICE) and other smaller experiments (TOTEM, LHCf, MoEDAL) dedicated to different
105 topics.

106 The LHC can also accelerate particles other than protons, such as lead or xenon, in order
107 to collect data samples for specific studies.

2.1. The Large Hadron Collider

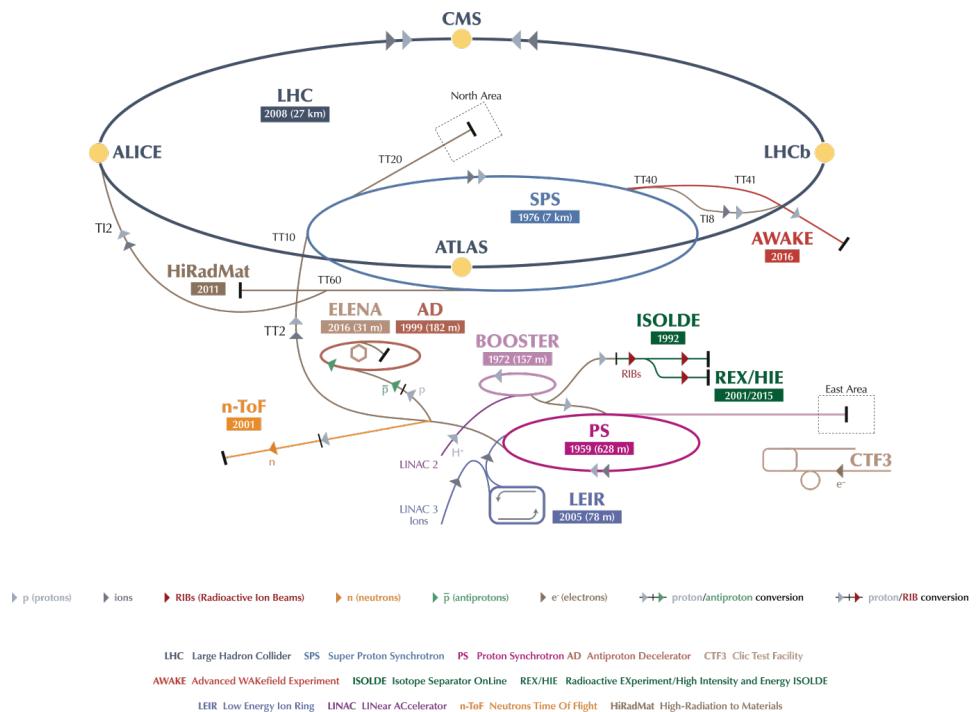


Figure 2.1 – Overview of the CERN accelerators complex.

¹⁰⁸

2.2 The LHCb experiment

¹⁰⁹ The *Large Hadron Collider beauty* (LHCb) experiment [22] is a single-arm forward
¹¹⁰ spectrometer (see Fig. 2.3) that exploits the forward production of the b - and \bar{c} - quarks
¹¹¹ pairs from pp collisions (Fig. 2.2). The LHCb angular coverage is comprised between
¹¹² 15 mrad and 250 (300) mrad in the vertical (horizontal) plane. The LHCb coordinate
¹¹³ system consists of an orthogonal set of coordinates, x , y , z , where z is aligned with the
¹¹⁴ beam pipe, y is perpendicular to the LHCb cavern ground and x is orthogonal to the
¹¹⁵ other two.

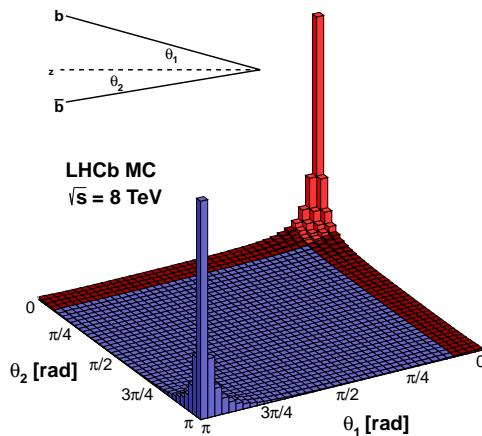


Figure 2.2 – b and \bar{b} pair simulated production angles with respect to the beam direction at $\sqrt{s} = 8\text{TeV}$.

¹¹⁶ The LHCb experiment is composed by different sub-detectors. The tracking system
¹¹⁷ includes a vertex and tracking detector called *VErtex LOcator* (VELO), the *Tracker*
¹¹⁸ *Turicensis* (TT), located upstream a magnetic dipole with an integrated field of 4 Tm, the
¹¹⁹ *Inner Tracker* (IT), situated downstream the magnet in three separated stations around
¹²⁰ the beryllium beam-pipe, and the *Outer Tracker* (OT), installed in the same stations of
¹²¹ the IT. The *Particle IDentification* (PID) system comprises two *Ring Imaging CHerenkov*
¹²² detectors (RICH), an *Electromagnetic CALorimeter* (ECAL), which also includes a *Pre-*
¹²³ *Showr* (PS) and *Scintillator Pad Detector* (SPD), a *Hadronic CALorimeter* (HCAL)
¹²⁴ and five *muon detectors* (M1-M5).

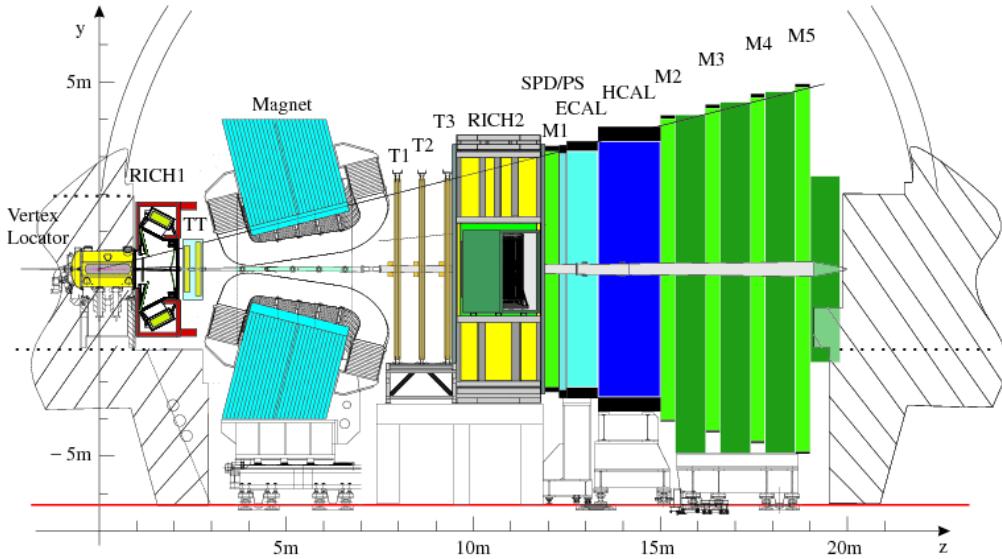


Figure 2.3 – Side view of the LHCb detector.

¹²⁵ 2.2.1 The tracking system

¹²⁶ The VErtex LOcator (VELO)

¹²⁷ The VELO [23] (Fig. 2.4) is a silicon micro-strip detector for charged particles surrounding
¹²⁸ the interaction point which identifies decay vertices and performs the first track recon-
¹²⁹ struction step. The sensitive region of the VELO is composed of n-on-n silicon micro-strip
¹³⁰ half-disk sensors with two different read-out strip geometries, called r -type and ϕ -type,
¹³¹ which measure the radial (r) and angular (ϕ) position in polar coordinates. The VELO
¹³² consists of 21 stations. Each station has two independent halves that can be moved apart
¹³³ during the beam injection period and then closed again when the beam orbit is stabilised.
¹³⁴ Each half-station is composed by one r -type and one ϕ -type sensor. The total length of
¹³⁵ the VELO detector is about 1 m.

¹³⁶ The silicon sensors, which provide an excellent measurement of vertex positions and
¹³⁷ impact parameters (IP), are 8.4 cm in diameter and have an inner hole with radius
¹³⁸ 0.8 cm. The IP resolution of a track is measured to be $\sigma_{\text{IP}} = 11.6 \pm 23.4/p_T \mu\text{m}$ in x and
¹³⁹ $\sigma_{\text{IP}} = 11.2 \pm 23.2/p_T \mu\text{m}$ in y , where p_T is the *transverse momentum* (in GeV/c) of the
¹⁴⁰ particle with respect to the beam axis. The strip pitch ranges from 38 to 108 μm (38 to
¹⁴¹ 97 μm) for r (ϕ) sensors, while the sensor thickness is 300 μm .

¹⁴² The Tracker Turicensis (TT)

¹⁴³ The TT [24] (Fig. 2.5) is a silicon micro-strip detector covering a total area of about
¹⁴⁴ 7.9 m^2 upstream the magnet and divided into two separate stations (TTa, TTb). Each

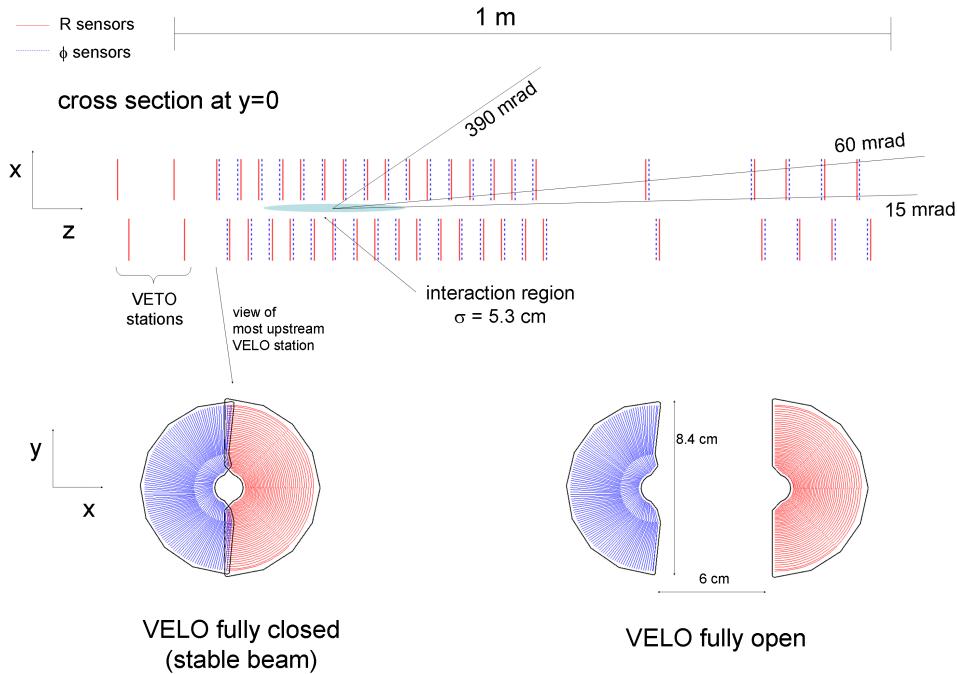


Figure 2.4 – Schematic overview of the VELO detector and its sensors.

station has two layers. The TT helps in improving the track momentum resolution and detecting long-lived particles that decay outside the VELO acceptance. TTa is composed by X-U layers, while TTb includes V-X layers. The X layers have read-out strips aligned vertically, whereas the U and V *stereo* layers are rotated by $+5^\circ$ and -5° with respect to the vertical in the xy plane, respectively.

The TT active area is made of p-on-n silicon micro-strip sensors. Since the sensors are exposed to a significant radiation due to a high track multiplicity, they are cooled to 0° in order to minimise the damage.

The TT read-out module contain from one to four sensors in a row, resulting in read-out strips up to 37 cm long. The strip pitch is 183 μm and the sensor thickness is 500 μm . The hit resolution is about 50 μm .

156 The Inner Tracker (IT)

The IT [25] (Fig. 2.6) is also a silicon micro-strip detector. Together with the TT, it's a part of the *Silicon Tracker* (ST). It's dedicated to detect charged particles in the high track density region around the beam-pipe downstream the magnet. It's separated into three stations, where each station consists of four boxes. Each box has four layers made

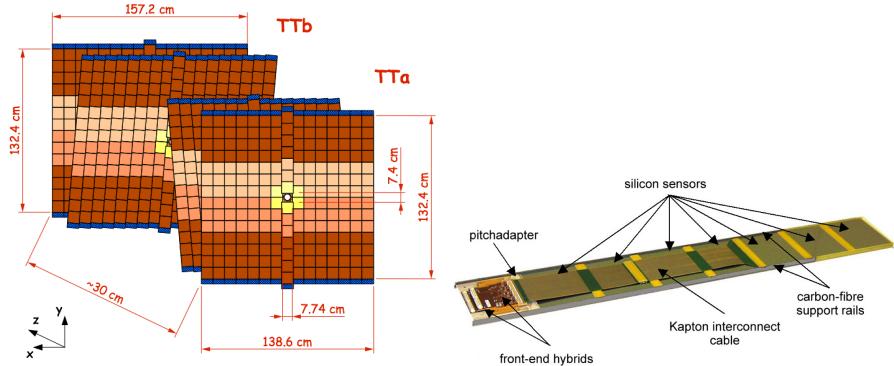


Figure 2.5 – Schematic overview of the TT stations/layers and one of the TT read-out modules.

of seven read-out modules arranged in a X1-U-V-X2 layout similar to that of the TT.
The total coverage of the IT is about 4.2 m^2 . The boxes directly above and below the beam-pipe are made of single-sensor modules, called *short modules*, whereas the remaining boxes are made of two bonded silicon sensor modules, called *long modules*. The IT strip pitch is $198 \mu\text{m}$, while the p-on-n sensor thickness is 320 (410) μm for the short (long) modules. The hit resolution is about $50 \mu\text{m}$.

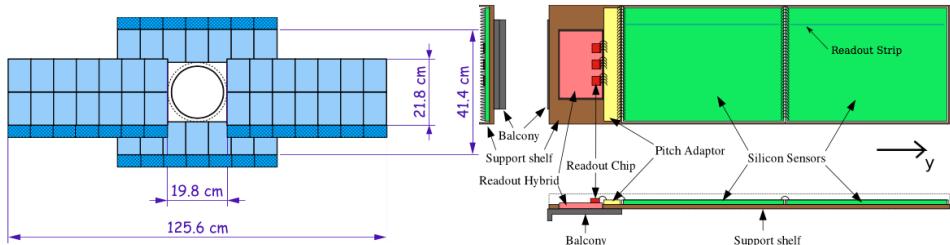


Figure 2.6 – Schematic overview of an IT station and one of the long IT read-out modules.

167 The Outer Tracker (OT)

The OT [26] (Fig. 2.8) is a gaseous straw-tube detector filled with an Ar/CO₂/O₂ (70%/28.5%/1.5%) gas mixture. It's dedicated to the detection of charged particles in the low track density region outside the IT acceptance and covers a large area of about 340 m^2 . The OT is composed by three stations, where each station has four layers in a X1-U-V-X2 configuration. Each station is separated physically for the left and right sides with respect to the beam-pipe into two C-frames (X1-U) and (V-X2). Each layer is divided into two mono-layers. The OT has different types of modules, the long F modules and the S1, S2, S3 modules that are cut in two pieces to leave space for the IT. The straw tube and anode wire diameters are 5 mm and $25 \mu\text{m}$ respectively. The hit resolution is

177 about 200 ?m.

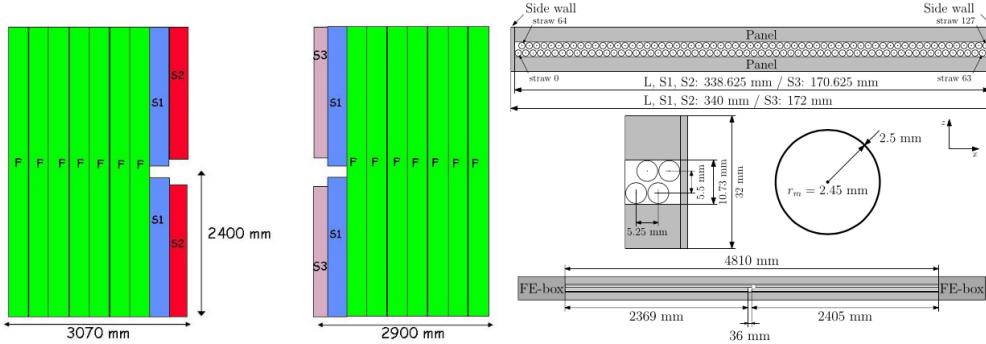


Figure 2.7 – Schematic overview of an OT layer and an OT module layout.

178 Spillover noise in the Silicon Tracker

179 Starting from 2015, the time spacing between proton *bunches* in LHC collision is 25 ns,
 180 half of the value adopted before. This had a direct impact on the front-end electronics
 181 of silicon detectors (VELO, TT, IT), because the width in time of the analogue signal
 182 produced by the front-end electronics is of the same order of magnitude. This means that
 183 it is possible to still have a non negligible amount of signal in the subsequent collision,
 184 which can be misidentified as coming from particles produced in that event. This source
 185 of noise is called *spillover*. The starting seeds of tracking algorithms, called *clusters*,
 186 can be polluted by spillover clusters which may increase the number of fake (or *ghost*)
 187 reconstructed tracks.

188 In the first part of my Phd activity, I studied the effect of spillover clusters in the ST
 189 using both simulated events and real collision data. This study showed that this time
 190 spacing has a little impact on the detector *occupancy*, and that the increase of ghost
 191 tracks is negligible. Moreover, in the same study it was showed that the charge deposited
 192 by particles in the detector can be exploited as a *feature* in multivariate analyses in order
 193 to further reduce the ghost track contamination. These results were documented in an
 194 internal note [27].

195 2.3 Particle Identification (PID)

196 The Ring Imaging Cherenkov (RICH) detectors

197 When a charged particle is travelling faster than the speed of light in a medium, Cherenkov
 198 light is produced at an angle that depends on the velocity of the particle and the refractive
 199 index of the medium (*radiator*). By knowing the momentum from the tracker and the
 200 velocity from the RICH detectors, the mass can be determined and therefore provide

2.3. Particle Identification (PID)

201 particle identification. Two RICH detectors [28] (Fig. 2.9) are used in order to provide
 202 PID in different momentum ranges.

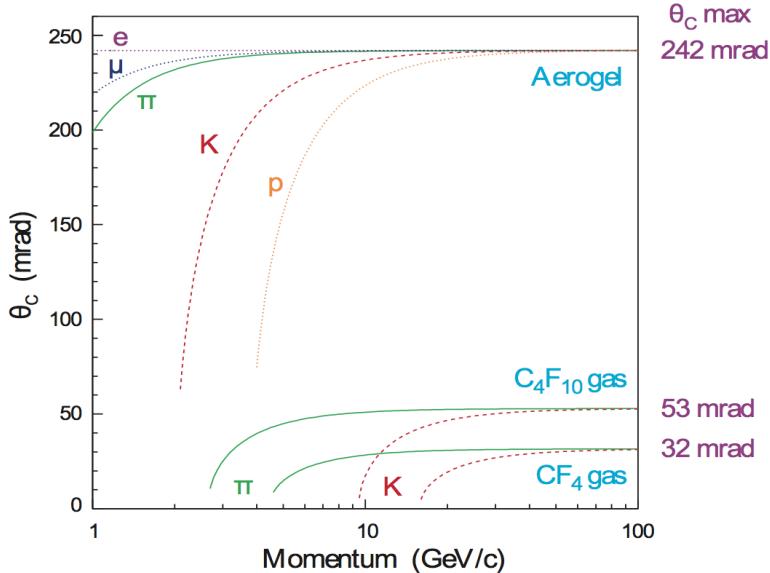


Figure 2.8 – Cherenkov angle as a function of the momentum for different particles and media.

203 RICH1 is responsible for providing PID in the momentum range from 1 to 60 GeV/c. The
 204 angular acceptance ranges from 25 mrad to 50 (300) mrad in the vertical (horizontal)
 205 plane. The adopted radiator is a mixture of aerogel and fluorobutane (C_4F_{10}). RICH1 is
 206 located between the VELO and the TT. The Cherenkov photons are guided to Hybrid
 207 Photon Detectors (HPD) via dedicated mirrors.

208 RICH2 is optimised for the momentum range from 15 to 100 GeV/c. The angular accep-
 209 tance ranges from 15 mrad to 100 (120) mrad in the vertical (horizontal) plane where
 210 most of the high-momentum tracks are produced. RICH2 uses tetrafluoromethane (CF_4)
 211 as radiator.

212 The Electromagnetic CALorimeter (ECAL), Pre-Shower (PS) and Scintilla- 213 tor Pad Detector (SPD)

214 The ECAL [29] is used for the detection and measurement of the energy of electrons and
 215 photons. The ECAL is built as a sandwich of alternating scintillators and lead layers in
 216 the xy plane. Scintillation light produced by the shower of particles generated by the
 217 lead plates is read out by Wave-Length Shifter (WLS) fibres coupled to PhotoMultiplier
 218 Tubes (PMTs). The SPD is installed upstream the ECAL to separate electrons from
 219 photons. The PS is installed between the SPD and the ECAL. Both SPD and PS use
 220 scintillator pads read out by WLS fibres coupled to MultiAnode PhotoMultiplier Tubes

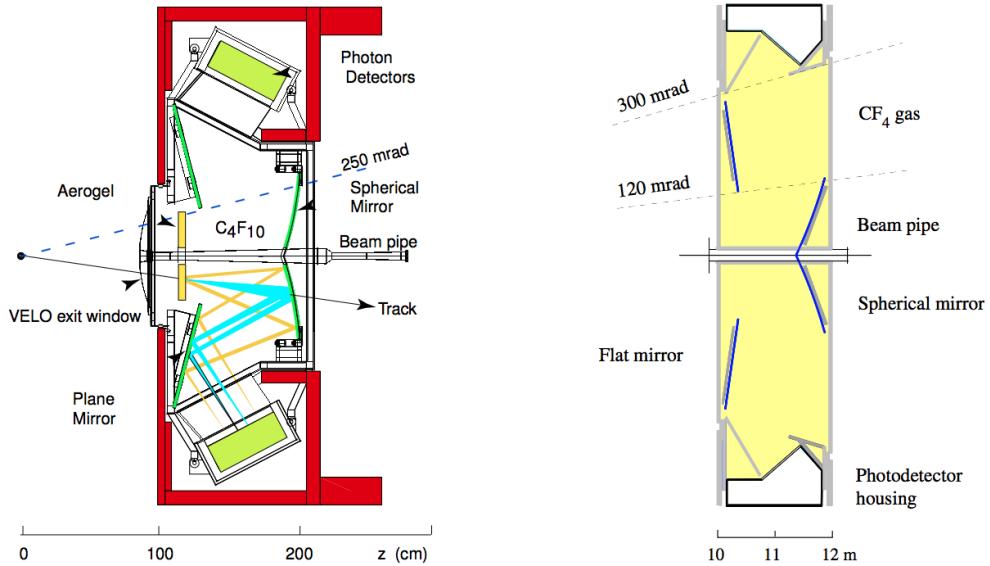


Figure 2.9 – Side view of RICH1 (left) and top view of RICH2 (right).

221 (MAPMT). The acceptance range of the ECAL is from 25 mrad up to 300 (250) mrad in
 222 the horizontal (vertical) plane. The relative energy resolution of the ECAL is given by
 223 $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$, where E is given in GeV.

224 The Hadronic CALorimeter (HCAL)

225 The HCAL [29] is used for the detection and measurement of the energy of hadrons (such
 226 as pions and kaons) for the first level trigger. A HCAL cell is a sampling device made
 227 of alternating iron and scintillator tiles, where the latters are located along the beam
 228 direction. The HCAL has the same acceptance coverage of the ECAL. The relative energy
 229 resolution of the HCAL is given by $\sigma_E/E = (69 \pm 5\%)/\sqrt{E} \oplus (9 \pm 2)\%$, where E is given
 230 in GeV.

231 Muon detector

232 The muon system [30] (Fig. 2.10) is a gaseous detector divided in five stations (M1 to M5)
 233 interleaved by 80 cm thick iron filters. The gaseous detectors are Multi-Wire Proportional
 234 Chambers (MWPC), except for the innermost part of M1, where triple GEM detectors
 235 are used to cope with the higher track density. The angular acceptance ranges from 20
 236 (16) to 308 (256) mrad in the horizontal (vertical) plane. The muon detector has 1380
 237 chambers and covers a total area of 435 m². Each muon chamber is composed by four
 238 layers of MPWC, except for M1, where two layers are used. The hit efficiency of the
 239 chambers is higher than 99% and allows to achieve a trigger efficiency greater than 95%

240 for muons. The adopted gas mixture ($\text{Ar}/\text{CO}_2/\text{CF}_4$ 40/55/5%) allows a fast triggering
 241 on muons (40 MHz).

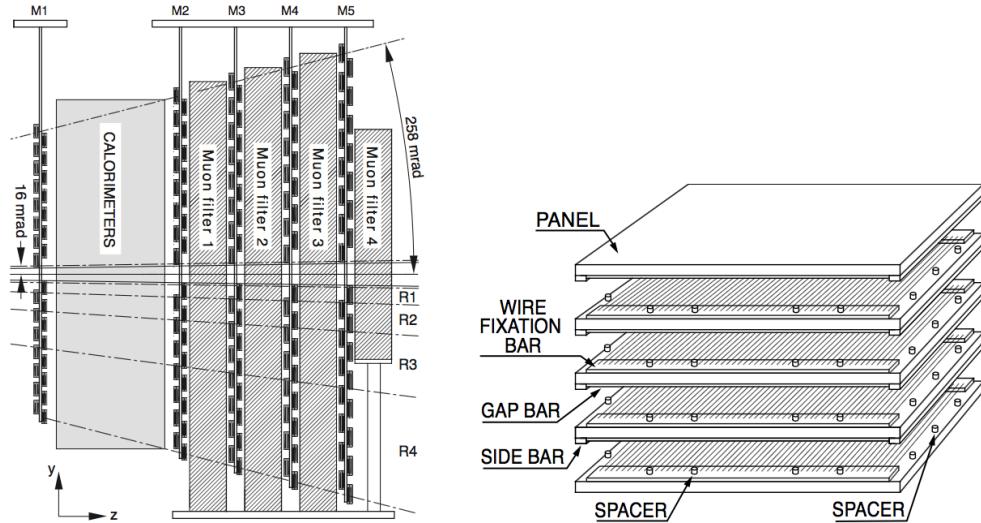


Figure 2.10 – Schematic overview of the muon system and a MWPC.

242 2.3.1 The trigger system

243 The collision rate of LHC is very high (40 MHz) because more than 99% of the $p\bar{p}$ collisions
 244 do not produce interesting events. It's not possible to record events with such a high
 245 rate: therefore, a trigger system [31] is required to reduce the rate from 40 MHz down to
 246 few kHz. The rate reduction is achieved via selection criteria which ensure that events
 247 containing heavy flavour decays are stored. The signatures of these interesting decays
 248 include high p_T and transverse energy of the decay products, as well as displaced decay
 249 vertices (with large IP) due to the relative long lifetimes of b - and c -hadrons. The trigger
 250 is divided into two sequential stages: a hardware stage called Level-0 (L0), and a software
 251 stage called the High Level Trigger (HLT). Different trigger decisions are separated into
 252 various *lines*, each of which provides informations on different physics process (e.g. decay
 253 topology, presence of muons etc...). All the trigger steps are summarised in Fig. 2.11.

254 Two types of trigger response are assigned offline, when some physics channel is analysed.
 255 The TOS (*Trigger On Signal*) trigger occurs when the event contains the user-defined
 256 signal channel. The TIS (*Trigger Independent of Signal*) trigger response is given regardless
 257 of the presence of the signal in that event.

258 After the trigger stage, the data go through further offline selection steps, where exclusive
 259 (e.g. $B^0 \rightarrow D^-\pi^+$, $B^+ \rightarrow J/\psi K^+$) and inclusive (e.g. $J/\psi \rightarrow \mu^+\mu^-$) decays are
 260 reconstructed. This offline selection step is known as *stripping*, and each set of selection
 261 requirement is called *stripping line*.

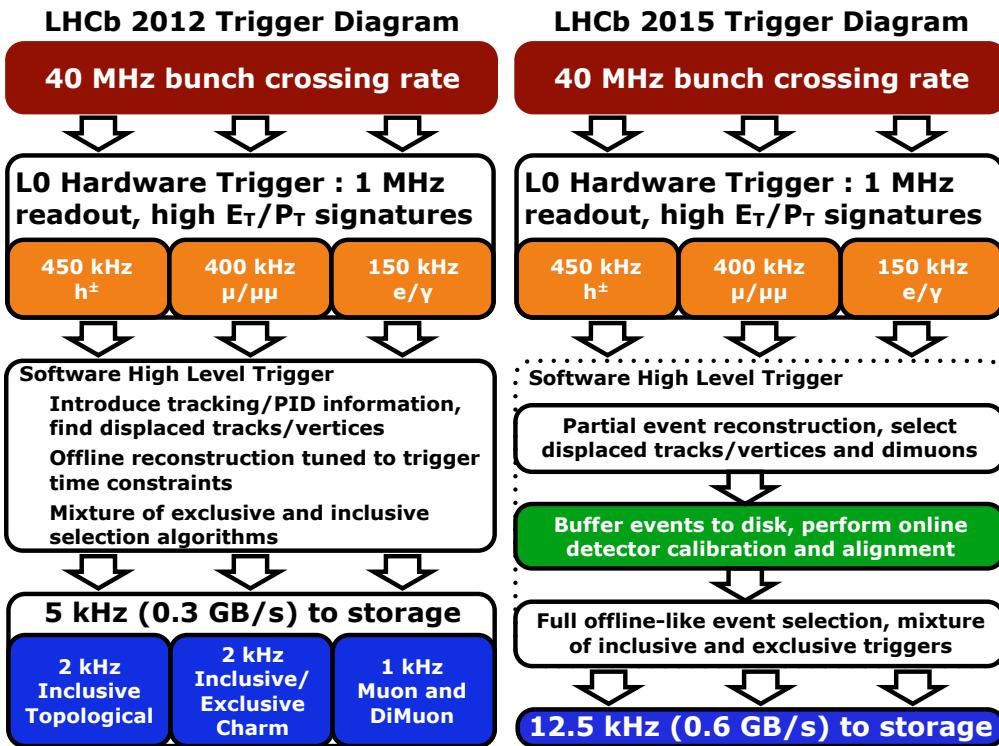


Figure 2.11 – Summary of the trigger strategy followed during Run 1 data-taking period (2011-2012, left) and Run 2 (2015-2017, right). During Run 2, an online detector and calibration alignment was introduced, plus full event selections (both inclusive and exclusive).

262 Level-0 (L0)

263 The L0 trigger mainly exploits the calorimeters and muon chambers. The idea behind the
 264 L0 trigger is to select events that contain high pT muons and high ET hadrons, electrons
 265 and photons, which very likely come from b - and c -hadron decays. The L0 trigger reduces
 266 the data rate from 40 MHz down to 1 MHz.

267 High Level Trigger (HLT)

268 The HLT is separated into two stages, HLT1 and HLT2, and runs on about 29000
 269 commercial CPU cores.

270 At the HLT1 level, the full detector information is read out, and then vertex/track
 271 reconstruction and PID are performed. The exploited signatures are mainly the presence
 272 of high p_T tracks, high transverse energy calorimeter clusters (photons and π^0), high
 273 di-muon invariant mass and tracks with large IP. All the HLT1 trigger lines are *inclusive*,
 274 meaning that only decay products common to various decay processes are selected rather

275 than specific decays. After the HLT1, the rate goes down to about 70 kHz.

276 The HLT2 is a combination of mainly inclusive selections and algorithms that reconstruct
 277 entirely (*exclusively*) some decay processes. The main lines are topological lines using
 278 Multi-Variate Analysis (MVA) with different sets of kinematic and position features as
 279 input, exclusive charm lines and high mass displaced di-hadron/lepton lines. After the
 280 HLT2, the events are finally stored on tape for further offline analysis.

281 **2.3.2 Event reconstruction, simulation and software**

282 **Track and vertex reconstruction**

283 Starting from the *hits* in the tracking detectors, tracks and vertices are reconstructed
 284 via dedicated algorithms. Different track types are reconstructed, as shown in Fig. 2.12.
 285 Each track is characterised by hits collected in different sub-detectors. For example,
 286 downstream tracks, with no hits in the VELO, are typically associated to long-lived
 287 particles such as Λ and K_s^0 . Because of the presence of a magnetic field, tracks are bend
 288 in the xz plane. By knowing the reconstructed particle trajectory and the magnetic field
 289 map, it is possible to measure the momentum of the particle.

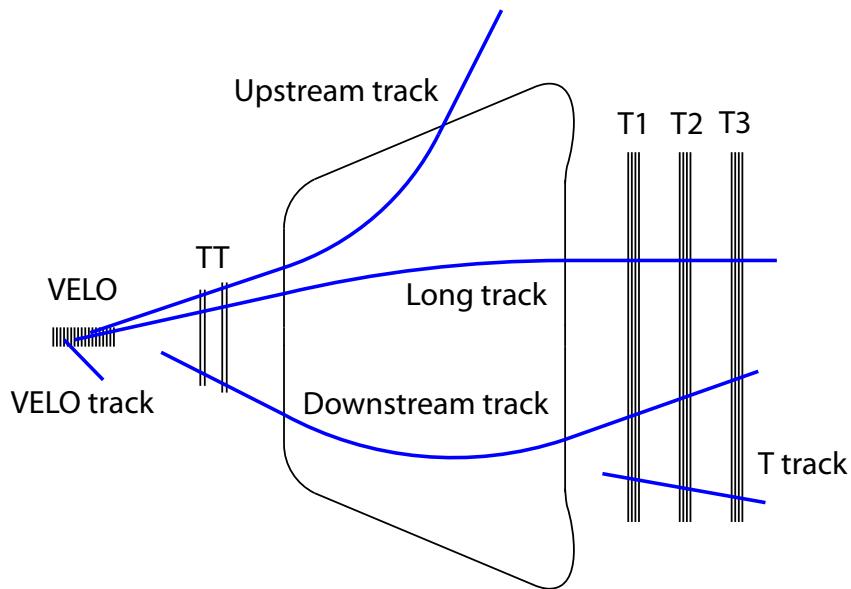


Figure 2.12 – Schematic description of the different track types reconstructed in LHCb.

290 **PID**

The Cherenkov radiation emitted by charged particles in the RICH radiators produces *rings* in the RICH acceptance, which are reconstructed via pattern recognition algorithms. For each reconstructed pattern, the *likelihood* \mathcal{L}_π for the ring to be produced by a pion (the most common particle in the LHCb environment) is computed. The momentum of the particle is also used in the likelihood computation. Then, likelihood functions for other hypotheses (kaon, proton, electron, muon) are computed and compared with the pion likelihood. For a given particle X ($X = e, \mu, p, k$), the PIDX observable is defined as:

$$\text{PIDX} = \ln \mathcal{L}_X - \ln \mathcal{L}_\pi \quad (2.1)$$

291 In typical LHCb analyses, requirements on 2.1 are applied to suppress physical backgrounds
292 due to wrongly-identified particles.

293 **Simulation**

294 The Monte Carlo (MC) simulation of pp collision, particle decays, and interaction with
295 the detector are crucial in the validation of physics analyses. The parton-parton collision
296 and hadronisation simulation is performed by PYTHIA [32], interfaced to EvtGen [33] for
297 the decay of the hadrons and leptons for standard productions. The QED corrections to
298 the decay (i.e. the emission of radiation photons) is generated by the PHOTOS package [34].
299 The interactions of particles with detector material and their tracking in the magnetic
300 field are simulated by GEANT4 [35].

301 2.3.3 Data collected by LHCb

The collision rate R [s^{-1}] in LHC can be expressed in terms of the *cross-section* σ [cm^2] and the *luminosity* \mathcal{L} [$\text{s}^{-1}\text{cm}^{-2}$] as:

$$R = \mathcal{L}\sigma. \quad (2.2)$$

302 For a given data taking interval, the *integrated luminosity* L [cm^{-2}] is a measure of the
303 amount of recorded data. The typical unit for luminosity is the inverse *barn*, which
304 corresponds to 10^{-24} cm^{-2} . The LHCb integrated luminosity are of the order of the
305 inverse *femtobarn* (fb^{-1}); one femtobarn corresponds to the production of about $10^{11} b\bar{b}$
306 quark pairs.

307 The LHCb detector collected data produced mainly from pp collisions in the 2010-2017
308 period, so far. During the 2011-2012 data taking period, called *Run 1*, about 3fb^{-1} of data
309 was collected. The center-of-mass energy \sqrt{s} of the pp interactions was 7 TeV and 8 TeV
310 in 2011 and 2012 respectively, and the time spacing between *bunches* of protons in LHC

2.3. Particle Identification (PID)

311 was 50 ns. The 2013-2014 period, known as *Long Shutdown 1* (LS1), was dedicated to
312 some maintenance work for the LHC components. The *Run 2* data taking period started
313 in 2015, and it is planned to last until the end of 2018. The center-of-mass energy of the
314 pp collisions during Run 2 is 13 TeV, whereas the time spacing between proton bunches
315 is 25 ns. In the 2015-2017 data taking period, about 3.7 fb^{-1} of data was collected. The
316 data collected for each year is summarised in Fig. 2.13.

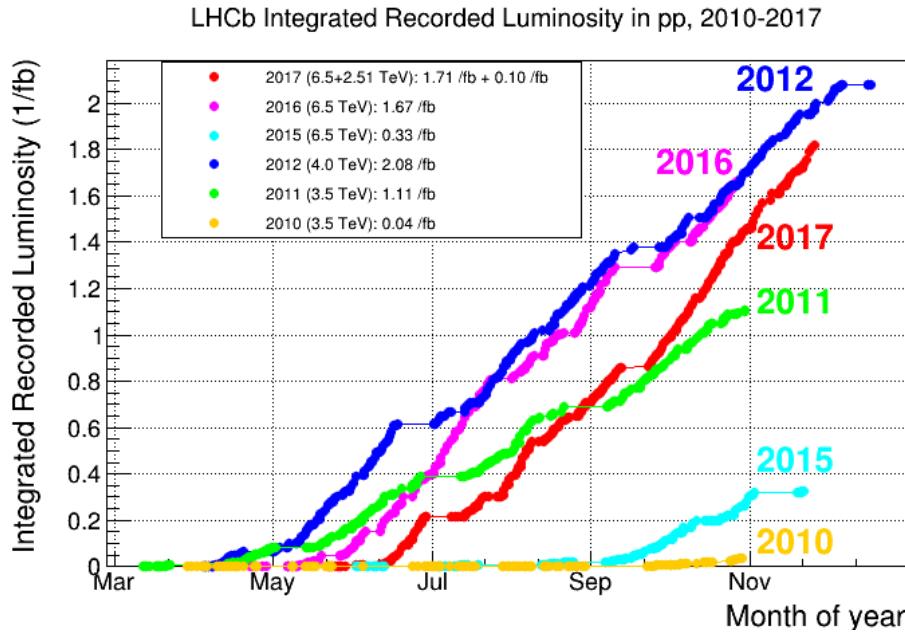


Figure 2.13 – Breakdown of the pp collision data collected by LHCb.

3 Flavour Tagging algorithms and ³¹⁸ optimisation of the opposite-side electron tagger

319 **3.1 Introduction to Flavour Tagging**

320 In this chapter, a description of the *flavour tagging* techniques at LHCb is reported. After
321 a brief introduction to the methods, the calibration of the *Opposite Side* (OS) and *Same*
322 *Side* (SS) algorithms for the $B^0 \rightarrow D^\pm \pi^\mp$ time-dependent analysis are described. Finally,
323 a reoptimisation of the OS *electron* (OSe) tagger on both Run 1 and Run 2 data is
324 reported. This work was made in collaboration with the University of Dortmund. During
325 my Phd work activity, I was mainly focused on the OS calibration and OSe reoptimistion.

326 **3.1.1 Flavour tagging algorithms**

327 The identification of the flavour at production time of a neutral B meson is a key element
328 for any time-dependent analysis which aims to measure oscillations and CP -asymmetries,
329 and in particular for the $B^0 \rightarrow D^\pm \pi^\mp$ analysis reported in this thesis. In fact, this
330 information is needed in order to measure experimentally the decay rates or asymmetries
331 introduced in Sec. 1.3. Complications arise from two facts:

- 332 • neutral B meson oscillate, so the flavour at the production time might differ from
333 the flavour at the decay time;
- 334 • many final states, such as $D^\pm \pi^\mp$, are not CP -eigenstates. This means that both B
335 and \bar{B} mesons can decay to the same final state particles, and so the flavour cannot
336 be obtained from the charge of these particles, even if there were no oscillations.

337 For these reasons, the flavour has to be reconstructed by exploiting particles not produced
338 by the neutral B meson being analysed.

339 Techniques to infer the initial flavour of a reconstructed candidate, *i.e.* whether it contained
340 a b or a \bar{b} quark at production, are called flavour tagging algorithms. Several flavour
341 tagging algorithms exist in LHCb; they can be categorised into same side taggers (SS
342 taggers) and opposite side taggers (OS taggers). A schematic representation of the taggers
343 that can be used for tagging B^0 mesons is shown in fig. 3.1.

344 The SS taggers infer the production flavour of the signal B meson by selecting charged
345 particle candidates that have a high chance of being remnants of the hadronisation process
346 of the B candidate [36, 37]. For B^0 mesons, the same side pion tagger, which exploits
347 π^\mp mesons produced in the hadronisation of the B^0 meson, and the same side proton
348 tagger, which looks for co-produced protons, have been developed. For both taggers, the
349 charge of the pion or proton particle is correlated with the production flavour of the signal
350 B^0 meson. The response of the two taggers is combined into a common SS response.

351 In contrast, the OS taggers exploit the predominant production process of B mesons
352 via $b\bar{b}$ quark pair production [38]. They partially reconstruct the decay of the (*other*)

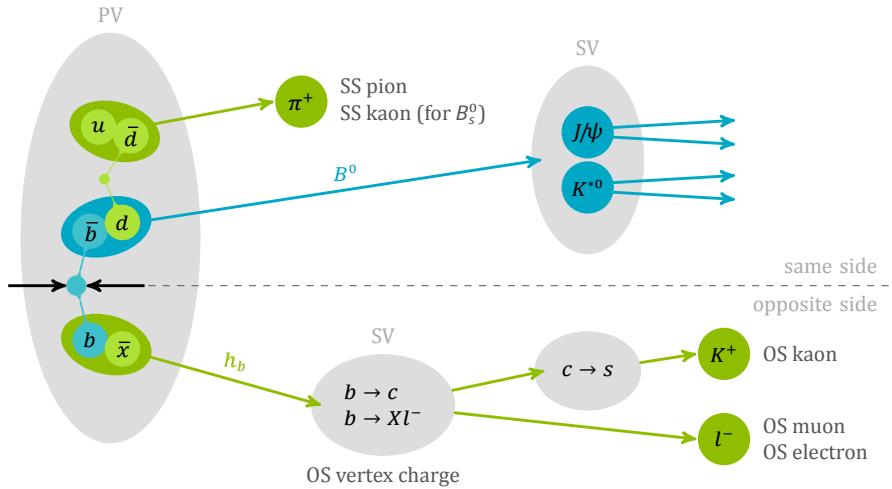


Figure 3.1 – Flavour tagging algorithms used in LHCb. In this cartoon, the signal channel is considered to be $B^0 \rightarrow J/\psi K^{*0}$.

353 b -hadron produced along with each reconstructed signal B meson, and infer its initial
 354 flavour. In fact, the flavour of the signal B meson and the other b -hadron produced in the
 355 same event are anti-correlated. Several OS taggers have been developed in LHCb, where
 356 the combination of the OS kaon, muon, electron and vertex charge tagging algorithms
 357 represents the current standard OS combination. An additional OS tagger, the OS Charm
 358 tagger [39], can be exploited, and can be combined with the OS standard combination.

359 Given a reconstructed candidate, each flavour tagging algorithm provides a flavour tag d
 360 and a prediction η for the probability of the tag to be wrong. This mistag probability η is
 361 defined on the range $[0, 0.5]$ and is based on the output of multivariate classifiers, which
 362 are trained on datasets of flavour specific decays, and combine several kinematic and
 363 geometric information on the tagging particle(s) and the event. The flavour tag takes the
 364 values $d = +1$ for an initial B^0 , $d = -1$ for an initial \bar{B}^0 , and $d = 0$ when no tag could
 365 be assigned; this happens, for example, if the tagging particle do not pass the selection
 366 criteria of a given tagger algorithm, or if its trajectory lies outside the LHCb acceptance.

367 More details on the flavour tagging at LHCb can be found in Refs. [40, 41, 42].

368 Performance characteristics

The performance of flavour tagging algorithms can be characterised by different quantities. If N_U is the number of untagged candidates and N_W (N_R) is the number of wrongly (correctly) tagged candidates, the *tagging efficiency* (*i.e.* the fraction of tagged candidates)

Chapter 3. Flavour Tagging algorithms and optimisation of the opposite-side electron tagger

can be defined as:

$$\varepsilon_{\text{tag}} = \frac{N_R + N_W}{N_R + N_W + N_U}. \quad (3.1)$$

The fraction of wrongly tagged candidates, or *mistag fraction*, is given by:

$$\omega = \frac{N_W}{N_R + N_W}. \quad (3.2)$$

A non-zero mistag fraction dilutes the time-dependent asymmetries, reducing the experimental sensitivity to them. For instance, the measured decay rates for a $B \rightarrow f$ decay and its CP -conjugated are:

$$\frac{d\Gamma^{\text{meas}}}{dt} = (1 - \omega) \frac{d\Gamma}{dt} + \omega \frac{d\bar{\Gamma}}{dt}, \quad (3.3)$$

$$\frac{d\bar{\Gamma}^{\text{meas}}}{dt} = \omega \frac{d\Gamma}{dt} + (1 - \omega) \frac{d\bar{\Gamma}}{dt}. \quad (3.4)$$

As a consequence, the measured CP -asymmetry is:

$$A^{\text{meas}}(t) = \frac{\frac{d\bar{\Gamma}^{\text{meas}}}{dt} - \frac{d\Gamma^{\text{meas}}}{dt}}{\frac{d\bar{\Gamma}^{\text{meas}}}{dt} + \frac{d\Gamma^{\text{meas}}}{dt}} = (1 - 2\omega) A^{\text{phys}}(t) = D A^{\text{phys}}(t). \quad (3.5)$$

- 369 The quantity $D = (1 - 2\omega)$ is known as average *dilution*. If $\omega = 0$ (perfect tagger), then
370 $D = 1$ and no asymmetry dilution occur. If $\omega = 0.5$ (random tagger), then $D = 0$, and it
371 is not possible to measure the asymmetry anymore.

The quantity that can be interpreted as the figure of merit to optimise a tagging algorithm is the *effective tagging efficiency*, also called *tagging power*:

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} (1 - 2\omega)^2 = \varepsilon_{\text{tag}} D^2. \quad (3.6)$$

In fact, it can be shown that the statistical uncertainty on the physical asymmetry is given by:

$$\sigma_{A^{\text{phys}}} = \frac{\sqrt{1 - A^{\text{meas}}^2}}{\sqrt{N\varepsilon_{\text{tag}}}(1 - 2\omega)} = \frac{\sqrt{1 - A^{\text{meas}}^2}}{\sqrt{N\varepsilon_{\text{eff}}}}, \quad (3.7)$$

where N is the total number of candidates. So, according to Eq. 3.7, the greater the tagging power, the smaller the resulting statistical uncertainty on the CP -asymmetry. Instead of using an average mistag fraction or *probability* ω , it is possible to exploit the mistag probability η estimated by the tagging algorithm. This probability η is evaluated for each B candidate individually, rather than being a global quantity. Usually, η needs to be *calibrated* via a function $\omega(\eta)$ in order to return the true mistag probability (details

3.1. Introduction to Flavour Tagging

in Sec. 3.1.1). So, the tagging power can be rewritten as:

$$\varepsilon_{\text{eff}} = \frac{1}{N} \sum_{i=1}^N D_i^2 = \frac{1}{N} \sum_{i=1}^N (1 - 2\omega(\eta_i))^2, \quad (3.8)$$

where $\omega(\eta_i) = 0.5$ ($D_i = 0$) for untagged candidates.

An example of dilution effect can be seen in Fig. 3.2, which shows how the measured amplitude of an asymmetry gets smaller for increasing values of η .

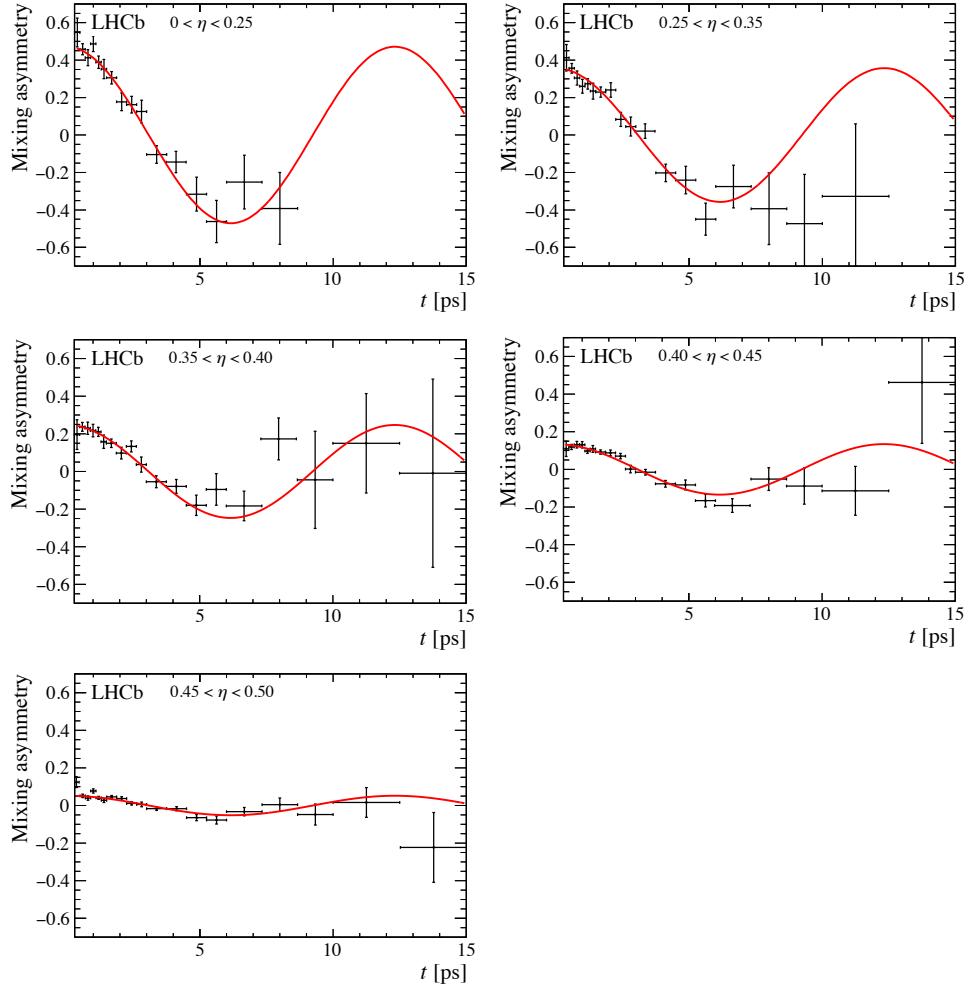


Figure 3.2 – Mixing asymmetry for SS pion tagged $B_s^0 \rightarrow D_s\pi$ candidates in bins of increasing estimated mistag η [37].

Chapter 3. Flavour Tagging algorithms and optimisation of the opposite-side electron tagger

375 Calibration of the tagging output

376 The output of the flavour tagging algorithms is the result of training multivariate classifiers
377 (MVA) using datasets of flavour specific B decays, and transforming the classifier output
378 into mistag estimates η through regression. However, as the training and validation
379 samples can be different from the signal sample used in the CP measurement (*e.g.* in
380 terms of trigger and selection criteria that affect the distribution of the MVA input
381 features), the output needs to be calibrated. Again, using control samples of flavour
382 specific decays, calibration functions $\omega(\eta)$ are obtained to transform the mistag estimate
383 η of the algorithm to the mistag probability ω measured in the control sample.

In literature, a common choice for the calibration function is a linear function,

$$\omega(\eta) = p_0 + p_1 (\eta - \langle \eta \rangle). \quad (3.9)$$

384 The use of the arithmetic mean $\langle \eta \rangle$ of the η distribution aims for a decorrelation of p_0
385 and p_1 , hence a perfect calibration of the taggers would result in $p_0 = \langle \eta \rangle$ and $p_1 = 1$.

The performance of the flavour taggers is not necessarily independent of the initial flavour of the B^0 . The charged decay products, like the K^\pm which are used by the OS kaon tagger, can have significantly different interaction rates with the detector material and therefore different reconstruction efficiencies. This can result in different tagging efficiencies ε_{tag} and mistag probabilities ω for initial B^0 and \bar{B}^0 . These tagging asymmetries can dilute or enhance the observed raw tag asymmetry and need to be corrected for. The asymmetries of the mistag probability, *i.e.* the difference of the tagging calibration parameters p_0 and p_1 for initial B^0 and \bar{B}^0 , can be parameterised by two independent calibration functions

$$\begin{aligned} \omega(\eta)^{B^0} &= p_0^{B^0} + p_1^{B^0} (\eta - \langle \eta \rangle), \\ \omega(\eta)^{\bar{B}^0} &= p_0^{\bar{B}^0} + p_1^{\bar{B}^0} (\eta - \langle \eta \rangle). \end{aligned} \quad (3.10)$$

Equivalently, we can parameterise the calibration parameters p_i (with $i = 0, 1$) as

$$p_i^{B^0} = p_i + \frac{\Delta p_i}{2}, \quad p_i^{\bar{B}^0} = p_i - \frac{\Delta p_i}{2} \quad (3.11)$$

The difference between the mistag of B^0 and \bar{B}^0 can be written as

$$\begin{aligned} \Delta\omega(\eta) &= \omega^{B^0}(\eta) - \omega^{\bar{B}^0}(\eta) \\ &= \Delta p_0 + \Delta p_1 (\eta - \langle \eta \rangle) \end{aligned} \quad (3.12)$$

386 In this thesis, new models for the calibration functions are adopted instead of the standard
387 linear calibrations. These different parameterisations are called *Generalised Linear Models*
388 (GLM), and are implemented in the EPM (*Espresso Performance Monitor*) package [43].

In general, a GLM which relates the predicted mistag probability η to the calibrated probability ω can be written as follows:

$$\omega(\eta) = g(h(\eta)) = g \left(g^{-1}(\eta) + \sum_{i=1}^N \left(p_i + \frac{d\Delta p_i}{2} \right) f_i(\eta) \right). \quad (3.13)$$

- 389 The functions $f_i(\eta)$ are called *basis functions*, and they can be chosen as polynomials
- 390 or spline functions. The set on basis functions is automatically orthogonalised by the
- 391 EPM by using the Gram-Schmidt method; this ensures that the corresponding calibration
- 392 parameters p_i and Δp_i are less correlated as possible.
- 393 The parameter d is the tagging decision, which is incorporated into the model in order to
- 394 parameterise $\omega(\eta)$ for the two possible flavours.
- 395 The function g is known as *link function*. Usually, this is chosen as the inverse of a
- 396 cumulative distribution function in order to map input values into the interval $[0, 1]$, such
- 397 that the output can be naturally interpreted as a probability.

For the $B^0 \rightarrow D^\pm \pi^\mp$ analysis presented in this thesis, the adopted link function g is a *modified* logistic function, defined as:

$$g(h) = \frac{1}{2(1 + e^h)}, \quad (3.14)$$

- 398 where h is defined in Eq. 3.13. This link function is built such that the calibrated mistag
- 399 probability is defined in the $[0, 0.5]$ interval. This choice allows to solve a numerical issue
- 400 that often occurs when standard link functions (e.g. identity or logistic) are adopted. In
- 401 fact, if $\omega > 0.5$, then an arbitrary prescription has to be taken (e.g., label the candidate
- 402 as untagged, or flip the tagging decision and take $1 - \omega$ as new calibrated mistag). If the
- 403 calibration parameters are free parameters in a time-dependent fit, this choice has to be
- 404 made during the minimisation process, according to the values ω takes at each iteration.
- 405 This means that the relative amount of B and \bar{B} , or the relative amount of tagged and
- 406 untagged candidates, may change during the fit, which yields to numerical instabilities
- 407 due to discontinuous changes in the likelihood function.
- 408 The EPM estimates the calibration parameters via an unbinned maximum likelihood
- 409 method called *binomial regression*; this is an improvement over traditional, binned least-
- 410 squares methods, which are affected by a systematic uncertainty due to the binning
- 411 choice.

412 **Combination of multiple taggers**

When more than one tagger is available per event, the tagging decisions and mistag probabilities provided by each tagger can be combined into a single decision and a single probability using the following equations:

$$p(b) = \prod_i \left(\frac{1 + d_i}{2} - d_i (1 - \eta_i) \right), \quad p(\bar{b}) = \prod_i \left(\frac{1 - d_i}{2} + d_i (1 - \eta_i) \right) \quad (3.15)$$

where $p(\bar{b}/b)$ is the probability that the signal B^0 contains a \bar{b}/b , d_i is the decision taken by the i -th tagger and η_i is the predicted mistag probability of the i -th tagger. These probabilities are normalised as

$$P(\bar{b}) = \frac{p(\bar{b})}{p(\bar{b}) + p(b)}, \quad P(b) = 1 - P(\bar{b}). \quad (3.16)$$

413 If $P(\bar{b}) > P(b)$ the combined tagging decision is $d = +1$ and the final mistag probability
414 is $\eta = 1 - P(\bar{b})$. Otherwise if $P(b) > P(\bar{b})$ the combined tagging decision and the mistag
415 probability are $d = -1$ and $\eta = 1 - P(b)$.

416 Eq. 3.15 is valid under the assumption that all taggers in the combination are independent.
417 In the $B^0 \rightarrow D^\pm \pi^\mp$ analysis presented in this thesis, the OS taggers are combined in a
418 single OS combination, and the same is done for the SS taggers. Effects due to correlations
419 among taggers within a combination are corrected by calibrating the combined predicted
420 mistag.

421 **3.2 Flavour tagging strategy for the $B^0 \rightarrow D^\pm \pi^\mp$ time-dependent
422 analysis**

423 In the $B^0 \rightarrow D^\pm \pi^\mp$ analysis presented in this thesis, the OS combination (including the OS
424 charm tagger) and the SS combination are used. The implementation of the OS algorithms
425 used in the combination are the same described in Refs. [38] and [39]; the OS algorithms
426 other than OS charm were built as neural networks trained on $B^+ \rightarrow J/\psi K^+$ Run 1
427 data, whereas the OS charm tagger was implemented with a BDT trained on a cocktail of
428 simulated $B^+ \rightarrow J/\psi K^+$, $B^0 \rightarrow J/\psi K^*$ and $B_s^0 \rightarrow J/\psi \phi$ decays. The SS taggers have
429 been reimplemented for this specific analysis by exploiting $B^0 \rightarrow J/\psi K^{*0}$ decays. The
430 functional form of the tagging calibrations is studied in control samples of flavour-specific
431 decays properly corrected to resemble the signal decay. The calibration parameters are
432 determined directly in the decay-time-dependent fit of the signal described in Sec. 4.7; they
433 are nuisance parameters of the likelihood function. Determining the calibration parameters
434 from the data along with the CP observables is possible because the CP coefficients C_f
435 and $C_{\bar{f}}$ of Eqs. 1.36–1.39 are fixed in this analysis (to 1 and -1 respectively). Hence, the

3.2. Flavour tagging strategy for the $B^0 \rightarrow D^\pm \pi^\mp$ time-dependent analysis

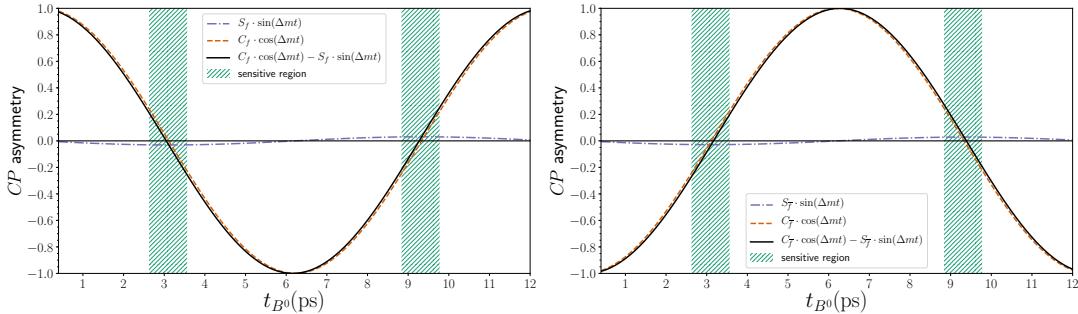


Figure 3.3 – \bar{B}^0 versus B^0 time-dependent asymmetries for the $D^- \pi^+$ (left) and $D^+ \pi^-$ (right) final states. The values of C_f , $C_{\bar{f}}$, S_f and $S_{\bar{f}}$ are the ones used in simulation. The sensitivity to S_f and $S_{\bar{f}}$ is maximised in the intervals called “sensitive regions”, since the $\sin(\Delta m)$ amplitude becomes of the same order of the $\cos(\Delta m)$ amplitude, which is close to zero. In the outer regions, since C_f ($C_{\bar{f}}$) is fixed to 1 (-1) in the fit, the mistag dilution (which depends on the flavour tagging calibration parameters) adapts to fit the $\cos(\Delta m)$ amplitude, giving sensitivity to the calibration parameters.

436 cosine terms give sensitivity to the calibration parameters independently from the S_f
 437 and $S_{\bar{f}}$ coefficients, which are proportional to the sine terms. An heuristic explanation is
 438 represented in Fig. 3.3. This strategy allows to avoid any assumption on the portability to
 439 the signal sample of the calibration parameters determined from the control data. Such a
 440 strategy was studied extensively on simulation: the increase of the statistical uncertainty
 441 of the S_f and $S_{\bar{f}}$ coefficients given by the additional degrees of freedom of the calibration
 442 parameters is smaller than the systematic uncertainties associated to the calibrations
 443 portability. Moreover, the use of the calibration parameters from the control decays
 444 causes biases on S_f and $S_{\bar{f}}$ of the order of their statistical uncertainty; when letting the
 445 calibration parameters floating in the fit, such biases are suppressed or disappear, at the
 446 cost of a moderate increase of the statistical uncertainty. In addition, while the precision
 447 of the OS tagger calibration from the control decay is similar to the one from the signal
 448 channel, the calibration of the SS tagger derived from the signal data (Tab. 4.20) is much
 449 more precise than that from the control decay (Tab. 3.4).

450 In what follows, the study of the tagging calibration from the control decays is presented.
 451 For all reasons discussed above, these studies are not meant for determining the calibration
 452 parameters to use in the time-fit to the signal data (usual strategy adopted in all flavour-
 453 tagged time-dependent analyses), but they serve the purpose of: i) determining the best
 454 functional form of the calibration functions to be used in the fit to the signal; ii) having
 455 some reference values for the calibration parameters for a comparison with those extracted
 456 from the signal.

457 The calibration for the OS combination are determined using $B^+ \rightarrow D^0 \pi^+$ decays, as
 458 described in section 3.2.1. The SS pion and the SS proton taggers were developed using

Chapter 3. Flavour Tagging algorithms and optimisation of the opposite-side electron tagger

459 $B^0 \rightarrow D^\mp \pi^\pm$ data and assuming negligible CP violation. The use of these algorithms in
460 this analysis could bias the measurement. Therefore, the SS taggers are retrained using
461 $B^0 \rightarrow J/\psi K^{*0}$ decays. The calibration of the SS combination is described in Sec. 3.2.2.

462 3.2.1 Calibration of the Opposite Side Tagger Combination

463 Data sample selection

464 The calibration parameters of the OS taggers combination (namely combination of the
465 OS electron, muon, kaon, vertex charge, and charm algorithms) are determined using
466 $B^+ \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ candidates reconstructed in 3 fb^{-1} of data. Such a control decay
467 mode provides very high statistics (more than 300K OS tagged signal candidates) and is
468 very similar to the signal decay $B^0 \rightarrow D^-\pi^+$.

469 Candidates $B^+ \rightarrow D^0\pi^+$ are selected through the `B2DOPiD2HHBeauty2CharmLine` strip-
470 ping line, versions `S21r1` (2011 data) and `S21` (2012 data), of the `BhadronCompleteEvent`
471 stream. The B^+ candidates are required to be TOS on `Hlt1TrackAllL0Decision` at the
472 HLT1 stage, and on one between `Hlt2Topo2BodyBBDTDecision`, `Hlt2Topo3BodyBBDTDecision`,
473 or `Hlt2Topo4BodyBBDTDecision`, at HLT2. The additional requirements listed in Tab. 3.1
474 are applied to further suppress backgrounds and enhance the signal purity.

475 A fit to the mass distribution of B^+ candidates is done to calculate `sWeights`, used in
476 the subsequent steps of the analysis to subtract the backgrounds surviving the selection.
477 This fit is described in details in App. 4.9.1.

478 Event-by-event weights are calculated to equalise the $B^+ \rightarrow D^0\pi^+$ and $B^0 \rightarrow D^\pm\pi^\mp$
479 distributions of variables the tagging calibration can depend on. The procedure and the
480 results of this reweighting are reported in Appendix 4.9.2. Additionally, the number of
481 B^+ and B^- candidates are made equal in the sample to avoid any spoil of the calibration
482 parameters due to a B^+/B^- production asymmetry or a detection asymmetry. All these
483 weights, along with `sWeights`, are applied during the calibration procedure.

484 Calibration

485 The calibration of the estimated mistag η is performed on the fully reweighted $B^+ \rightarrow D^0\pi^+$
486 dataset. A GLM model with NSpline basis function [43] is adopted.

487 The projection of the fitted calibration function over the $B^+ \rightarrow D^0\pi^+$ dataset is shown
488 in Fig. 3.4, whereas the fitted calibration parameters are listed in Tab. 3.2

489 The number of free parameters in the adopted GLM model (10) has been chosen in order
490 to have satisfactory goodness-of-fit (GOF) metrics (more details in Appendix 4.9.3).

3.2. Flavour tagging strategy for the $B^0 \rightarrow D^\pm \pi^\mp$ time-dependent analysis

Table 3.1 – Selection requirements for the $B^+ \rightarrow D^0 \pi^+$ data sample.

Variable	Description	Requirement
bachelor track		
IsMuon	muon identification criteria	= 0
g_{prob}	ghost probability	< 0.1
$\chi^2_{\text{track}}/\text{ndof}$	quality of track	< 2
D^0 daughter tracks		
PID k	$\log L_K - \log L_\pi$	> -2 (kaon), < 8 (pion)
g_{prob}	ghost probability	< 0.1
$\chi^2_{\text{track}}/\text{ndof}$	quality of track	< 2.5
B^+ candidate		
τ_{B^+}	decay time	$0.2 \text{ ps} \leq \tau_{B^+} \leq 15 \text{ ps}$
MIN $\chi^2_{\text{IP PV}}$	minimum IP χ^2 w.r.t PV	< 15
D^0 selection cuts		
$m_{K\pi}$	D^0 invariant mass	$1830 \text{ MeV}/c^2 < m_{K\pi} < 1904 \text{ MeV}/c^2$

Table 3.2 – Fitted OS calibration parameters on the $B^+ \rightarrow D^0 \pi^+$ reweighted dataset.

Parameter	Fitted value
p_0	-0.13596 ± 0.018861
p_1	-0.0056847 ± 0.022384
p_2	-0.010699 ± 0.0083159
p_3	-0.45139 ± 0.10295
p_4	-0.84808 ± 0.46093
Δp_0	-0.12897 ± 0.037722
Δp_1	0.041859 ± 0.044769
Δp_2	-0.020072 ± 0.016632
Δp_3	0.42483 ± 0.20589
Δp_4	1.91 ± 0.92186

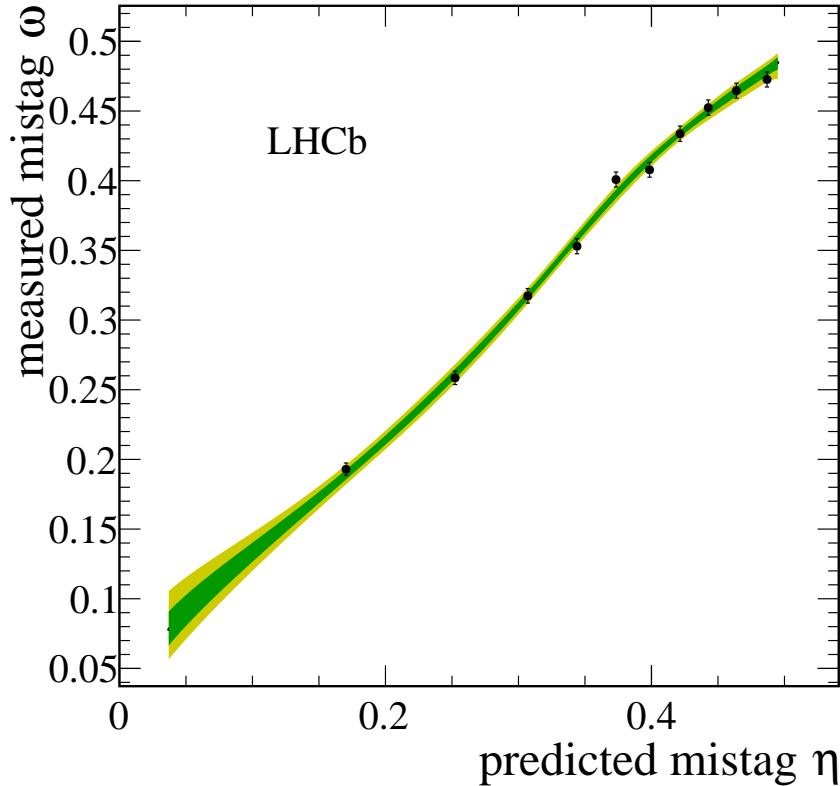


Figure 3.4 – Projection of the fitted OS calibration function over the $B^+ \rightarrow D^0\pi^+$ reweighted dataset.

491 Calibration portability

492 The aim of the calibration is to return a mistag ω as close as possible to the *true* mistag,
493 which would be given by a *true calibration*. The latter is not defined on data, but it is
494 possible to estimate it on MC by calibrating the predicted mistag η on the truth-matched
495 signal decay $B^0 \rightarrow D^-\pi^+$. In fact, since, unlike data, the true flavour of B^0 is known on
496 MC, this MC calibration can be done in the same way as $B^+ \rightarrow D^0\pi^+$, where the true
497 flavour is given by the B charge.

498 This $B^0 \rightarrow D^-\pi^+$ calibration is performed after equalising the number of B^0 and \bar{B}^0 in
499 the sample, in order to disentangle tagging asymmetries from CP violation and production
500 asymmetries.

501 The $B^+ \rightarrow D^0\pi^+$ MC calibration is performed in exactly the same way as described
502 in Sec. 3.2.1, except that no *sWeights* are considered, since MC samples contain signal
503 candidates only.

3.2. Flavour tagging strategy for the $B^0 \rightarrow D^\pm \pi^\mp$ time-dependent analysis

504 The projections of the two calibrations over $B^0 \rightarrow D^- \pi^+$ and $B^+ \rightarrow D^0 \pi^+$ MC samples
 505 are shown in Fig. 3.5. A side-by-side comparison of the two calibrations is reported
 506 in Tab. 3.3. A more robust comparison is obtained from a χ^2 function describing the
 507 discrepancy between the two calibrations by taking covariance matrices into account. The
 508 overall discrepancy (corresponding to the χ^2 minimum) is around 2.03σ .

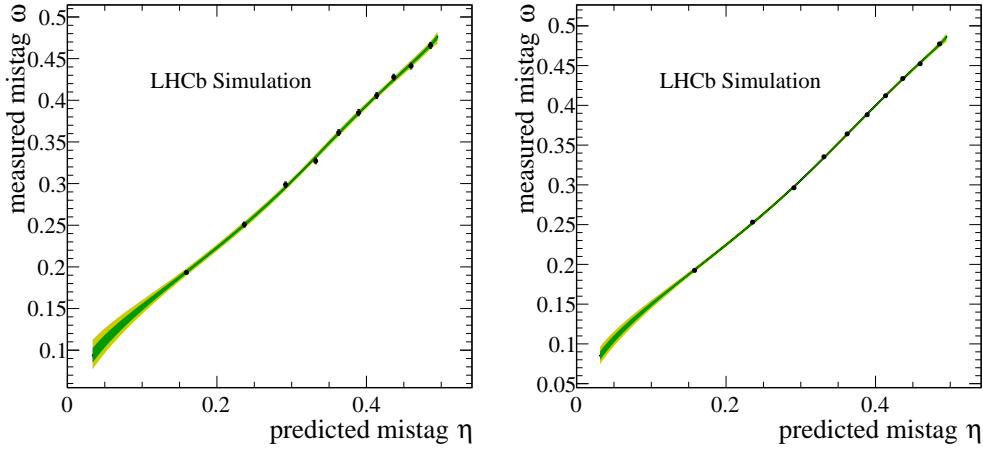


Figure 3.5 – Projection of the fitted OS calibration function over the $B^+ \rightarrow D^0 \pi^+$ (left) and $B^0 \rightarrow D^- \pi^+$ (right) truth-matched Monte Carlo samples.

Table 3.3 – Comparison between fitted calibration parameters on truth-matched $B^+ \rightarrow D^0 \pi^+$ and $B^0 \rightarrow D^- \pi^+$ MC samples. The discrepancy is computed assuming uncorrelated parameters and independent datasets.

Parameter	$B^+ \rightarrow D^0 \pi^+$	$B^0 \rightarrow D^- \pi^+$	Discrepancy (σ)
p_0	-0.065 ± 0.011	-0.0996 ± 0.0066	2.70
p_1	-0.190 ± 0.012	-0.1492 ± 0.0077	-2.84
p_2	-0.0105 ± 0.0044	-0.0191 ± 0.0029	1.63
p_3	-0.295 ± 0.054	-0.234 ± 0.036	-0.93
p_4	-0.42 ± 0.26	-0.14 ± 0.20	-0.85
Δp_0	-0.059 ± 0.022	-0.058 ± 0.013	-0.03
Δp_1	0.044 ± 0.024	0.030 ± 0.015	0.46
Δp_2	-0.0012 ± 0.0088	-0.0126 ± 0.0058	1.08
Δp_3	-0.08 ± 0.11	-0.046 ± 0.073	-0.25
Δp_4	-0.34 ± 0.53	-0.29 ± 0.39	-0.08

509 3.2.2 Calibration of the Same Side Tagger Combination

510 The same side algorithms determine the flavour content of the initial B meson by exploiting
 511 its production mechanism, *e.g.* in the B^0 meson production from the strong decay of
 512 an excited meson $B^{**+} \rightarrow B^{(*)0} \pi^+$, the B^0 flavour can be determined by the charge of

Chapter 3. Flavour Tagging algorithms and optimisation of the opposite-side electron tagger

Table 3.4 – SS calibration parameters obtained on the $B^0 \rightarrow J/\psi K^{*0}$ data sample (calibration sub-sample).

p_0	p_1	Δp_0	Δp_1
-0.091 ± 0.059	-0.027 ± 0.065	0.034 ± 0.084	0.032 ± 0.094

513 the companion pion. Additionally, a charge correlation can arise from the fragmentation
514 process of a b quark.

515 As described in Ref. [37], the SS pion and SS proton tagger were both trained on the
516 2012 data sample of the decay $B^0 \rightarrow D^\mp \pi^\pm$. As the effect of CP violation was neglected
517 during the training the algorithms and the underlying MVAs cannot be blindly used
518 when measuring CP violation in the same decay channel. Thus, $B^0 \rightarrow J/\psi K^{*0}$ decays
519 are chosen instead, as they represent a flavour-specific B^0 decay with a considerably large
520 signal yield.

521 Once the SS pion and proton taggers are implemented, they are combined into a single
522 SS combination as described in Sec. 3.1.1.

523 Calibration

524 The calibration is performed on a data $B^0 \rightarrow J/\psi K^{*0}$ subsample which is not used for
525 the SS pion and proton training. A GLM model having a first order polynomial is chosen
526 as basis function and a modified logistic function (Eq. 3.14) is used as link. The number
527 of free parameters in this model (4) is tuned in order to have satisfactory goodness-
528 of-fit (GOF) metrics. Together with the *sWeights*, additional weights to correct the
529 $B^0 \rightarrow J/\psi K^*$ data to resemble those of $B^0 \rightarrow D^\pm \pi^\mp$ are applied during the calibration.
530 The resulting calibration parameters are listed in Table 3.4 and a graphical representation
531 of the calibration and the distribution of the combined mistag is presented in Fig. 3.6.

532 Calibration portability

533 In the same way as done for the OS taggers (Sec. 3.2.1), the portability is checked on
534 Monte Carlo for the SS tagging algorithms. For $B^0 \rightarrow D^- \pi^+$ the calibration is performed
535 using the true flavour of the B^0 meson after equalising the number of B^0 and \bar{B}^0 in the
536 sample, in order to disentangle tagging asymmetries from CP violation and production
537 asymmetries. Also on $B^0 \rightarrow J/\psi K^{*0}$ the true flavour of the B^0 meson is used for the
538 calibration, and no *sWeights* are needed, since the simulated data only contain signal
539 candidates.

540 The projections of the two calibrations over $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow J/\psi K^{*0}$ Monte
541 Carlo samples are shown in Fig. 3.7. A side-by-side comparison of the two calibrations is

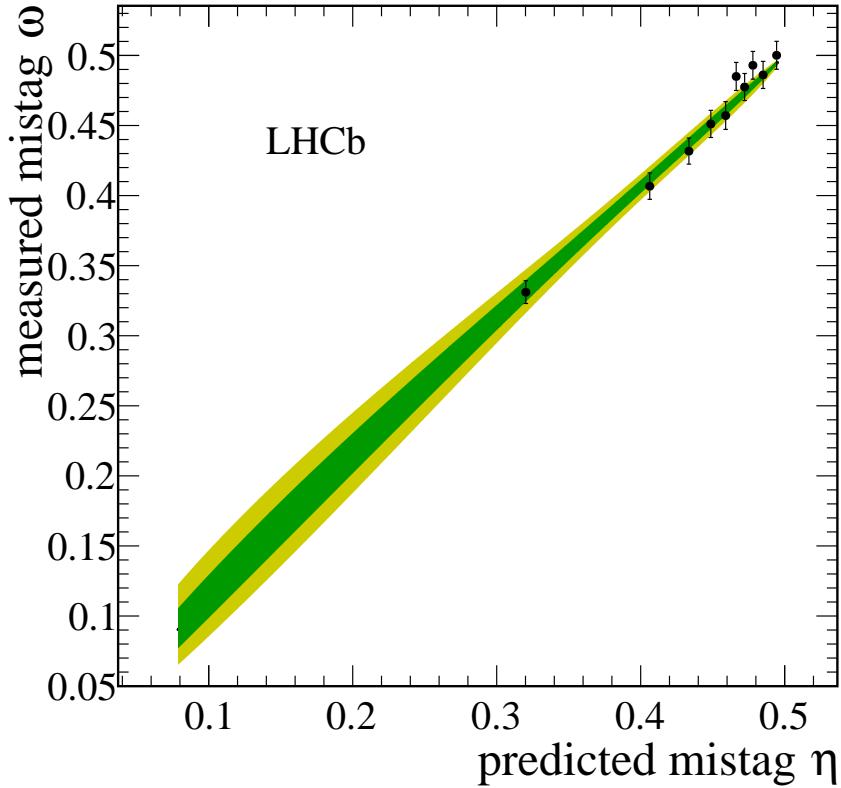


Figure 3.6 – Projection of the fitted SS calibration function over the $B^0 \rightarrow J/\psi K^*$ reweighted dataset.

542 reported in Tab. 3.5. A full comparison that takes into account the correlation between
 543 the parameters is obtained from a χ^2 test similar to the one described in Sec. 3.2.1. The
 544 agreement is around 0.09σ . Despite this test doesn't hint to issues of portability between
 545 the decay modes, the same strategy used for the OS calibrations is followed, i.e fitting the
 546 parameter directly in data with the CP asymmetries. This is motivated by the fact that
 547 the $B^0 \rightarrow D^- \pi^+$ signal sample has much more sensitivity to determine the parameters
 548 than the $B^0 \rightarrow J/\psi K^{*0}$ sample. In addition, with this approach no systematic related to
 549 calibration portability is necessary, consistently with the case of the OS tagger.

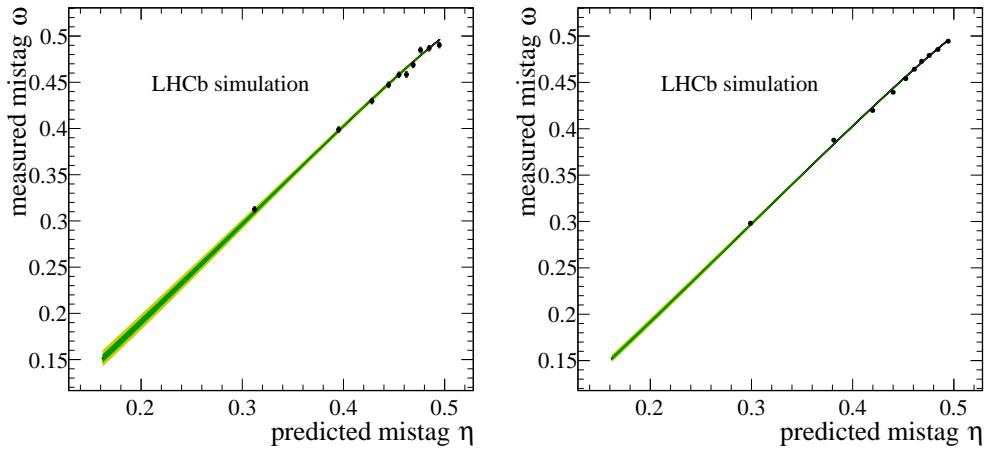


Figure 3.7 – Projection of the fitted SS calibration function over the $B^0 \rightarrow J/\psi K^{*0}$ (left) and $B^0 \rightarrow D^- \pi^+$ (right) truth-matched Monte Carlos samples.

Table 3.5 – Comparison between fitted SS tagger calibration parameters on truth-matched $B^0 \rightarrow J/\psi K^{*0}$ and $B^0 \rightarrow D^- \pi^+$ Monte Carlo samples. The discrepancy is computed assuming uncorrelated parameters and independent datasets.

Parameter	$B^0 \rightarrow J/\psi K^{*0}$	$B^0 \rightarrow D^- \pi^+$	Discrepancy (σ)
p_0	-0.016 ± 0.017	-0.019 ± 0.008	-0.19
p_1	0.063 ± 0.021	0.060 ± 0.010	-0.14
Δp_0	-0.029 ± 0.033	-0.027 ± 0.015	0.04
Δp_1	-0.026 ± 0.041	0.015 ± 0.019	0.90

4 Measurement of CP violation in ₅₅₁ $B^0 \rightarrow D^\mp \pi^\pm$ decays

552 All the work presented in this chapter has been carried out in cooperation with the
 553 University of Dortmund. The main points for which I was responsible during my Phd
 554 time are:

- 555 • OS flavour tagging (already presented in the previous chapter, Sec. 3.2.1);
 556 • PID correction (Sec. 4.3.1);
 557 • mass fit for *sWeights* calculation (Sec. 4.4);
 558 • time-dependent analysis, in particular the acceptance parameterisation (Sec. 4.6),
 559 the decay-time fit (Sec. 4.7) and the estimation of systematic uncertainties (Sec. 4.8);

560 **4.1 Introduction**

561 In this chapter, a decay-time dependent analysis of the decay $B^0 \rightarrow D^\pm\pi^\mp$ is presented,
 562 where the D meson is reconstructed as $D^\pm \rightarrow K^\mp\pi^\pm\pi^\pm$. The pion produced together with
 563 the D meson will be named as *bachelor* or *companion* particle hereafter. The objective
 564 of this study is to perform a measurement of CP asymmetries, which allow to constrain
 565 the CKM angle γ [15, 16]. Measurements of γ via tree-level decays such as $B^0 \rightarrow D^\pm\pi^\mp$
 566 are theoretically clean [44, 45, 46]. The CP-violation appears in the interference between
 567 the Cabibbo favoured $b \rightarrow c$ amplitude without mixing, $A(B^0 \rightarrow D^-\pi^+)$, and the
 568 Cabibbo suppressed $b \rightarrow u$ amplitude with mixing, $A(B^0 \rightarrow \bar{B}^0 \rightarrow D^-\pi^+)$. Two of the
 569 corresponding Feynman diagrams for these amplitudes are depicted in fig. 4.1.

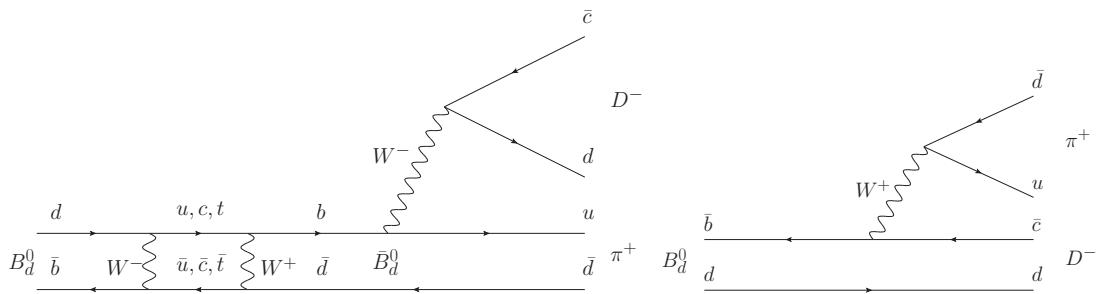


Figure 4.1 – Feynman diagrams contributing to $B^0 \rightarrow D^\pm\pi^\mp$

The measurement is performed by analysing the four decay-time dependent decay rates $\frac{d\Gamma(B^0 \rightarrow D^-\pi^+)}{dt}$, $\frac{d\Gamma(B^0 \rightarrow D^+\pi^-)}{dt}$, $\frac{d\Gamma(\bar{B}^0 \rightarrow D^-\pi^+)}{dt}$ and $\frac{d\Gamma(\bar{B}^0 \rightarrow D^+\pi^-)}{dt}$. Identifying the final state as $f = D^-\pi^+$ or $\bar{f} = D^+\pi^-$, and assuming CPT symmetry and no CP violation in mixing ($|q/p| = 1$), the time-dependent decay rates for B mesons initially produced as B^0 or \bar{B}^0

can be written as follows:

$$\frac{d\Gamma(B^0 \rightarrow f)}{dt}(t) = \frac{1}{4\tau} e^{-t/\tau} [1 + C_f \cos(\Delta m_d t) - S_f \sin(\Delta m_d t)], \quad (4.1)$$

$$\frac{d\Gamma(\bar{B}^0 \rightarrow f)}{dt}(t) = \frac{1}{4\tau} e^{-t/\tau} [1 - C_f \cos(\Delta m_d t) + S_f \sin(\Delta m_d t)], \quad (4.2)$$

$$\frac{d\Gamma(B^0 \rightarrow \bar{f})}{dt}(t) = \frac{1}{4\tau} e^{-t/\tau} [1 + C_{\bar{f}} \cos(\Delta m_d t) - S_{\bar{f}} \sin(\Delta m_d t)], \quad (4.3)$$

$$\frac{d\Gamma(\bar{B}^0 \rightarrow \bar{f})}{dt}(t) = \frac{1}{4\tau} e^{-t/\tau} [1 - C_{\bar{f}} \cos(\Delta m_d t) + S_{\bar{f}} \sin(\Delta m_d t)], \quad (4.4)$$

where ΔM and $\tau = 1/\Gamma$ are given by Eq. 1.23 and 1.24, while $\Delta\Gamma = 0$. The CP-coefficients $C_{f/\bar{f}}$ and $S_{f/\bar{f}}$, defined in Eq. 1.40 and 1.41, can be expressed as

$$S_f = -\frac{2r_{D\pi} \sin[\delta - (\gamma + 2\beta)]}{1 + r_{D\pi}^2}, \quad S_{\bar{f}} = \frac{2r_{D\pi} \sin[\delta + (\gamma + 2\beta)]}{1 + r_{D\pi}^2}, \quad (4.5)$$

$$C_f = -C_{\bar{f}} = C = \frac{1 - r_{D\pi}^2}{1 + r_{D\pi}^2}, \quad (4.6)$$

570 where β (Eq. 1.9) is related to the B^0 mixing phase, $r_{D\pi} = |A(\bar{B}^0 \rightarrow D^- \pi^+)/A(B^0 \rightarrow D^- \pi^+)|$
 571 is the magnitude of the ratio between doubly Cabibbo suppressed and favoured amplitudes,
 572 and δ is the strong phase difference between these amplitudes.

573 A measurement of γ can be obtained by measuring the CP-coefficients and taking external
 574 measurement of β and $r_{D\pi}$ as input. The angle β is known with very high precision, both
 575 theoretically and experimentally [47]. An estimation of $r_{D\pi}$ was performed by BaBar
 576 and Belle experiments [48, 49], by measuring the branching fraction of $B^0 \rightarrow D_s^{(*)+} \pi^-$
 577 decays and assuming SU(3) symmetry, yielding an average of about 1.7% with a relative
 578 error around 15%, mainly due to SU(3) symmetry breaking. For this analysis, C_f ($C_{\bar{f}}$) is
 579 simply fixed to 1 (-1) due to a lack of *sensitivity*.

580 The small value for the $r_{D\pi}$ parameter, which reduces the sensitivity on $S_{f/\bar{f}}$, make this
 581 measurement challenging compared to similar analyses like $B_s^0 \rightarrow D_s^\mp K^\pm$; however, the
 582 expected signal yield (about 5×10^5 signal events) is sufficiently large to compensate for
 583 this lack of sensitivity.

584 Measurements of $\sin(2\beta + \gamma)$ in $B^0 \rightarrow D^{(*)\pm} \pi^\mp$ were performed by the BaBar and Belle
 585 experiments [50, 51, 52, 53]. A measurement of $2\beta_s - \gamma$ was performed by LHCb in
 586 $B_s^0 \rightarrow D_s^\mp K^\pm$ decays by analysing 1 fb^{-1} of data [54].

587 Analysis strategy

588 The measurement is performed in terms of a flavour tagged, decay-time dependent analysis
 589 of the Run 1 dataset. The dataset includes two sub-samples recorded with opposite

590 directions of the magnetic field ("up" and "down") in the spectrometer dipole. The
591 selection of the data, which is explained in detail in Sec. 4.2, includes the use of vetoes
592 to reduce the number of components that must be modelled in the sample, and a BDT
593 to reduce the amount of *combinatorial* background. The expected sample composition
594 after the selection is discussed in Sec. 4.3 based on studies with simulated samples. A
595 fit to the invariant mass distribution of the resulting dataset is performed to extract
596 sWeights for the signal component. The fit is described in detail in Sec. 4.4. The training
597 and calibration of the flavour tagging algorithms, which infer the initial flavour of the
598 reconstructed B^0 candidates, is summarised in Sec. 3.1. The final estimation of the
599 CP-coefficients is the result of an unbinned, sWeighted likelihood fit to the distributions
600 of the decay time and the flavour tagging observables.

601 4.2 Data sample and selection

602 The sample of data is passed through the following selection steps:

- 603 1. stripping and trigger requirements;
- 604 2. a *cut-based* preselection;
- 605 3. vetoes for misidentified backgrounds and wrongly associated *primary vertices* (PVs);
- 606 4. a multivariate classification (MVA);
- 607 5. a final randomised multiple candidate selection.

608 In what follows, the details of each step are provided.

609 4.2.1 Stripping and trigger requirements

610 Signal $B^0 \rightarrow D^\pm \pi^\mp$ candidates are reconstructed using a dedicated stripping line (called
 611 `B02DPiD2HHHBeauty2CharmLine`). Each event is required to have less than 500 Long tracks. The criteria that the charged tracks have to fulfill are listed in Tab. 4.1. Three

Table 4.1 – Stripping requirements applied in the selection of charged tracks. The more stringent requirements for the bachelor track are given in brackets.

χ^2/ndof	$< 3.0 (< 2.5)$
momentum p	$> 1 \text{ GeV}/c (> 5 \text{ GeV}/c)$
transverse momentum p_T	$> 100 \text{ MeV}/c (> 500 \text{ MeV}/c)$
$\text{IP}\chi^2$ w.r.t. any PV	> 4.0
ghost probability	< 0.4

612
 613 of these hadrons have to form a common vertex to build a D meson. The (transverse)
 614 momentum of one of the daughter hadrons has to exceed ($500 \text{ MeV}/c$) $5 \text{ GeV}/c$ and its
 615 track χ^2/ndof has to be less than 2.5. The further requirements on the D^\pm combination are
 616 given in Tab. 4.2. The B candidates are built by combining a D candidate and a bachelor
 617 particle if the requirements listed in Tab. 4.3 are fulfilled. Finally, a bagged boosted
 618 decision tree (BDT) classifier, which is trained on simulated candidates, is applied [55]. A
 619 minimum value of 0.05 is required for the output value of the BDT. Stripping candidates
 620 are then filtered according to how they were selected at the trigger level: no specific
 621 requirements are made at L0; at HLT1, B candidates are required to be TOS on the
 622 `Hlt1TrackAllL0Decision` trigger line; at HLT2, B candidates are required to be TOS on
 623 one of the following lines: `Hlt2Topo2BodyBBDTDecision`, `Hlt2Topo3BodyBBDTDecision`
 624 or `Hlt2Topo4BodyBBDTDecision`. The details of these trigger lines are described in
 625 Refs. [56, 57].

Table 4.2 – Stripping requirements on the D^\mp combination. DOCA is the Distance Of Closest Approach of the daughter particles w.r.t. each other, and DIRA indicates the cosine of the angle between the momentum of the D meson and the direction from the best PV to the decay vertex.

$\sum p_T(hhh)$	$> 1800 \text{ MeV}/c$
DOCA	$< 0.5 \text{ mm}$
m_D^\mp	1769.62 to 2068.49 MeV/c^2
$D^\mp \chi_{\text{vtx}}^2/\text{ndof}$	< 10.0
vertex separation χ^2 to any PV	> 36
DIRA	> 0.0

Table 4.3 – Stripping requirements on the B^0 combination.

SV χ^2/ndof	< 10.0
reconstructed decay time t	$> 0.2 \text{ ps}$
IP χ^2 w.r.t. the associated PV	< 25.0
DIRA	> 0.999

Figure 4.2 shows the B and D mass distributions of the reconstructed candidates after the stripping and trigger selection. In the B mass distribution the signal peak is already visible at around $5280 \text{ MeV}/c^2$. The structure at masses lower than the B peak originates from partially reconstructed $B^0 \rightarrow D\rho$ and $B^0 \rightarrow D^*\pi$ decays. The D mass distributions features a clearly visible D peak at $1870 \text{ MeV}/c^2$ and a D^* peak around $2010 \text{ MeV}/c^2$.

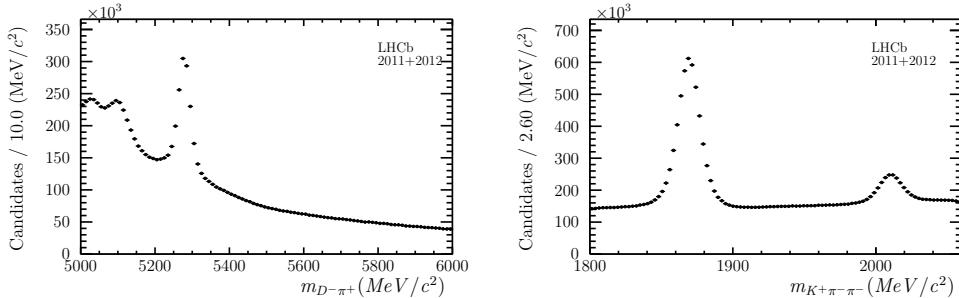


Figure 4.2 – B and D mass distributions of the reconstructed candidates after the stripping and trigger selection.

630

4.2.2 Preselection and samples definition

Additional preselection criteria (shown in Tab. 4.4) are applied offline. In order to obtain the correct correlations between and uncertainties on vertex positions, particle momenta, flight distances, decay times, and invariant masses, a Kalman filter, known as

635 DecayTreeFitter (DTF), is used [58]. The decay time related observables are derived
 636 from a DTF fit where the knowledge about the primary vertex has been used to constrain
 637 the production vertex of the B^0 meson. To determine the momenta and the invariant
 638 mass of the B^0 meson, the invariant mass of the D meson is constrained to the PDG
 value ($m_{D^\pm}^{\text{PDG}} = 1869.61 \text{ MeV}/c^2$ [20]) in a separate DTF fit. The PIDK variable 2.1 is

Table 4.4 – Offline preselection criteria

Cuts on B^0 candidate	
decay time	$> 0.2 \text{ ps}$
Cuts on D meson	
$ m(K^+\pi^-\pi^-) - m_D $	$< 35 \text{ MeV}/c^2$
Cuts on D daughters	
PIDK for pions	< 8
PIDK for kaon	> -2

639 used to identify the kaon and the pions from the D decays, and to identify the bachelor
 640 pion from the B decay. The requirement on the PIDK of the bachelor pion define two
 641 categories of candidates: those of the so-called the *pion sample* ($\text{PIDK} \leq 5.0$) and those
 642 of the *kaon sample* ($\text{PIDK} > 5.0$). This distinction will be useful in the fit to the B
 643 mass distribution for determining the sample composition described in Sec. 4.4. All the
 644 following step of the selection are applied to data without distinction in pion and kaon
 645 samples.

647 4.2.3 Veto for physics backgrounds

648 Misidentification of kaons and protons as pions lead to exclusive backgrounds in the data
 649 sample. These are suppressed by means of explicit *veto*s.

650 In order to reduce contributions from semileptonic decays such as $B^0 \rightarrow D^- (\rightarrow K^+\pi^-\pi^-) \mu^-\bar{\nu}_\mu$,
 651 the bachelor pion is required to have no hits in the muon chambers.

652 For the D^\pm mesons which are reconstructed in the $D^\pm \rightarrow K^\mp\pi^\pm\pi^\pm$ final state, a $p \rightarrow \pi$
 653 mis-ID can lead to background contributions from $\Lambda_b^0 \rightarrow \Lambda_c^\pm (\rightarrow K^\mp\pi^\pm p) \pi^\mp$. To reduce
 654 these contributions, the proton mass hypothesis is applied separately to both pions from
 655 the D^\pm final state. The invariant mass of the three hadrons is recalculated and if the
 656 candidate is inside a $30 \text{ MeV}/c^2$ window around the Λ_c mass, $m_{\Lambda_c} = 2286.46 \text{ MeV}/c^2$ [20],
 657 it is required to have $\text{PIDp} < -8.0$ and if it is inside a $50 \text{ MeV}/c^2$ window around the Λ_c
 658 mass it is required to have $\text{PIDp} < -5.0$. A plot showing the distributions before and
 659 after applying the veto is given in Fig. 4.3. This requirement shows a signal efficiency
 660 of $(93.48 \pm 0.06)\%$. For the rejection of $\Lambda_b^0 \rightarrow \Lambda_c^\pm\pi^\mp$, the efficiency on a signal MC
 661 for this decay is checked. After stripping and preselection alone, $(99.720 \pm 0.004)\%$ of

the $\Lambda_b^0 \rightarrow \Lambda_c^\pm \pi^\mp$ candidates are rejected; this veto rejects another $(76.6 \pm 0.6)\%$ of the $\Lambda_b^0 \rightarrow \Lambda_c^\pm \pi^\mp$ candidates.

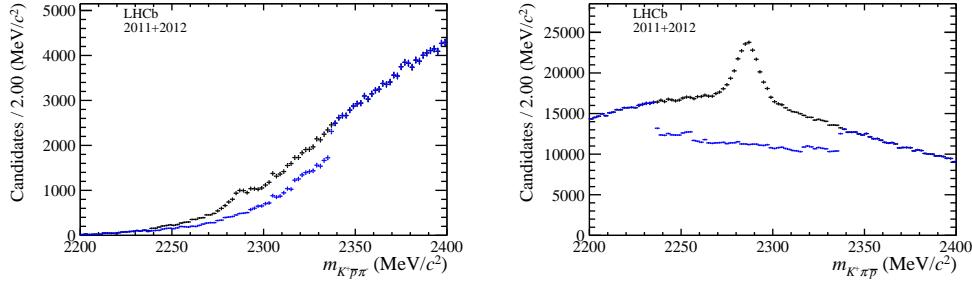


Figure 4.3 – Plots of the invariant mass of the $K\pi\bar{p}$ combinations for both daughter pions of the D meson misidentified as protons. The distribution is given without (black) and with (blue) the Λ_c^\pm veto described in the text. The distributions look different for both plots because they are sorted by p_T . (Left the proton mass hypothesis is applied to pions with lower p_T , right to pions with higher p_T).

In the same way as protons are misidentified as pions, it is possible for kaons to be misidentified as pions. Such a mis-ID would lead to background contributions from $B_s^0 \rightarrow D_s^\pm (\rightarrow K^\mp K^\pm \pi^\pm) \pi^\mp$. To check for these contributions, the kaon mass hypothesis is applied separately to both pions from the D^\pm final state. The invariant mass of the resulting $K^\mp K^\pm \pi^\pm$ system is recalculated and plotted for two different ranges of the B^0 mass: The first range, from 5330 to 5400 MeV/c^2 , covers the invariant mass signal region of the B_s^0 meson as possible background contribution. The second range, from 5500 to 5700 MeV/c^2 is the upper mass *sideband* for this possible background contamination. As can be seen in Fig. 4.4 the distribution which covers the invariant mass of the B_s^0 meson does not show any significant peaking structure compared to the upper mass sideband region. The visible differences are expected as the distributions arise from different kinematic ranges. To double check for possible resonant contributions from a kaon mis-ID, the decay of the D meson after applying the kaon mass hypothesis is investigated. Possible resonant decays of the D meson can take place via a K^* or ϕ resonance. These resonances would be visible in the $K\pi$ and KK invariant mass distributions, which are plotted in the same two ranges in fig. 4.5. As the distributions in the signal and background range look compatible, the D_s^+ contamination is negligible and no veto is applied.

In the same way as for the D meson daughters, it is also possible that the bachelor pion candidate is actually a misidentified kaon. This mis-ID could lead to background contributions of $B^0 \rightarrow D^0 (\rightarrow K^\mp \pi^\pm) K^+ \pi^-$. The same background contribution could arise if one D meson daughter pion is misidentified as a kaon and combined with the bachelor pion. To check for this contribution, the kaon mass hypothesis is applied to the bachelor pion and the D meson daughter pions, and the invariant mass distributions for the four possible $K\pi$ systems are plotted after applying the BDT classifier described in Sec. 4.2.5 (fig. 4.6). As the distributions show no significant peaking structures, this

4.2. Data sample and selection

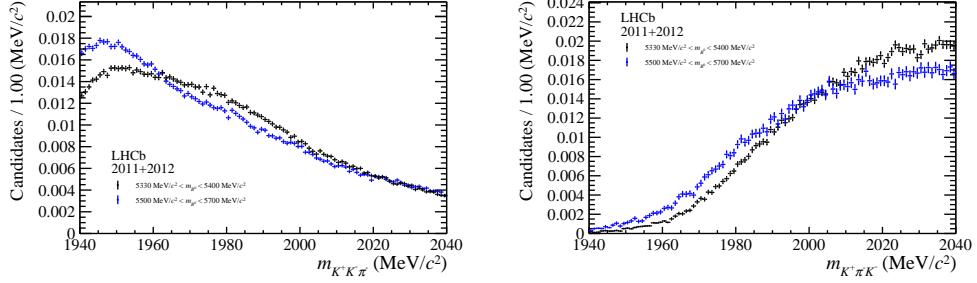


Figure 4.4 – Plots of the invariant mass of the $KK\pi$ combinations for both daughter pions of the D meson misidentified as kaons. The distribution is given in the B^0 meson mass range from 5300 to 5400 MeV/ c^2 (black) and in the B^0 meson mass range from 5400 to 5500 MeV/ c^2 (blue) after applying the Λ_c^+ veto. The distributions look different for both plots because they are sorted by p_T . (Left the kaon mass hypothesis is applied to pions with lower p_T , right to pions with higher p_T).

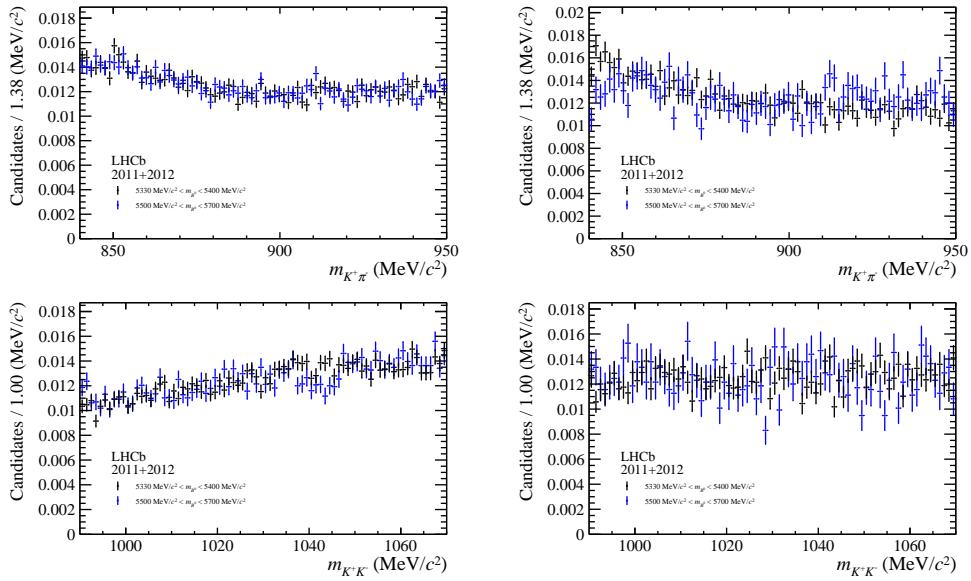


Figure 4.5 – Plots of the invariant mass of the $K\pi$ combination (top) and for the KK combination (bottom) for both daughter pions of the D meson. The distributions are shown in the B^0 meson mass range from 5330 to 5400 MeV/ c^2 (black) and in the B^0 meson mass range from 5500 to 5700 MeV/ c^2 (blue) as described in the text. Additionally, for all plots a cut on the D_s^+ mass hypothesis from 1940 to 2040 MeV/ c^2 is applied. The distributions look different for both plots because they are sorted by p_T . (Left the kaon mass hypothesis is applied to pions with lower p_T , right to pions with higher p_T).

contribution is neglected at this stage and no specific cuts are applied.

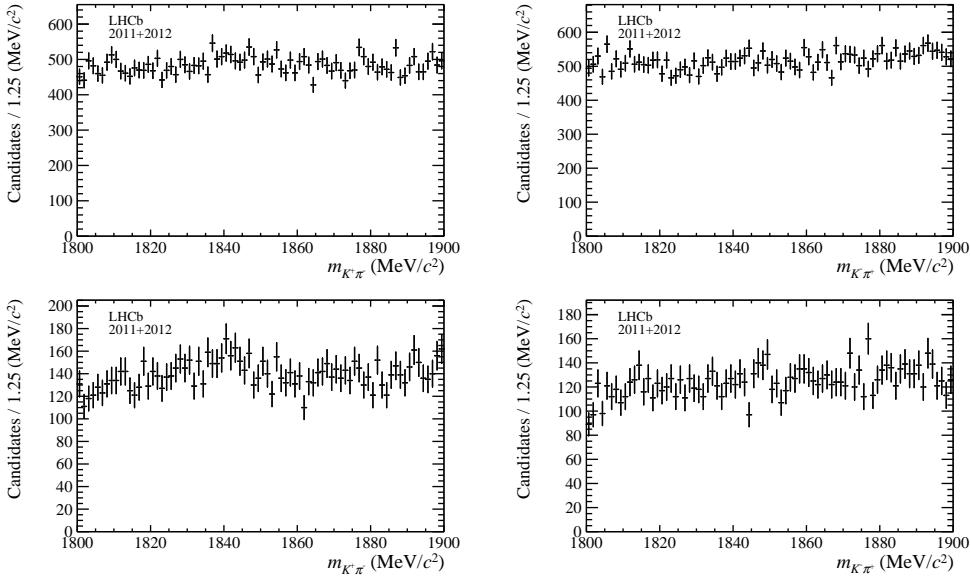


Figure 4.6 – Plots of the invariant mass of the four possible $K\pi$ combinations after applying a kaon mass hypothesis to the bachelor pion. In the top plots the bachelor pion is combined with the D meson daughter pion with lower p_T , in the bottom plots the combination with the D meson daughter pion with the higher p_T is shown. In the left plots the kaon mass hypothesis is applied to the bachelor pions and in the right plots to the D meson daughter pions.

689

690 4.2.4 Wrongly associated PVs

691 Given an average number of pp interactions per bunch crossing of $\nu = 2.5$, a large fraction
 692 of events have more than one reconstructed PV. The PV to which the B^0 candidate
 693 has the smallest χ_{IP}^2 (*best* PV) is chosen. Furthermore, It is possible that more than
 694 one B^0 candidate is reconstructed per event: in this case, the B^0 candidate is chosen
 695 randomly following the prescription of Ref. [59], which allows to prevent any bias due to
 696 an arbitrary choice.

697 In events where the association of the B^0 candidate to the best PV B^0 is wrong, the
 698 decay time of this candidate will be incorrect. These wrongly associated candidates are
 699 responsible for a large tail that can be observed in the decay time distribution, particularly
 700 in signal MC where the true decay time is known. Weighting each candidate with an
 701 exponential with the true lifetime shows an excess of candidates at high decay times.
 702 To remove these incorrect associations in MC, one can compare the z -position of the
 703 associated PV with the z position of the true PV and reject the candidate if the distance
 704 between those positions exceeds 5 times its uncertainty (fig. 4.7). In real data, the the

4.2. Data sample and selection

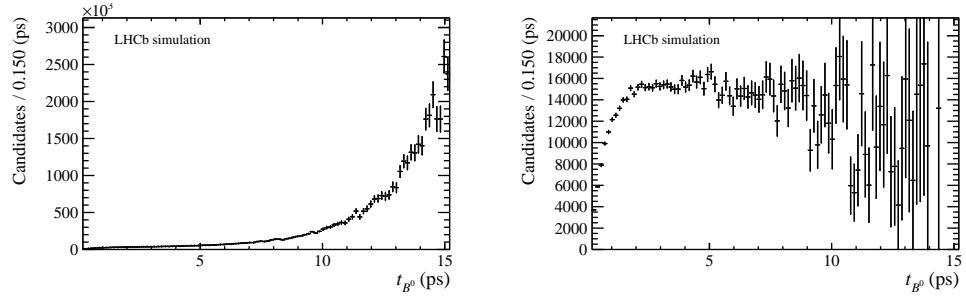


Figure 4.7 – Left: Signal MC decay time distribution weighted with an exponential with the true lifetime. At high decay times an excess of candidates can be observed. Right: Signal MC decay time distribution weighted with an exponential with the true lifetime. In addition, it is required that the difference of the PV z position with the true PV z position is within 5 times the PV uncertainty. As the excess of candidates at high decay times vanishes (from left to right), it can be concluded that these candidates are wrongly associated to their PV.

705 true PV in unknown, so a selection cut on the smallest B^0 impact parameter $\chi^2_{\text{DTF,PV}}$ (called $\text{MinIP}\chi^2$) with respect to any other PV in the event is applied (fig. 4.8). Events

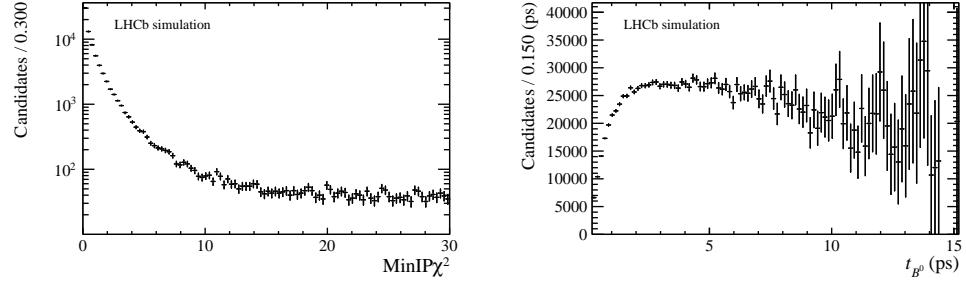


Figure 4.8 – Left: distribution of the smallest B^0 impact parameter χ^2_{DTF} with respect to any other primary vertex in the event. Right: Signal MC decay time distribution weighted with an exponential with the true lifetime after requiring that the smallest B^0 impact parameter χ^2_{DTF} to any other PV in the event is larger than 16.5.

706
 707 with only a single PV are not affected. The main advantage of this $\text{MinIP}\chi^2$ variable
 708 is that all PVs are treated equally, without any biasing choice. The requirement on
 709 $\text{MinIP}\chi^2$ is optimised to retain 98 % of the truth-matched signal candidates in MC. The
 710 optimal cut is then found at $\text{MinIP}\chi^2 > 16.5$. A plot showing the weighted decay time
 711 distribution after applying the cut on the smallest B^0 impact parameter $\chi^2_{\text{DTF,PV}}$ to any
 712 other PV is given in Fig. 4.8.

713 4.2.5 Development of an MVA classifier

714 The combinatorial background, consisting of fake signal candidates created from random
 715 combinations of tracks, is rejected by using a Boosted Decision Tree (BDT) classifier [60,
 716 61]. The signal input to the training stage consists of signal MC candidates simulated
 717 under 2012 data taking conditions, while the upper mass side band above 5500 MeV/ c^2
 718 from the 2012 data sample is used as template for the combinatorial background. The
 719 BDT is trained on one half of these samples, the other half being used to test the
 720 performance of the BDT. Before the BDT training, all previous steps (the cut-based
 721 preselection, the mass vetoes and the wrongly associated PV veto) are applied. To reduce
 722 the number of input features, the ones with a correlation larger than 97 % are removed.
 The 16 final input features are listed in Tab. 4.5. The correlation matrices between the

Table 4.5 – List of input features used in the training of the BDT

B^0 candidate	$\cos \varphi [PV, B^0 Vtx], \vec{p}(B^0)]$ χ^2 on B^0 vertex
	χ^2_{IP} w.r.t. B^0 vertex
	χ^2_{IP} w.r.t. associated PV
	radial flight distance
D candidate	flight distance χ^2 w.r.t. B^0 vertex
	χ^2/ndof on D vertex
	transverse momentum
	$\cos \varphi [B^0 Vtx, D Vtx], \vec{p}(D)]$
bachelor π	χ^2_{IP} w.r.t. the associated PV
	transverse momentum
	track χ^2/ndof
D daughters	χ^2_{IP} of associated primary vertex
DTF	χ^2 of DTF with PV constraint

723
 724 input features in the signal and the background sample are shown in Fig.4.9, while the
 725 distributions of the input features can be found in Appendix 4.10.1.

726 The BDT implementation of TMVA [62] is used. The BDT is built out of 1700 trees,
 727 with a depth limited to four. For each node, at least 2.5 % of the training events have to
 728 be present. The variables are scanned at 20 points to find the optimal cut value. The
 729 chosen boosting method is the AdaBoost [63] algorithm with a boost factor $\beta = 0.5$. The
 730 number of trees and the maximal depth of trees have been increased iteratively until no
 731 significant increase of the performance or overtraining was observed. The BDT is tested
 732 on the events which are not used in the training. The plot of this overtraining check is
 733 given in Fig.4.10.

4.2. Data sample and selection



Figure 4.9 – Correlation matrices of the input features used in the training of the BDT for signal (left) and background (right).

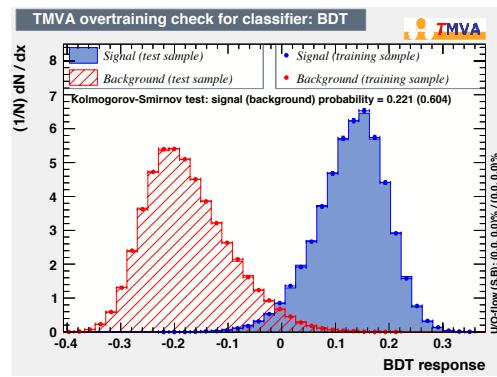


Figure 4.10 – Comparison of BDT response on training and test sample.

⁷³⁴ **4.2.6 BDT selection optimisation**

⁷³⁵ To estimate the best requirement on the output of the BDT classifier, the statistical
⁷³⁶ uncertainty of the CP-coefficients derived from the analysis of simulated samples is used
⁷³⁷ as the figure of merit (FoM). To determine the sensitivity, the preselection, the mass
⁷³⁸ vetoes and the wrongly associated PV veto are applied and the BDT classifier is calculated
⁷³⁹ for every candidate. The BDT cut point is then scanned with a step size of 0.01 from
⁷⁴⁰ -0.15 to 0.10 and a step size of 0.05 in the outer regions. For each cut point of the BDT
⁷⁴¹ classifier, a simulated (*toy*) sample is generated. This simulated sample contains the same
⁷⁴² signal and combinatorial background yields as determined from the real dataset via a
⁷⁴³ maximum likelihood fit of the B^0 mass distribution. Finally, a time-dependent analysis of
⁷⁴⁴ each toy dataset is performed in order to estimate the statistical uncertainty on S_f and
⁷⁴⁵ $S_{\bar{f}}$. These statistical uncertainties as a function of the BDT cut are shown in Fig. 4.11.
 Based on these distributions, the BDT cut point is chosen to be at 0.05.

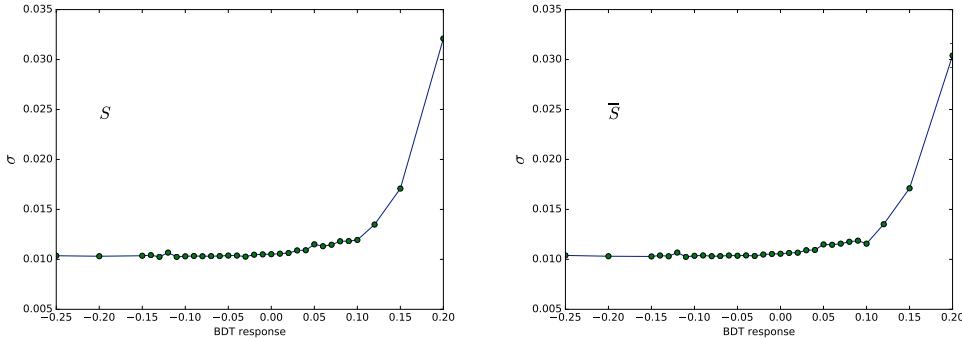


Figure 4.11 – Sensitivity of S_f (left) and $S_{\bar{f}}$ (right) as a function of the BDT output classifier.

⁷⁴⁶

⁷⁴⁷ **4.2.7 Multiple candidates**

⁷⁴⁸ After the stripping selection, approximately 9 % of events contain at least two B^0 can-
⁷⁴⁹ didates, and 18 to 20 % of all B^0 candidates share an event. If the offline selection is
⁷⁵⁰ also applied, around 0.4 % of events contain multiple B^0 candidates, and 0.8 % of all B^0
⁷⁵¹ candidates share an event. More details are given in Appendix 4.10.2. In order to be
⁷⁵² consistent with the prescription used in the stripping and trigger requirements, only the
⁷⁵³ best PV is chosen; all events in which the best PV is no longer present after the offline
⁷⁵⁴ selection are removed. Finally, since the remaining B^0 candidates are considered to be
⁷⁵⁵ equally likely signal candidates, one remaining B^0 candidate is chosen randomly.

756 4.2.8 Selection performance

757 The offline selection performances are listed in Tab. 4.6. They are determined by using
 758 data candidates of the 2012 sample with an invariant B^0 mass above $5500 \text{ MeV}/c^2$ to
 759 represent combinatorial background, and signal MC candidates (see Sec. 4.3) to represent
 760 the signal. In fig. 4.12 a comparison of the mass distribution of the selected 2011 and
 2012 data is shown. Additionally, the BDT performances are quoted in Tab. 4.7 split

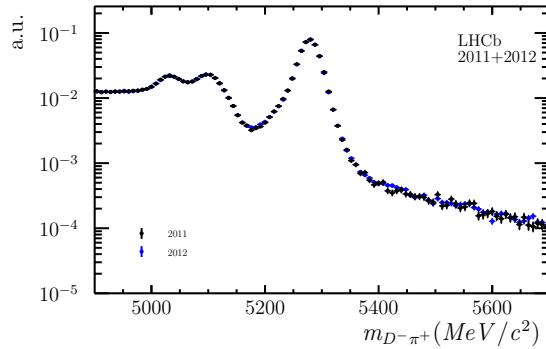


Figure 4.12 – B mass distributions for the 2011 (blue) and 2012 (black) samples after the full offline selection.

Table 4.6 – Offline selection performances. The signal efficiencies and background rejections of the different selection steps are given with respect to the previous selection step, while the last row shows the overall selection performance.

Selection step	ε_{sig}	$1 - \varepsilon_{\text{bkg}}$
preselection	$(93.61 \pm 0.06) \%$	$(85.20 \pm 0.02) \%$
Λ_c^\pm -veto	$(93.48 \pm 0.06) \%$	$(9.85 \pm 0.03) \%$
semileptonic veto	$(98.96 \pm 0.03) \%$	$(7.66 \pm 0.03) \%$
mass vetoes combined	$(92.51 \pm 0.07) \%$	$(16.77 \pm 0.04) \%$
wrongly associated PVs veto	$(97.75 \pm 0.04) \%$	$(15.81 \pm 0.04) \%$
BDT selection	$(83.63 \pm 0.10) \%$	$(97.18 \pm 0.01) \%$
total	$(70.7 \pm 0.1) \%$	$(99.911 \pm 0.002) \%$

761 by both polarity and year of data taking. Finally, in order to check the contribution
 762 of non-resonant decays $B^0 \rightarrow K^+\pi^-\pi^-\pi^+$, the B^0 and D^- invariant mass distributions
 763 are analysed after applying the full offline selection in two ways. First, the D^- mass
 764 distribution after applying a signal window cut on the B^0 mass is plotted (Fig. 4.13).
 765 From this plot, the maximal contamination with non-resonant decays can be estimated as
 766 roughly 1 %. Then, the B^0 distribution after excluding the D^- signal window is plotted.
 767 To quantify the non-resonant $B^0 \rightarrow K^+\pi^-\pi^-\pi^+$ decays, the sum of an exponential and a

Chapter 4. Measurement of CP violation in $B^0 \rightarrow D^\mp \pi^\pm$ decays

Table 4.7 – BDT performance split by polarity and year of data taking. The given efficiencies contain signal and background.

	2011, MU	2011, MD	2012, MU	2012, MD
# cand. before BDT	398 357	569 853	1 301 800	1 316 597
# cand. after BDT	210 844	285 137	601 345	609 880
$\varepsilon_{\text{sig+bkg}}$	$(50.67 \pm 0.08)\%$	$(50.04 \pm 0.07)\%$	$(46.19 \pm 0.04)\%$	$(46.32 \pm 0.04)\%$

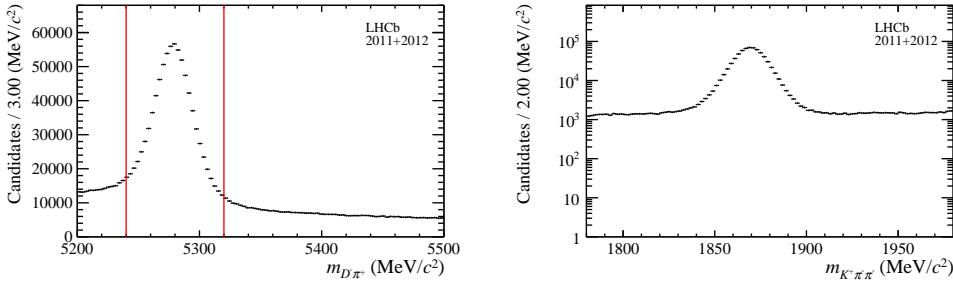


Figure 4.13 – Left: B^0 mass distribution. The red vertical lines indicate the selected signal window. Right: resulting D^- mass distribution.

769 Gaussian with a fixed shape is used to fit the resulting B^0 mass distribution, as shown in
 770 Fig. 4.14. As the fitted B^0 yield is 645 ± 242 , the non-resonant contribution is assumed
 to be negligible.

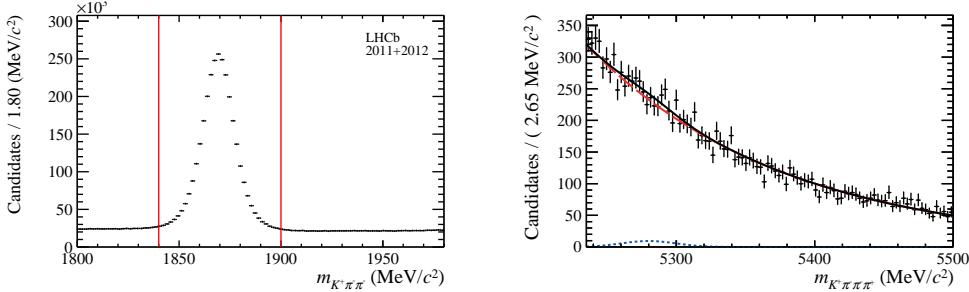


Figure 4.14 – Left: D^- mass distribution, The red vertical lines indicate the excluded range. Right: resulting B^0 mass distribution with the fitting function overlaid.

771

4.3 Simulation and expected sample composition

Simulated samples are used to (i) gain a detailed overview of all sources of b -hadron decays that contribute to the sample and (ii) model the relevant distributions studied in the analysis. Simulated data undergoes the same processing of reconstruction and selection as real data. Each sample is split into "up" and "down" magnet polarities and into 2011 and 2012 data taking conditions subsamples, in proportions similar to those present in real data.

The used simulated samples are listed in Tab. 4.8, together with the number of true signal events passing the final selection and the corresponding total efficiencies. Any PID cut on the bachelor cut is applied in order to compute these efficiencies.

Sample	Event type	N_{sel}	Efficiency [%]
$B^0 \rightarrow D\pi$	11164003	101096	1.966 ± 0.006
$B^0 \rightarrow DK$	11264011	19300	1.833 ± 0.013
$B^0 \rightarrow D\rho$	11164401	2408	0.1178 ± 0.0024
$B^0 \rightarrow D^*\pi$	11164404	14901	0.721 ± 0.006
$B_s^0 \rightarrow D_s^+\pi^-$	13264021	7942	0.1531 ± 0.0017
$\Lambda_b^0 \rightarrow \Lambda_c^0\pi^-$	15164001	325	0.0155 ± 0.0009
$B^0 \rightarrow DK^*$	11164470	361	0.0358 ± 0.0019

Table 4.8 – Samples of simulated data used in the analysis, number of candidates, N_{sel} , after selection, and selection efficiency. Efficiencies include generator level, stripping, offline selection and tagging efficiencies.

781

4.3.1 PIDK correction

783 The PIDK distributions in data and MC differ. To correct for that, the PIDK distributions
 784 in MC are resampled using the PIDK distributions of dedicated calibration samples as
 785 (binned) probability density functions. These calibration samples consist of kinematically-
 786 clean $D^* \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ data, for which no requirement on RICH information is
 787 applied in the reconstruction.

788 The need for this resampling is due to the fact that, if the same cut is applied on data and
 789 MC, the resulting distributions in other observables may differ if the PIDK distributions
 790 in data and MC are different. Moreover, a correct PIDK distribution in MC allow to
 791 properly estimate (on MC) the efficiency or misidentification rate for a given PIDK cut,
 792 which is an essential ingredient in the fit to the B invariant mass distribution (as described
 793 in Sec. 4.4).

794 The following strategy is adopted. A two-dimensional binning in momentum, p , and
 795 pseudorapidity, η , is defined. For each bin, the corresponding PIDK distribution in the

796 calibration sample is built and for each event in the MC sample, a random PIDK value is
797 sampled from the PIDK distribution associated to the corresponding bin in the calibration
798 sample. More details are given in Appendix 4.11.

799 Because of the Λ_c veto described in Sec. 4.2.3, the PIDp variable for the D daughter
800 particles is resampled as well in a similar manner using $\Lambda^0 \rightarrow p\pi^-$ decays as calibration
801 samples.

802 4.3.2 Surviving physics backgrounds

803 Some physics background candidates that survive the selection chain described in the
804 previous section are expected. In the pion sample, these are:

- 805 • $B^0 \rightarrow D^\mp K^\pm$: Peaking background due to the bachelor kaon being wrongly
806 identified as a pion.
- 807 • $B^0 \rightarrow D^\mp \rho^\pm (\rightarrow \pi^\pm \pi^0)$: Low mass background due to missing a pion in the
808 reconstruction.
- 809 • $B^0 \rightarrow D^\mp (\rightarrow D^\mp \gamma/\pi^0) \pi^\pm$: Low mass background due to a missed neutral
810 particle.

811 In the kaon sample, the following backgrounds are expected:

- 812 • $B^0 \rightarrow D^\mp \pi^\pm$: Signal candidates having the bachelor pion wrongly identified as a
813 kaon.
- 814 • $B^0 \rightarrow D^\mp \rho^\pm (\rightarrow \pi^\pm \pi^0)$: Low mass background where, in addition to the missed
815 pion in the final state, a reconstructed pion is wrongly identified as a kaon.
- 816 • $B^0 \rightarrow D^\mp K^{\pm*} (\rightarrow \pi^0 K^\pm)$: Low mass background where the neutral pion is missed
817 in the final state.

818 The fraction of the backgrounds expected in the pion sample with respect to the $B^0 \rightarrow D\pi$
819 signal are reported in Tab. 4.9. These fractions are computed using the branching fractions
820 of the expected decay as inputs and from the ratio of efficiencies estimated from MC
821 (when relevant we consider also the ratio of the fragmentation probabilities of b quarks to
822 different b hadron). These expectation will be compared with the results from the fit to
823 data described in the next section.

824 The $B_s \rightarrow D_s \pi$ and $\Lambda_b \rightarrow \Lambda_c (\rightarrow p K \pi) \pi$ backgrounds are suppressed to a negligible
825 fraction by the offline selection described in Sec. 4.2, and are thus ignored in the description

4.3. Simulation and expected sample composition

Decay	\mathcal{B} [%]	ϵ_{bkg} [%]	$f_{Bz/B_s^0/\Lambda_b^0}$ [%]	f_{bkg} [%]
$B^0 \rightarrow DK$	0.00186 ± 0.00020	0.684 ± 0.008	33.9 ± 3.9	2.61 ± 0.31
$B^0 \rightarrow D\rho(\rightarrow \pi\pi)$	0.071 ± 0.011	0.1149 ± 0.0024	33.9 ± 3.9	16.7 ± 2.8
$B^0 \rightarrow D^*(\rightarrow D\gamma/\pi^0)\pi$	0.0080 ± 0.0004	0.705 ± 0.006	33.9 ± 3.9	11.6 ± 0.8
$\Lambda_b \rightarrow \Lambda_c(\rightarrow pK\pi)\pi$	0.032 ± 0.004	0.0150 ± 0.0008	21.2 ± 6.9	0.62 ± 0.24
$B_s^0 \rightarrow D_s\pi$	0.0164 ± 0.0014	0.1493 ± 0.0017	11.1 ± 1.4	1.64 ± 0.32

Table 4.9 – Background contributions expected in the pion sample. The $B^0 \rightarrow D\pi$ branching ratio and total selection efficiency in the pion sample are $(0.254 \pm 0.014)\%$ and $(1.924 \pm 0.006)\%$ respectively.

826 of the sample composition. Moreover, in the kaon sample, the $B^0 \rightarrow D^*\pi$ and $B^0 \rightarrow D^*K$
827 components, which are expected to be negligible, are neglected as well ¹.

¹More precisely, these components are taken into account by the PDF describing $B^0 \rightarrow DK^*(\rightarrow \pi^0 K)$, since they are expected to sit in the same mass region.

4.4 Fits to the B^0 invariant mass

The *sPlot* technique [64] is applied in order to statistically disentangle signal from background for the decay time fit. The $D\pi$ invariant mass, where the D mass is constrained to its known value in order to improve the mass resolution, is adopted as discriminating observable thanks to its low correlation with the B^0 decay time (see Appendix 4.12).

In a first step, a binned extended maximum likelihood fit (“Fit A”) is performed in order to define the PDFs describing signal and background components. The choice of a binned fit is justified by the very high statistics of the data sample. The invariant mass range of the fit is $[5090, 6000]$ MeV/c². Only tagged candidates are considered, i.e. candidates with at least a nonzero tagging decision among OS or SS taggers. The reason for this is that untagged candidates do not contribute to the sensitivity on CP-coefficients. The fit is performed simultaneously for the pion sample and the kaon sample (see section 4.2.2). This approach is adopted in order to control the contamination of $B^0 \rightarrow DK$ background in the pion sample. The number of $B^0 \rightarrow DX$ candidates in the Y sample (with $X, Y = \pi, K$), $N_{B^0 \rightarrow DX}^Y$, can be defined via the following relations:

$$\begin{aligned} N_{B^0 \rightarrow D\pi}^K &= \frac{\epsilon_{\text{PID}}(B^0 \rightarrow D\pi)_K}{\epsilon_{\text{PID}}(B^0 \rightarrow D\pi)_\pi} \times N_{B^0 \rightarrow D\pi}^\pi \\ &= \frac{1 - \epsilon_{\text{PID}}(B^0 \rightarrow D\pi)_\pi}{\epsilon_{\text{PID}}(B^0 \rightarrow D\pi)_\pi} \times N_{B^0 \rightarrow D\pi}^\pi, \end{aligned} \quad (4.7)$$

$$\begin{aligned} N_{B^0 \rightarrow DK}^\pi &= \frac{\epsilon_{\text{PID}}(B^0 \rightarrow DK)_\pi}{\epsilon_{\text{PID}}(B^0 \rightarrow DK)_K} \times N_{B^0 \rightarrow DK}^K \\ &= \frac{1 - \epsilon_{\text{PID}}(B^0 \rightarrow DK)_K}{\epsilon_{\text{PID}}(B^0 \rightarrow DK)_K} \times N_{B^0 \rightarrow DK}^K. \end{aligned} \quad (4.8)$$

The quantities $\epsilon_{\text{PID}}(B^0 \rightarrow DX)_Y$ are the fractions of true $B^0 \rightarrow DX$ candidates that are selected in the Y sample by applying the relative PIDK cut. These fractions (or efficiencies) are estimated on $B^0 \rightarrow D\pi$ and $B^0 \rightarrow DK$ MC samples where the PIDK distributions are resampled from calibration data, as described in section 4.3.1; the results of these estimations are reported in table 4.10.

Table 4.10 – Fractions of true $B^0 \rightarrow D\pi$ and $B^0 \rightarrow DK$ decays that are selected in the π or K sample.

Sample	PIDK cut	fraction
$B^0 \rightarrow D\pi$	< 5.0 (π sample)	0.979 ± 0.004
$B^0 \rightarrow D\pi$	> 5.0 (K sample)	0.0211 ± 0.0005
$B^0 \rightarrow DK$	< 5.0 (π sample)	0.373 ± 0.005
$B^0 \rightarrow DK$	> 5.0 (K sample)	0.637 ± 0.007

Finally, an unbinned extended maximum likelihood fit (“Fit B”) is performed on data using the reduced mass interval [5220, 5600] MeV/ c^2 in order to extract *sWeights*. In this second fit, all the parameters are fixed to the values found in Fit A, except for the normalisations. The reduced mass window avoids diluting the *sWeights* with components that are neither combinatoric nor close to the signal component and therefore cannot be meaningfully subtracted. This has the added advantage of reducing the dataset size used in the decay time fit.

4.4.1 Probability density functions

Before Fit A, the PDFs used to describe both the pion and kaon sample composition are first estimated on MC samples. The parameters of the combinatorial background PDFs are instead determined directly on data. The PDFs used for the pion sample are:

- $B^0 \rightarrow D\pi$: sum of a double-sided Hypatia [65] and a Johnson SU [66] functions ($PDF_{B^0 \rightarrow D\pi}^\pi$).
- $B^0 \rightarrow DK$: double-sided Hypatia function ($PDF_{B^0 \rightarrow DK}^\pi$).
- $B^0 \rightarrow D\rho$: Johnson SU function ($PDF_{B^0 \rightarrow D\rho}^\pi$).
- $B^0 \rightarrow D^*\pi$: sum of a single-sided Crystal Ball function [67] and a Gaussian function ($PDF_{B^0 \rightarrow D^*\pi}^\pi$).
- Combinatorial: sum of an exponential and a constant function.

For the kaon sample they are:

- $B^0 \rightarrow D\pi$: double-sided Hypatia function ($PDF_{B^0 \rightarrow D\pi}^K$).
- $B^0 \rightarrow DK$: single-sided Hypatia function ($PDF_{B^0 \rightarrow DK}^K$).
- $B^0 \rightarrow D\rho$: double Gaussian function ($PDF_{B^0 \rightarrow D\rho}^K$).
- $B^0 \rightarrow DK^*$: Gaussian function ($PDF_{B^0 \rightarrow DK^*}^K$).
- Combinatorial: sum of an exponential and a constant function.

The definitions of all the PDFs listed above are reported in Appendix 4.13. The fits to the MC samples are shown in figs. 4.15 and 4.16.

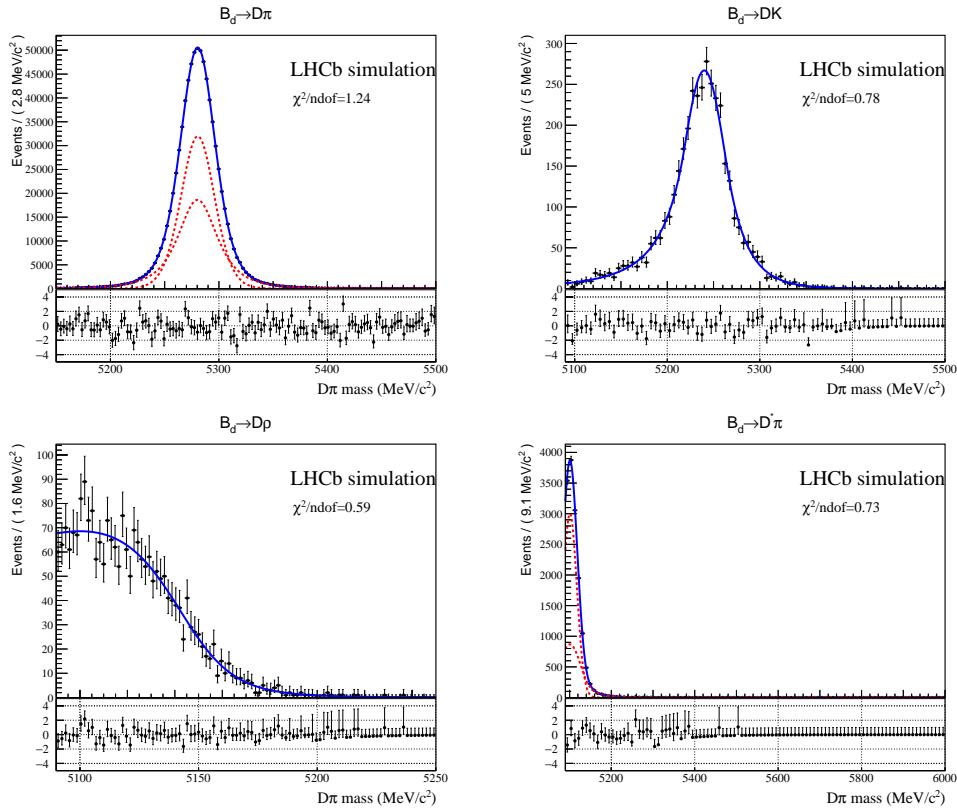


Figure 4.15 – Fits to MC samples of PDFs describing the π sample composition.

864 4.4.2 Fit to data

In order to perform Fit A, two extended PDFs are defined as follows:

$$\begin{aligned} PDF_\pi = & N_{B^0 \rightarrow D\pi}^\pi PDF_{B^0 \rightarrow D\pi}^\pi + N_{B^0 \rightarrow DK}^\pi PDF_{B^0 \rightarrow DK}^\pi \\ & + N_{B^0 \rightarrow D^*\pi}^\pi PDF_{B^0 \rightarrow D^*\pi}^\pi + N_{B^0 \rightarrow D\rho}^\pi PDF_{B^0 \rightarrow D\rho}^\pi \\ & + N_{\text{comb}}^\pi PDF_{\text{comb}}^\pi. \end{aligned} \quad (4.9)$$

$$\begin{aligned} PDF_K = & N_{B^0 \rightarrow DK}^K PDF_{B^0 \rightarrow DK}^K + N_{B^0 \rightarrow D\pi}^K PDF_{B^0 \rightarrow D\pi}^K \\ & + N_{B^0 \rightarrow D^*K}^K PDF_{B^0 \rightarrow D^*K}^K + N_{B^0 \rightarrow D\rho}^K PDF_{B^0 \rightarrow D\rho}^K \\ & + N_{\text{comb}}^K PDF_{\text{comb}}^K. \end{aligned} \quad (4.10)$$

865 Two likelihood functions are defined using data and PDFs related to both sample, \mathcal{L}_π
866 and \mathcal{L}_K , and the product $\mathcal{L}_\pi \mathcal{L}_K$ is maximised during the fit.

867 The following strategy is adopted to perform Fit A:

- 868 • The mean and width parameters ($\mu_{B^0 \rightarrow D\pi}^\pi, \sigma I_{B^0 \rightarrow D\pi}^\pi, \sigma J_{B^0 \rightarrow D\pi}^\pi, \mu_{B^0 \rightarrow DK}^K, \sigma_{B^0 \rightarrow DK}^K$)

4.4. Fits to the B^0 invariant mass

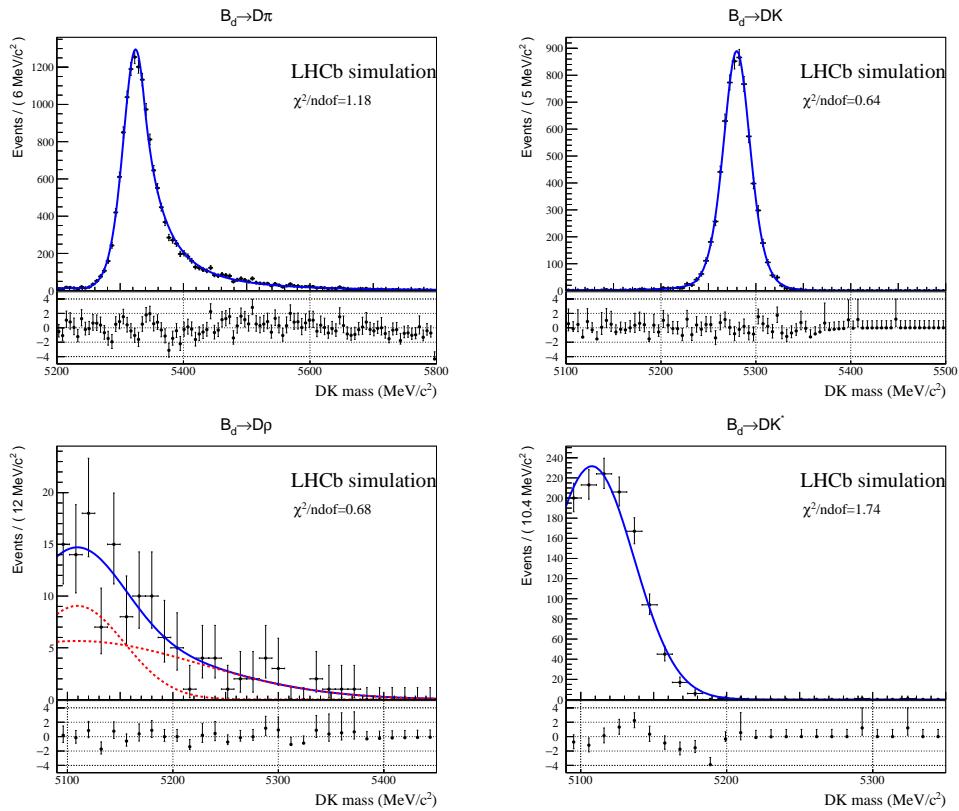


Figure 4.16 – Fits to MC samples of PDFs describing the K sample composition.

869 of $PDF_{B^0 \rightarrow D\pi}^\pi$ and $PDF_{B^0 \rightarrow DK}^K$ are floated in the fit.

- 870 • The tail parameters ($a1_{B^0 \rightarrow D\pi}^\pi, a2_{B^0 \rightarrow D\pi}^\pi, n1_{B^0 \rightarrow D\pi}^\pi, n2_{B^0 \rightarrow D\pi}^\pi$) of $PDF_{B^0 \rightarrow D\pi}^\pi$ are
871 constrained in the following way: $a1_{B^0 \rightarrow D\pi}^\pi, a2_{B^0 \rightarrow D\pi}^\pi$ are set to the values found
872 on MC and both multiplied by a floating scale factor $sa_{B^0 \rightarrow D\pi}^\pi$; the same constraint
873 is applied to $n1_{B^0 \rightarrow D\pi}^\pi, n2_{B^0 \rightarrow D\pi}^\pi$, where the scale factor is labelled as $sn_{B^0 \rightarrow D\pi}^\pi$
- 874 • The yield parameters $N_{B^0 \rightarrow D\pi}^K$ and $N_{B^0 \rightarrow DK}^\pi$ are constrained according to eqs. (4.7)
875 and (4.8). The efficiencies $\epsilon_{PID}(B^0 \rightarrow D\pi)_{D\pi}$ and $\epsilon_{PID}(B^0 \rightarrow DK)_{DK}$ are Gaussian-
876 constrained independently in the fit, using the values reported in Tab. 4.10 as
877 constraint parameters. The yield $N_{B^0 \rightarrow D\rho}^K$ is fixed to be 0.92 times the yield
878 $N_{B^0 \rightarrow DK^*}^K$, the latter being floated in the fit. This is done according to the expected
879 $B^0 \rightarrow D\rho$ to $B^0 \rightarrow DK^*$ ratio in the kaon sample, which is 0.92 ± 0.21 . All the
880 other yields appearing in Eqs. 4.9 and 4.10 are floated in the fit.
- 881 • The mean parameters ($\mu_{B^0 \rightarrow D^*\pi}^\pi, \mu_{B^0 \rightarrow D\rho}^{\pi/K}$) of $PDF_{B^0 \rightarrow D^*\pi}^\pi, PDF_{B^0 \rightarrow D\rho}^{\pi/K}$, are con-
882 strained to be shifted from $\mu_{B^0 \rightarrow D\pi}^\pi$ (in the π sample) and $\mu_{B^0 \rightarrow DK}^K$ (in the K
883 sample) by the same amount found in MC ². The mean parameters ($\mu_{B^0 \rightarrow DK^*}^K$,

²The shift of the component with respect to the $B^0 \rightarrow D\pi$ peak in the π sample/ $B^0 \rightarrow DK$ peak in the K sample is denoted as $\Delta\mu_{comp}^{K/\pi}$.

Chapter 4. Measurement of CP violation in $B^0 \rightarrow D^\mp \pi^\pm$ decays

884 $\mu_{B^0 \rightarrow DK}^\pi, \mu_{B^0 \rightarrow D\pi}^K$) of $PDF_{B^0 \rightarrow DK*}^K, PDF_{B^0 \rightarrow DK}^\pi, PDF_{B^0 \rightarrow D\pi}^K$ are floated in the
885 fit instead.

- 886 • The exponent parameters ($c_{\text{comb}}^{\pi/K}$) and fractions ($f_{\text{comb}}^{\pi/K}$) of $PDF_{\text{comb}}^{\pi/K}$ are floated in
887 the fit.

888 The projections of the fitted PDF_π and PDF_K PDFs in the $D\pi$ and DK invariant mass
889 observables (Fit A) are shown in fig. 4.17, together with the π and K data samples. A
890 list of all the parameters fixed in the Fit A is given in Tab. ???. The fitted parameters
(including yields and PID efficiencies) are listed in Tab. 4.13.

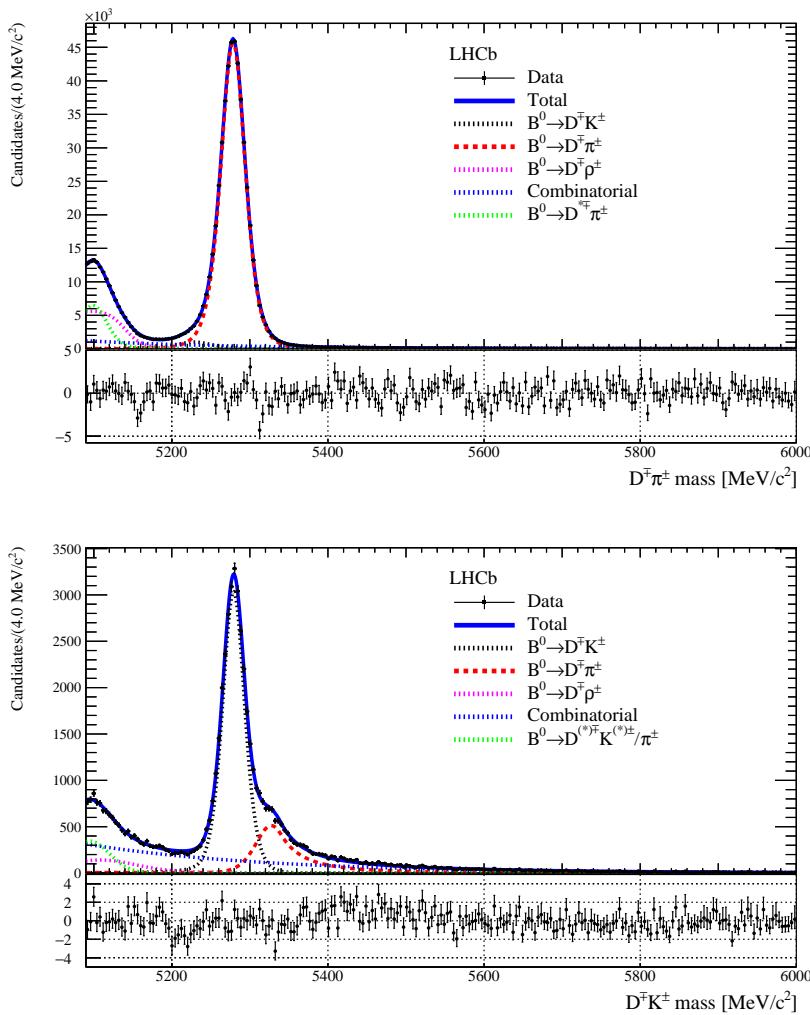


Figure 4.17 – Projections of the PDFs describing the π and K sample compositions fitted to the π and K data samples (Fit A).

891

4.4. Fits to the B^0 invariant mass

Table 4.11 – Parameters of PDF_π fixed in Fit A. These values are obtained in the fits to MC samples described in Sec. 4.4.1.

Parameter name	Value	Comment
$a1_{B^0 \rightarrow D\pi}^\pi$	0.722 ± 0.091	fitted in MC and constrained in data
$a2_{B^0 \rightarrow D\pi}^\pi$	0.96 ± 0.12	fitted in MC and constrained in data
$n1_{B^0 \rightarrow D\pi}^\pi$	5.92 ± 0.92	fitted in MC and constrained in data
$n2_{B^0 \rightarrow D\pi}^\pi$	5.83 ± 0.38	fitted in MC and constrained in data
$\beta_{B^0 \rightarrow D\pi}^\pi$	0.0	fixed
$\lambda_{B^0 \rightarrow D\pi}^\pi$	-1.240 ± 0.060	fitted in MC
$\zeta_{B^0 \rightarrow D\pi}^\pi$	0.0	fixed
$f_{B^0 \rightarrow D\pi}^\pi$	0.436 ± 0.060	fixed
$\sigma_{B^0 \rightarrow DK}^\pi$	$23.43 \pm 0.42 \text{ MeV}/c^2$	fitted in MC
$a1_{B^0 \rightarrow DK}^\pi$	0.898 ± 0.025	fitted in MC
$a2_{B^0 \rightarrow DK}^\pi$	1.092 ± 0.033	fitted in MC
$n1_{B^0 \rightarrow DK}^\pi$	3.83 ± 0.40	fitted in MC
$n2_{B^0 \rightarrow DK}^\pi$	22.0 ± 7.6	fitted in MC
$\beta_{B^0 \rightarrow DK}^\pi$	0.0	fixed
$\lambda_{B^0 \rightarrow DK}^\pi$	-24 ± 10	fitted in MC
$\zeta_{B^0 \rightarrow DK}^\pi$	0.0	fixed
$\nu_{B^0 \rightarrow D\rho}^\pi$	-2.01 ± 0.15	fitted in MC
$\mu_{B^0 \rightarrow D\rho}^\pi$	$4828 \pm 80 \text{ MeV}/c^2$	fitted in MC and constrained into $\Delta\mu_{B^0 \rightarrow D\rho}^\pi$
$\sigma_{B^0 \rightarrow D\rho}^\pi$	$550 \pm 190 \text{ MeV}/c^2$	fitted in MC
$\tau_{B^0 \rightarrow D\rho}^\pi$	1.163 ± 0.090	fitted in MC
$\alpha_{B^0 \rightarrow D^*\pi}^\pi$	-1.443 ± 0.031	fitted in MC
$n_{B^0 \rightarrow D^*\pi}^\pi$	4.65 ± 0.30	fitted in MC
$\mu_{B^0 \rightarrow D^*\pi}^\pi$	$5100.93 \pm 0.23 \text{ MeV}/c^2$	fitted in MC and constrained into $\Delta\mu_{B^0 \rightarrow D^*\pi}^\pi$
$\sigma G_{B^0 \rightarrow D^*\pi}^\pi$	$16.52 \pm 0.20 \text{ MeV}/c^2$	fitted in MC
$\sigma C B_{B^0 \rightarrow D^*\pi}^\pi$	$25.84 \pm 0.48 \text{ MeV}/c^2$	fitted in MC
$f_{B^0 \rightarrow D^*\pi}^\pi$	0.302 ± 0.011	fitted in MC

Table 4.12 – Fixed parameters of PDF_K in Fit A. These values are obtained from the fits to MC samples described in Sec. 4.4.1.

Parameter name	Value	Comment
$\sigma_{B^0 \rightarrow D\pi}^K$	$23.97 \pm 0.46 \text{ MeV/c}^2$	fitted in MC
$a1_{B^0 \rightarrow D\pi}^K$	3.14 ± 0.14	fitted in MC
$a2_{B^0 \rightarrow D\pi}^K$	0.569 ± 0.039	fitted in MC
$n1_{B^0 \rightarrow D\pi}^K$	0.05 ± 0.11	fitted in MC
$n2_{B^0 \rightarrow D\pi}^K$	2.81 ± 0.12	fitted in MC
$\beta_{B^0 \rightarrow D\pi}^K$	0.0	fixed
$\lambda_{B^0 \rightarrow D\pi}^K$	-3.77 ± 0.57	fitted in MC
$\zeta_{B^0 \rightarrow D\pi}^K$	0.0	fixed
$\sigma_{B^0 \rightarrow DK}^K$	$17.32 \pm 0.26 \text{ MeV/c}^2$	fitted in MC
$a_{B^0 \rightarrow DK}^K$	2.34 ± 0.19	fitted in MC
$n_{B^0 \rightarrow DK}^K$	1.56 ± 0.33	fitted in MC
$\beta_{B^0 \rightarrow DK}^K$	0.0	fixed
$\lambda_{B^0 \rightarrow DK}^K$	-3.45 ± 0.34	fitted in MC
$\zeta_{B^0 \rightarrow DK}^K$	0.0	fixed
$f_{B^0 \rightarrow D\rho}^K$	0.58 ± 0.17	fitted in MC
$\mu_{B^0 \rightarrow D\rho}^K$	$5109 \pm 24 \text{ MeV/c}^2$	fitted in MC and constrained into $\Delta\mu_{B^0 \rightarrow D\rho}^K$
$\sigma_{1_{B^0 \rightarrow D\rho}^K}$	$117 \pm 18 \text{ MeV/c}^2$	fitted in MC
$\sigma_{2_{B^0 \rightarrow D\rho}^K}$	$45 \pm 16 \text{ MeV/c}^2$	fitted in MC

892 4.4.3 $sWeight$ calculation

After that Fit A is performed, all the floating shape parameters in PDF_π are fixed, all the background components in the π sample are combined into a single PDF, and the B mass range is restricted to $[5220, 5600] \text{ MeV/c}^2$. Concretely, PDF_π is redefined as follows:

$$PDF_\pi = N_{B^0 \rightarrow D\pi}^\pi PDF_{B^0 \rightarrow D\pi}^\pi + N_{\text{bkg}}^\pi PDF_{\text{bkg}}^\pi. \quad (4.11)$$

The N_{bkg}^π coefficient describes the total number of background events in the new range. The PDF_{bkg}^π term is defined as:

$$\begin{aligned} PDF_{\text{bkg}}^\pi &= f_{\text{comb}}^\pi PDF_{\text{comb}}^\pi \\ &+ f_{B^0 \rightarrow DK}^\pi PDF_{B^0 \rightarrow DK}^\pi + f_{B^0 \rightarrow D\rho}^\pi PDF_{B^0 \rightarrow D\rho}^\pi \\ &+ (1 - f_{\text{comb}}^\pi - f_{B^0 \rightarrow DK}^\pi - f_{B^0 \rightarrow D\rho}^\pi) PDF_{B^0 \rightarrow D^*\pi}^\pi. \end{aligned} \quad (4.12)$$

Table 4.13 – Floating parameters of PDF_π and PDF_K in Fit A.

Parameter number	Parameter name	Fitted value	MC value
0	$\mu_{B^0 \rightarrow DK}^\pi$	5228.62 ± 0.92	5239.96 ± 0.52
1	$\sigma_{B^0 \rightarrow DK}^K$	17.17 ± 0.15	17.32 ± 0.26
2	$\mu_{B^0 \rightarrow DK^*}^K$	5094.8 ± 3.9	5107.7 ± 2.2
3	$\sigma_{B^0 \rightarrow DK^*}^K$	25.5 ± 2.6	28.9 ± 1.2
4	$c1_{\text{comb}}^\pi$	-0.00576 ± 0.00017	/
5	$c2_{\text{comb}}^\pi$	-0.0010 ± 0.0010	/
6	f_{comb}^π	0.899 ± 0.025	/
7	c_{comb}^K	-0.004397 ± 0.000066	/
8	$\mu_{B^0 \rightarrow DK}^K$	5279.19 ± 0.14	5279.84 ± 0.20
9	$\mu_{B^0 \rightarrow D\pi}^\pi$	5278.360 ± 0.032	5280.291 ± 0.025
10	$sa_{B^0 \rightarrow D\pi}^\pi$	0.684 ± 0.022	1.0
11	$sn_{B^0 \rightarrow D\pi}^\pi$	2.71 ± 0.80	1.0
12	$\sigma I_{B^0 \rightarrow D\pi}^\pi$	37.69 ± 0.69	43.9 ± 4.9
13	$\sigma J_{B^0 \rightarrow D\pi}^\pi$	17.01 ± 0.17	16.82 ± 0.15
14	$\mu_{B^0 \rightarrow D\pi}^K$	5327.32 ± 0.78	5324.37 ± 0.44
15	$\epsilon_{\text{PID}}(B^0 \rightarrow DK)_K$	0.6197 ± 0.0079	0.637 ± 0.013
16	$\epsilon_{\text{PID}}(B^0 \rightarrow D\pi)_\pi$	0.98048 ± 0.00041	0.979 ± 0.004
17	$N_{B^0 \rightarrow DK}^K$	28820 ± 242	/
18	$N_{B^0 \rightarrow DK^*}^K$	3164 ± 110	/
19	$N_{B^0 \rightarrow D\rho}^\pi$	73766 ± 1239	/
20	$N_{B^0 \rightarrow D^*\pi}^\pi$	52494 ± 819	/
21	N_{comb}^K	17469 ± 341	/
22	N_{comb}^π	56230 ± 1336	/
23	$N_{B^0 \rightarrow D\pi}^\pi$	483398 ± 1040	/

For each background component in the π sample, the fraction f_{comp}^π is determined by the following expression:

$$f_{\text{comp}}^\pi = \frac{N_{\text{comp}}^\pi \int_{5220\text{MeV}/c^2}^{5600\text{MeV}/c^2} PDF_{\text{comp}}^\pi dm}{\sum_i N_i^\pi}, \quad (4.13)$$

where the index i in the denominator runs over all the background components in the π sample (combinatorial, $B^0 \rightarrow DK$, $B^0 \rightarrow D\rho$, $B^0 \rightarrow D^*\pi$).

As cross-check, the fitted yields of $B^0 \rightarrow DK$, $B^0 \rightarrow D\rho$ and $B^0 \rightarrow D^*\pi$ in the pion sample are compared with the expected yields from Table 4.9, which are obtained, for each background, by multiplying the fitted $B^0 \rightarrow D\pi$ yield in the pion sample by the f_{bkg} fractions. These yields are reported in Table 4.14.

Once PDF_π is redefined as described above, an unbinned extended maximum likelihood fit (Fit B) is performed to the π sample. The only floating parameters are the yields

Table 4.14 – Expected and fitted yields for the physical background components in the pion sample. The uncertainties on the expected yields are propagated from 4.9.

Decay	Expected yield [10^4]	Fitted yield [10^4]
$B^0 \rightarrow DK$	1.26 ± 0.15	1.65 ± 0.05
$B^0 \rightarrow D\rho$	8.1 ± 1.4	7.38 ± 0.12
$B^0 \rightarrow D^*\pi$	5.6 ± 0.4	5.25 ± 0.08

Table 4.15 – Floating parameters of PDF_π in Fit B.

Parameter	Value
$N_{B^0 \rightarrow D\pi}^\pi$	479045 ± 732
N_{bkg}^π	34381 ± 300

901 $N_{B^0 \rightarrow D\pi}^\pi$ and N_{bkg}^π . The result of the fit is reported in Table 4.15.

902 Fit B is used as starting point to apply the *sPlot* technique and extract *sWeights* used
 903 to subtract the total background component from the π sample. The projection of the
 904 fitted PDF_π in Fit B, and a comparison between the weighted and unweighted datasets
 905 projected over the B decay time and D invariant mass observables, are reported in
 906 Fig 4.18. The distribution of *sWeights* is shown in fig. 4.19.

907 4.4.4 Fit split by subsample

908 In order to validate the data sample, selection and fit procedure, the mass fit is repeated
 909 in smaller subsamples. These subsamples are divided per year of data taking (2011, 2012),
 910 magnet polarity (up, down) and final state ($D^+\pi^-$, $D^-\pi^+$).

911 In order to cope with the reduced statistics in the 2011 subsample and simplify the “Fit
 912 A” strategy, the PDF_{comb}^K PDF is taken as simple exponential (instead of an exponential
 913 plus constant function).

914 The projections of the PDFs describing the pion and kaon samples for each data subsample
 915 (“Fit A”) are illustrated in Figs. 4.20 and 4.21 respectively.

916 The “Fit B” strategy is repeated exactly as before for each subsample. The corresponding
 917 signal and background fitted yields are listed in Table 4.16. The sum of the yields for each
 918 subsample has to be compared with the total signal yield in the total sample (reported in
 919 Table 4.15), which is $(4.790 \pm 0.007)10^5$.

4.4. Fits to the B^0 invariant mass

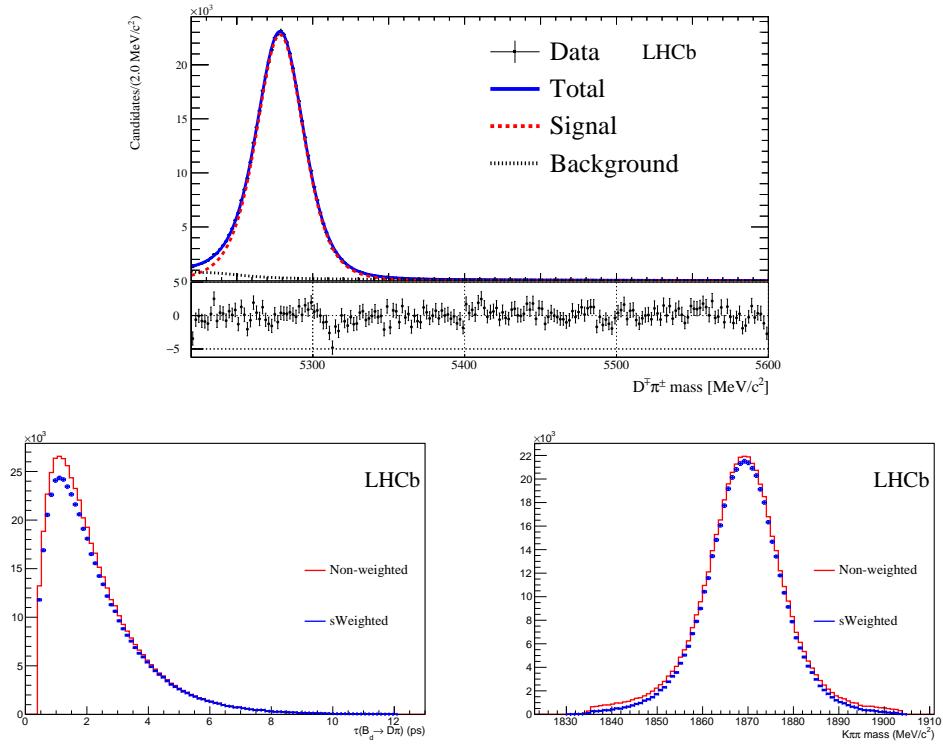


Figure 4.18 – *s*Weighted and unweighted π sample projected over B decay time and D invariant mass observables.

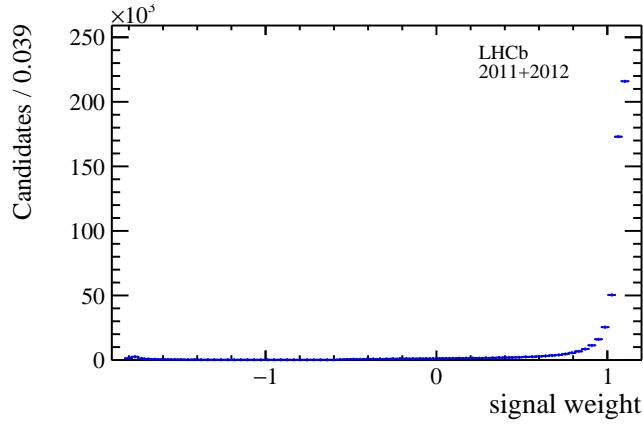


Figure 4.19 – Distribution of calculated *sWeights* by Fit B. No obvious outliers are visible.

Chapter 4. Measurement of CP violation in $B^0 \rightarrow D^\mp \pi^\pm$ decays

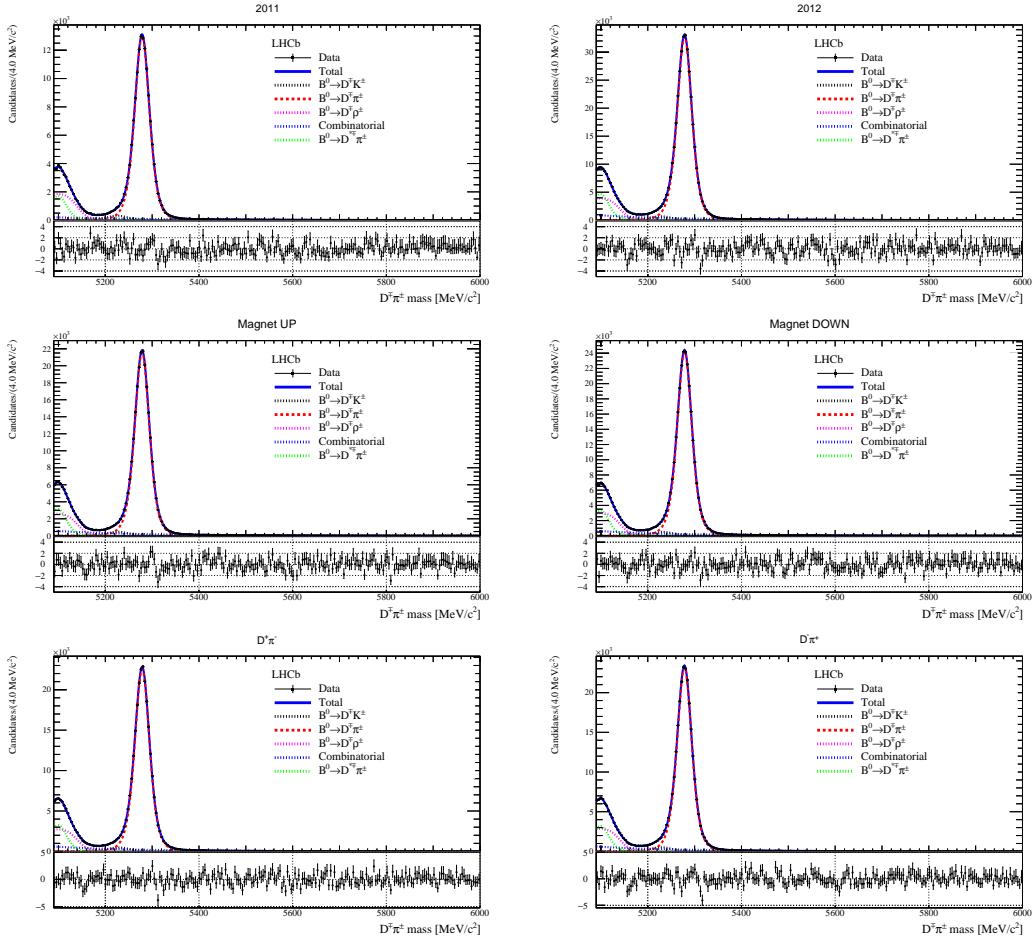


Figure 4.20 – Projections of the PDFs describing the pion sample for each data subsample.

Table 4.16 – Signal yields (in units of 10^5) in the pion sample for each subsample (obtained from “Fit B”).

2011	2012	Sum
1.383 ± 0.004	3.424 ± 0.006	4.807 ± 0.007
Magnet Up	Magnet Down	Sum
2.263 ± 0.005	2.523 ± 0.005	4.786 ± 0.007
$D^- \pi^+$	$D^+ \pi^-$	Sum
2.421 ± 0.005	2.373 ± 0.005	4.794 ± 0.007

4.4. Fits to the B^0 invariant mass

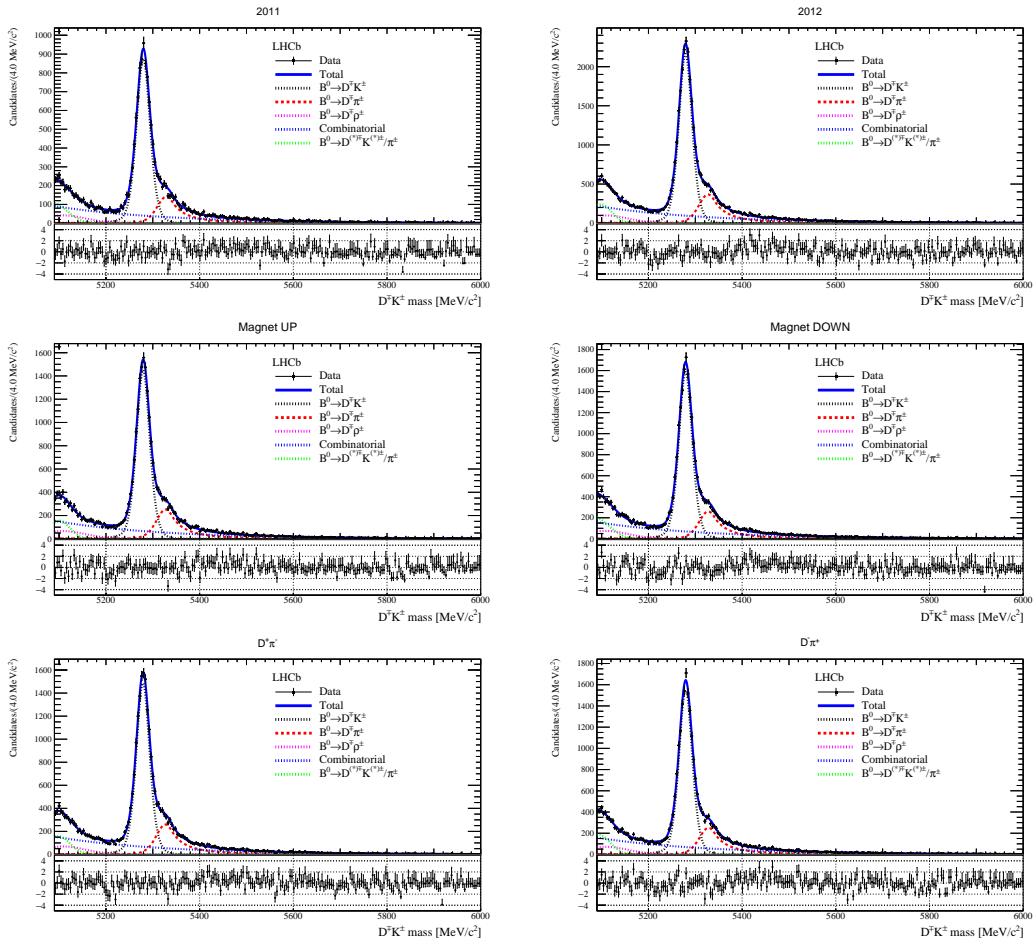


Figure 4.21 – Projections of the PDFs describing the kaon sample for each data subsample.

920 4.5 Time Resolution

921 The decay time resolution is determined from a sample of “fake” B^0 candidates formed from
 922 a prompt D^\pm candidate combined with a track originating from the PV. This is sample
 923 is referred to as “ D^\pm +track”. The candidates are selected from the B02DKLTUBD2HHH
 924 stripping line. These candidates are subjected to the same offline selection as that of the
 925 signal sample without a BDT cut, with two additional requirements: the D^\pm IP χ^2 with
 926 respect to the PV is less than 9 to reduce the nonresonant $K^\pm \pi^\mp \pi^\mp$ background, and the
 927 number of PVs per event are required to be 1 in order to reduce wrong-PV associations.
 928 The combined stripping and offline selection yields 51053 candidates. True D^\pm +track
 929 candidates are unfolded from combinatorial background and nonresonant decays by means
 930 of *sWeights* computed via a fit to the $K^\pm \pi^\mp \pi^\mp$ invariant mass distribution.

931 4.5.1 Companion track momentum reweighting

932 The decay time resolution is found to be dependent upon the companion track p_T which
 933 is considerably lower on average for the D^\pm +track candidates than it is for genuine
 934 $B^0 \rightarrow D^\pm \pi^\mp$ signal. This is corrected for by reweighting the prompt sample by the ratio of
 935 prompt and signal $B^0 \rightarrow D^\pm \pi^\mp$ companion $\log(p_T)$ distributions. The logarithm is taken
 936 to compress the high- p_T tails, and a binning scheme is chosen to have an equal number of
 937 signal events per bin entering the denominator. The p_T spectra for signal, prompt, and
 938 reweighted prompt candidates is shown in Fig. 4.22. Prior to reweighting the average
 939 resolution is determined to be ~ 71 fs, post-reweighting the resolution is found to be
 consistent with the value of 54 fs that was obtained in other LHCb analyses [68].

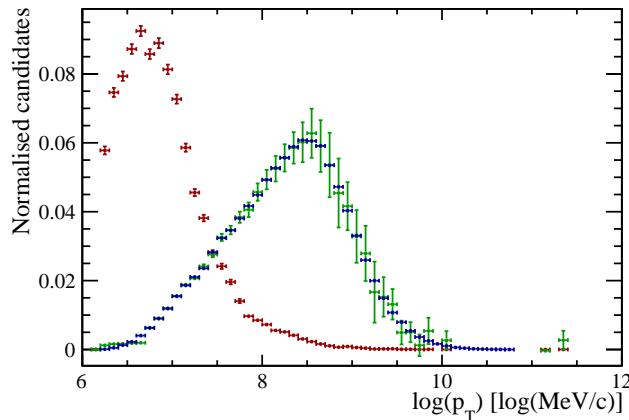


Figure 4.22 – Normalised $\log(p_T)$ distributions of signal $B^0 \rightarrow D^\pm \pi^\mp$ (blue), prompt D^\pm +track before reweighting (red) and after reweighting (green).

941 4.5.2 Resolution determination from decay time error parameterisa-
 942 tion

In order to study potential second order corrections to the decay time error distribution, fits to the prompt decay time distribution in bins of the per-event decay time error are performed. The binning scheme is chosen such that the sum of signal *sWeights* in each bin is equal. The fit is similar to that used to determine the resolution in Ref. [69], consisting of three components: a delta function convolved with a Gaussian accounts for the genuine prompt $D^\pm +$ track component; a pair of exponential functions convolved with the same Gaussian as the delta function accounts for signal candidates coming from b-hadron decays, and a Gaussian with a large width accounts for wrong-PV associated backgrounds. The time constant of the exponentials and the mean of the wrong-PV component are fixed from a global fit to the sample, while the mean and width of the resolution, the width of the wrong PV component and the relative fractions of the prompt, wrong-PV and from-b components are all free parameters in the fits to each decay time error bin. A likelihood fit is performed in 20 bins of the decay time error from which the average decay time error $\langle \delta \rangle_i$ and the measured resolution $\langle \sigma \rangle_i$ are determined. The results of these fits are presented in Table 4.17, and a representative fit is shown in fig 4.23. A χ^2 fit to the obtained values of the per-bin average error and resolution of the form:

$$\langle \sigma \rangle_i = \langle \sigma \rangle + p_1 (\langle \delta \rangle_i - \langle \delta \rangle) + p_2 (\langle \delta \rangle_i - \langle \delta \rangle)^2 \quad (4.14)$$

943 where $\langle \delta \rangle$ is the average per-event decay time error of the whole (unbinned) sample. In
 944 the prompt $D^\pm +$ track sample this is determined to be 0.0307 ± 0.0097 ps, in good
 945 agreement with the signal sample value of 0.034 ± 0.011 ps. The fit determines the average
 946 resolution, $\langle \sigma \rangle$ in addition to the trend. This fit is shown in fig. 4.23, the result of which
 947 is presented in Table 4.18. The global average resolution is determined from this fit to
 948 be $\sigma(\langle \delta \rangle) = 0.05491 \pm 0.00038$ ps. The procedure is found to be stable and to reproduce
 949 similar values with fits to 10 bins (0.05523 ± 0.00041 ps) and 30 (0.05464 ± 0.00037 ps)
 950 yielding compatible results.

951 This method, which accounts for second-order corrections to the decay time error, is used to
 952 define the width of a single Gaussian in the decay time fit to data, $\mathcal{R}(t-t') = G(t-t', \langle \sigma \rangle)$,
 953 with $\langle \sigma \rangle = 0.05491 \pm 0.00038$ ps. The uncertainty stated here is statistical. Systematic
 954 uncertainties will be considered in section 4.8.2.

955 4.6 Time-dependent selection efficiency

956 Because of some of the selection criteria described in Sec. 4.2, the B^0 decay time
 957 distribution is biased from the shape it would have with a selection efficiency equal
 958 to 100%. This efficiency, called here and after “acceptance”, is a function of time. In
 959 particular, it goes very rapidly to zero at low decay times due to the impact parameter

Table 4.17 – Parameters obtained from a fit to the p_T corrected sPlot of the decay time distribution, t , in bins of per-event decay time error, δ , for prompt $D^\pm +$ track signal.

Bin (i)	lower edge	$\langle \delta \rangle_i$	$\langle \sigma \rangle_i$
0	0.01	0.0142 ± 0.0016	0.01731 ± 0.00053
1	0.0165376	0.01801 ± 0.00075	0.02439 ± 0.00089
2	0.0192247	0.02038 ± 0.00063	0.0286 ± 0.0011
3	0.0214493	0.02248 ± 0.00052	0.0347 ± 0.0011
4	0.0232264	0.02388 ± 0.00036	0.0384 ± 0.0013
5	0.0245968	0.02528 ± 0.00033	0.0422 ± 0.0014
6	0.0257605	0.02641 ± 0.00034	0.0449 ± 0.0014
7	0.0269093	0.02753 ± 0.00033	0.0489 ± 0.0015
8	0.0280345	0.02857 ± 0.00028	0.0489 ± 0.0015
9	0.0290414	0.02955 ± 0.00030	0.0525 ± 0.0018
10	0.0301189	0.03054 ± 0.00024	0.0552 ± 0.0019
11	0.0309259	0.03138 ± 0.00027	0.0582 ± 0.0017
12	0.0318409	0.03229 ± 0.00032	0.0594 ± 0.0016
13	0.0328907	0.03347 ± 0.00036	0.0641 ± 0.0015
14	0.0341106	0.03482 ± 0.00039	0.0643 ± 0.0014
15	0.0354999	0.03638 ± 0.00052	0.0658 ± 0.0014
16	0.0372226	0.03830 ± 0.00063	0.0719 ± 0.0012
17	0.0395386	0.04096 ± 0.00086	0.0736 ± 0.0012
18	0.0424521	0.0447 ± 0.0014	0.0786 ± 0.0011
19	0.0473915	0.0561 ± 0.0095	0.0933 ± 0.0010

Table 4.18 – Parameters obtained from a fit to the per-bin decay time error and measured resolution.

Parameter	Result			
$\langle \delta \rangle$	0.0307	\pm	0.0097	ps
p_1	2.031	\pm	0.022	
p_2	-19.30	\pm	1.6	
$\sigma(\langle \delta \rangle)$	0.05491	\pm	0.00038	ps

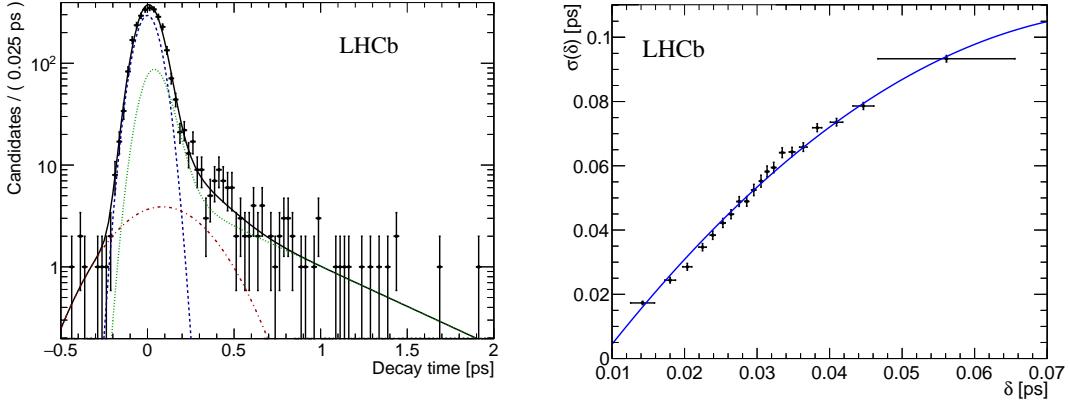


Figure 4.23 – Left: p_T -corrected and background-subtracted decay-time distribution for the 15th bins in per-event decay time error. A fit is overlaid in black. The wrong-PV component is shown in red, dot-dashed. The from-b component is shown in green, dotted, and the prompt component is shown in blue, dashed. The results are presented in Table 4.17. Right: measured resolution vs. average per-event decay time error determined from fits to the decay time in bins of decay time error. The x -axis uncertainties are the standard deviation of the average per-event decay time error in each bin. The overlaid fit is described in the text.

960 requirements which exclude short-lived B^0 candidates; then, it reaches a “plateau” at
 961 intermediate decay times; finally, it drops at high decay times because of reconstruction
 962 inefficiencies effects from VELO.

The acceptance function $a(t)$ is parameterised using splines defined analytically as described in Ref. [70]. These splines are cubic polynomials defined in sub-ranges of the decay time. The boundaries of each sub-range, called “knots”, are located at $[0.4, 0.5, 1.0, 1.5, 2.0, 2.3, 2.6, 3.0, 4.0, 10.0, 12.0]$ ps. The location of the knots and the higher density of knots at low decay times, where the acceptance is a strongly-varying function of t , ensure that the resulting acceptance is sufficiently smooth. For each knot t_i , a coefficient v_i is defined, which is the actual value of the acceptance function $a(t_i)$. In order to fix the overall scale of the acceptance function, the v_{10} coefficient is set to 1.0. Moreover, since statistical fluctuations at high decay times may strongly affect v_{11} , the latter is constrained to be the linear extrapolation from the previous two coefficients:

$$v_{11} = v_{10} + \frac{v_9 - v_{10}}{t_9 - t_{10}} \times (t_{11} - t_{10}) \quad (4.15)$$

The knots position and the number of knots are optimized in order to fit the $B^0 \rightarrow D^- \pi^+$ Monte Carlo decay time distribution with sufficient fit quality. The PDF adopted in this

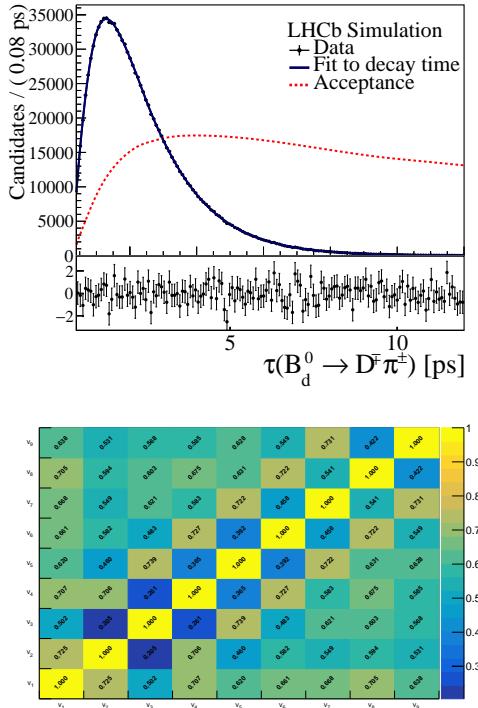


Figure 4.24 – (Top) Projection of the fitted acceptance function over the Monte Carlo sample and (bottom) correlation matrix.

fit is the following:

$$PDF(t) \propto a(t) \int dt' \mathcal{R}(t-t') e^{-t'/\tau_d}, \quad (4.16)$$

where $\mathcal{R}(t-t')$ is the average resolution model discussed in 4.5 and τ_d is the B^0 lifetime value used in the Monte Carlo generation. All acceptance coefficients are floating in the fit, while resolution and lifetime are fixed.

The fit projection is shown in Figure 4.24 together with the correlation matrix obtained from the fit, whereas the fitted coefficients are listed in Table 4.19.

4.7 Decay time fit to data

The CP coefficients S_f and $S_{\bar{f}}$ are determined using an unbinned maximum likelihood fit where each candidate is weighted with the *sWeights* extracted from the mass fit described in Sec. 4.4. Hence, the total PDF is given solely by the PDF describing the signal distribution. This is

$$a(t) \int dt' \mathcal{R}(t-t') P(t'|d_{OS}, d_{SS}, \eta_{OS}, \eta_{SS}) P(\eta_{OS}) P(\eta_{SS}), \quad (4.17)$$

Table 4.19 – Acceptance parameters fitted on the Monte Carlo sample.

Parameter name	Fitted value
v_1	0.1961 ± 0.0016
v_2	0.3348 ± 0.0032
v_3	0.6159 ± 0.0057
v_4	0.8667 ± 0.0073
v_5	0.9982 ± 0.0086
v_6	1.0747 ± 0.0091
v_7	1.1051 ± 0.0094
v_8	1.1590 ± 0.0086
v_9	1.188 ± 0.014

where $R(t - t')$ is the Gaussian resolution function, and $P(\eta_i)$ ($i = OS, SS$) is the PDFs of the predicted mistag probability of the taggers. The term $P(t|d_{OS}, d_{SS}, \eta_{OS}, \eta_{SS})$ represents the expected B decay-time distribution for a B^0 or a \bar{B}^0 decaying into a $D^-\pi^+$ or $D^+\pi^-$ final state. This is conditional on the tagging decision d_i and the mistag probability η_i , and it contains the decay-rates of Eqs. ??–??. A detailed description of the time PDF including the tagging parameters, and the detection and production asymmetries, is given in Appendix 4.14. The likelihood function is obtained by the sum of the logarithm of the PDF of Eq. 4.17 evaluated at the the decay time, tagging decisions, and mistag probabilities of each candidate. Each term of this sum is multiplied by the *sWeights* of the candidate, s_W^i , to statistically subtract the background. The likelihood function is then scaled by a correction factor,

$$s = \frac{(\sum_i s_W^i)^2}{\sum_i (s_W^i)^2}, \quad (4.18)$$

969 to take into account the dilution due to the background subtraction with the *sWeights*,
 970 so that correctly-estimated uncertainties from the fit are obtained [71].

971 In the PDF, $\Delta\Gamma$ is fixed to zero. The CP coefficients D_f and $D_{\bar{f}}$, which are directly
 972 proportional to $\Delta\Gamma$, are also fixed to zero; moreover, the C_f ($C_{\bar{f}}$) coefficient is fixed to 1
 973 (-1) because the value of $r_{D\pi}^2$ is such that sensitivity to C_f ($C_{\bar{f}}$) is negligible. Systematic
 974 uncertainties will be considered in Sec.?? for all these assumptions.

The following physics parameters are Gaussian-constrained to their measured values:

$$\tau = 1/\Gamma = (1.518 \pm 0.004) \text{ ps} \quad (4.19)$$

$$\Delta m = (0.5050 \pm 0.0023) \text{ ps}^{-1} \quad (4.20)$$

975 where τ is taken as the world average value [72], and Δm is the LHCb measurement from

976 semileptonic B^0 decays [73].

977 The free parameters of the fit are:

- 978 • the S_f and $S_{\bar{f}}$ coefficients;
- 979 • the detection asymmetry A_D ³ and the production asymmetry A_P ;⁴
- 980 • the calibration parameters for both OS and SS taggers;
- 981 • the time acceptance coefficients.

982 The tagging efficiency asymmetries A_{eff}^i parameters are found to be consistent to zero in
 983 the $B^0 \rightarrow D^- \pi^+$ Monte Carlo sample: for this reason, these coefficients are fixed to zero
 984 in the fit.

985 The value of the parameters obtained from the fit to data are listed in Table 4.20. The
 986 correlation matrix of the parameters is reported in Appendix 4.15. The projection of
 987 the PDF on the decay-time distribution is shown in Fig 4.25, while Fig 4.26 shows the
 988 CP asymmetries for the two final states.

It is possible to define the asymmetries between Cabibbo-favoured (CF) and Cabibbo-suppressed (CS) amplitudes as follows:

$$A_{\text{CF}} = \frac{\Gamma_{B^0 \rightarrow f}(t) - \Gamma_{\bar{B}^0 \rightarrow \bar{f}}(t)}{\Gamma_{B^0 \rightarrow f}(t) + \Gamma_{\bar{B}^0 \rightarrow \bar{f}}(t)}, \quad (4.21)$$

$$A_{\text{CS}} = \frac{\Gamma_{\bar{B}^0 \rightarrow f}(t) - \Gamma_{B^0 \rightarrow \bar{f}}(t)}{\Gamma_{\bar{B}^0 \rightarrow f}(t) + \Gamma_{B^0 \rightarrow \bar{f}}(t)}.. \quad (4.22)$$

989 The asymmetries defined above are plotted together with data in Fig. 4.27.

990 The projection of the PDF considering the decay-time distribution of the four independent
 991 decays rates, $B^0 \rightarrow D^- \pi^+$, $\bar{B}^0 \rightarrow D^- \pi^+$, $B^0 \rightarrow D^+ \pi^-$ and $\bar{B}^0 \rightarrow D^+ \pi^-$, for OS and SS
 992 tagged candidates, are reported in Fig. 4.28 and 4.29, respectively.

993 The 2D contour plot for the CP coefficients S_f and $S_{\bar{f}}$ and for the detection and production
 994 asymmetry in Fig. 4.30.

995 Using the flavour tagging calibrations from data, the tagging performances of the signal
 996 sample is computed. The average of the squared dilution is 9.53 % for the OS and 2.79 %
 997 for the SS. By combining the two taggers together, the total squared dilution is 6.55 %.

³The detection asymmetry is defined as $\frac{N(f) - N(\bar{f})}{N(f) + N(\bar{f})}$. This definition has the opposite sign compared to the one in Ref. [74].

⁴The production asymmetry is defined as $\frac{N(\bar{B}^0) - N(B^0)}{N(\bar{B}^0) + N(B^0)}$

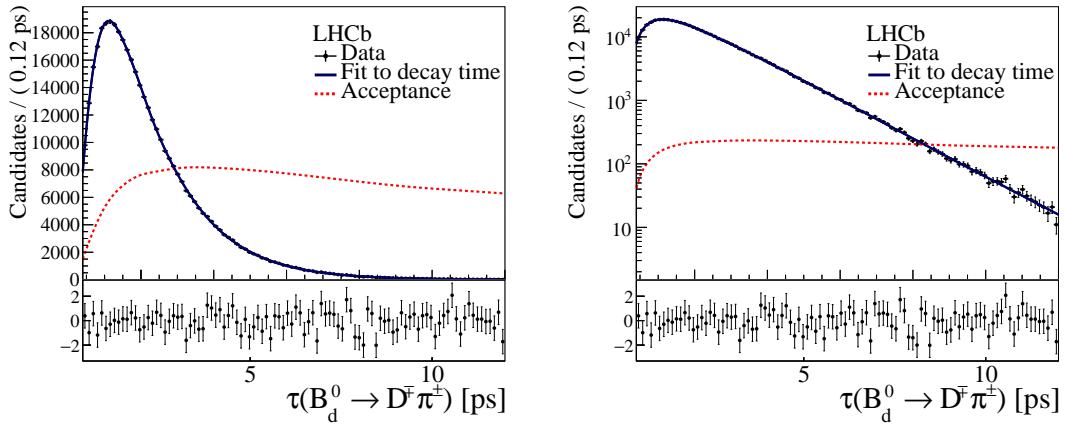


Figure 4.25 – Projections of time PDF onto the decay-time distribution of the *sWeighted* data sample (left: linear scale, right: logarithmic scale).

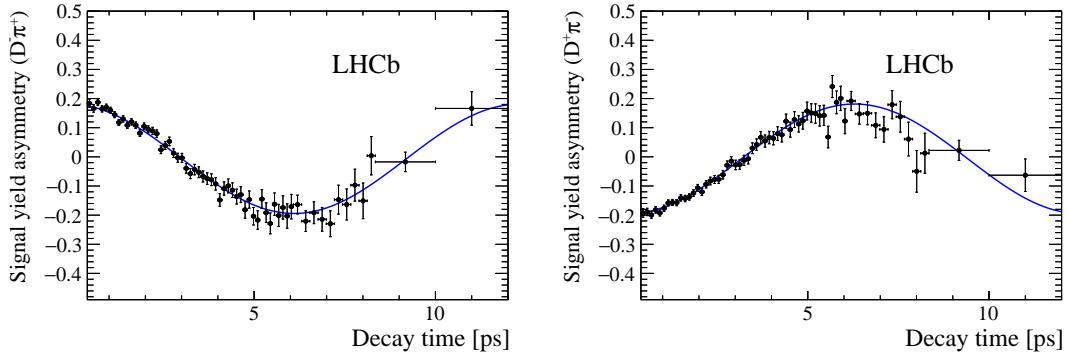


Figure 4.26 – Decay time dependent signal yield asymmetry for the (left) $D^- \pi^+$ and the (right) $D^+ \pi^-$ final states. The solid curve is the projection of the signal PDF.

998 Taking into account also untagged candidates, *i.e.* considering the tagging efficiency, the
 999 tagging power is 5.58 %.

1000 4.7.1 Fit validation

1001 Check of nuisance parameters

1002 The values of the nuisance parameters obtained in the fit (production/detection asymmetries,
 1003 flavour tagging calibrations) are compared with available external measurements.

The production asymmetry A_P is compared with the LHCb measurement of Ref. [?]. The production asymmetry is computed by weighting the production asymmetry measured from this paper in bins of p_T and η , $A_{P,i}$, with the signal fractions $\varepsilon_i = \frac{f_i}{\sum_i f_i}$ in each bin

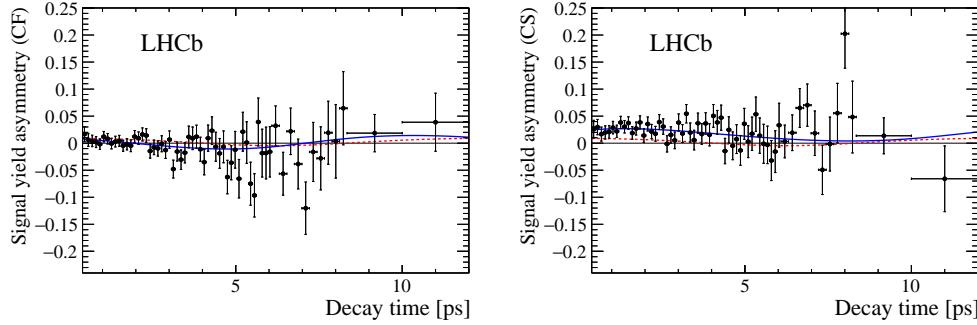


Figure 4.27 – Decay-time-dependent signal-yield asymmetries for (left) Cabibbo-favoured and (right) Cabibbo-suppressed decay topologies. The shown signal-yield asymmetries are defined in Eq. (4.21) and Eq. (4.22). The blue solid curve is the projection of the signal PDF, the red dotted curve indicates the projection of the fit when $S_{\bar{f}} \equiv -S_f$ is required.

i of $B^0 \rightarrow D^- \pi^+$ data:

$$A_P = \sum_i \varepsilon_i A_{P,i}, \quad (4.23)$$

where f_i is the number of signal candidates. This yields

$$A_P = -0.0100 \pm 0.0047 \text{ (stat)} \pm 0.0004 \text{ (syst)}. \quad (4.24)$$

1004 This value is compatible within 0.65σ with that obtain from the $B^0 \rightarrow D^- \pi^+$ decay-time
 1005 fit. The detection asymmetry A_D is again obtained from Ref. [?], and it's measured
 1006 using both $B^0 \rightarrow J/\psi K^{*0}$ to be 0.0098 ± 0.0046 and 0.0056 ± 0.0030 for 2011 and 2012,
 1007 respectively and $B_s^0 \rightarrow D_s^\pm \pi^\mp$ to be 0.0143 ± 0.0086 and 0.0103 ± 0.0058 for 2011 and
 1008 2012, respectively. The central value fitted on $B^0 \rightarrow D^- \pi^+$ are in agreement with this
 1009 set of results.

1010 The values of the parameters of the tagging calibrations are compared with those found
 1011 in the control samples as described in Sec. 3.2.1 and 3.2.2. The strategy presented in
 1012 section 3.2 is followed, and no perfect portability of the calibrations is assumed a priori,
 1013 as section 3.2.1 has shown this not to be the case. However, the values of the parameters
 1014 found in the signal fit are expected to be in the same ballpark interval of those from the
 1015 control channels. A full comparison that takes into account the correlation between the
 1016 parameters with a χ^2 test which uses the covariance matrices of the fit to data (reduced
 1017 to the tagging parameters only), and the fit of the calibration from the control channel,
 1018 is performed. The discrepancy (corresponding to the χ^2 minimum) is around 0.91σ for
 1019 the OS tagger, and 0.29σ for the SS tagger. The parameters which present the largest
 1020 disagreement are Δp_3^{OS} for the OS tagger and the Δp_0^{SS} for the SS tagger.

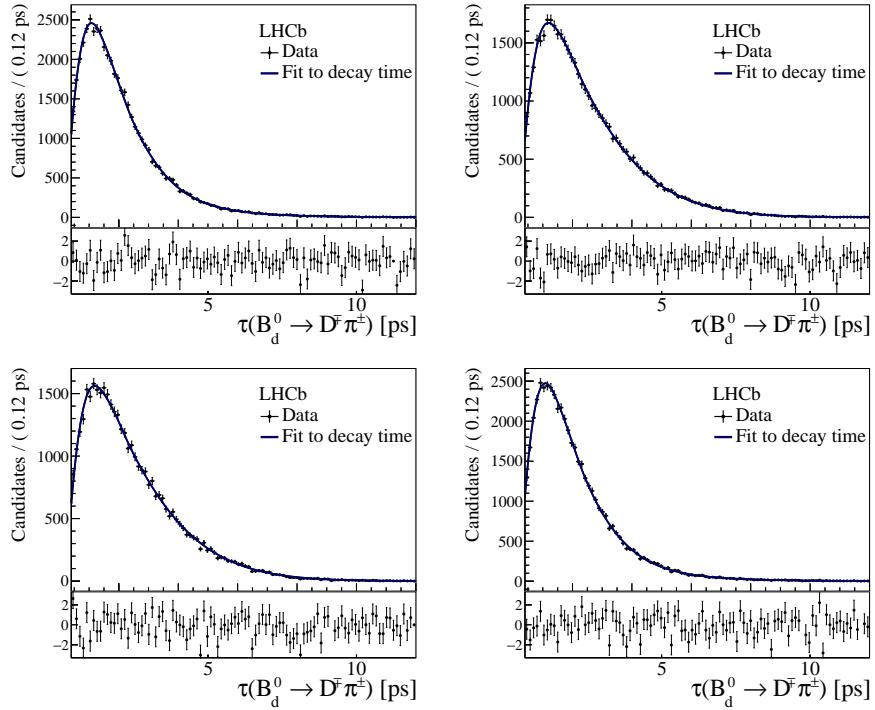


Figure 4.28 – Projections of time PDF onto the decay-time distribution of the *sWeighted* data sample for (top left) $B^0 \rightarrow D^- \pi^+$, (top right) $\bar{B}^0 \rightarrow D^- \pi^+$, (bottom left) $B^0 \rightarrow D^+ \pi^-$ and (bottom right) $B^0 \rightarrow D^+ \pi^+$ for OS inclusively tagged candidates.

1021 Fits in data subsamples

1022 A check of the stability of the results against the different data taking conditions is
1023 performed by repeating the fit in four subsamples of the data, namely data taken with
1024 magnet “up” and “down” polarities, and data taken in 2011 and in 2012. The *sWeights*
1025 for each subsample are obtained via the mass fits described in Sec. 4.4.4. The detailed
1026 results of these time fits are reported in Appendix 4.16. A comparison between the fitted
1027 values for S_f and $S_{\bar{f}}$ obtained in each subsample is shown in Fig. 4.31. In all cases, the
1028 parameters are in agreement and have good p -values; the smallest being the one between
1029 the values of $S_{\bar{f}}$ from the magnet polarity splits (2.7%). In addition, the average between
1030 the fitted values in each split (black line in Fig. 4.31) is always very close with the central
1031 value from the nominal fit (red hatched band).

1032 The stability of the results against the tagging algorithm adopted in the fit are also
1033 checked. In this case, the data sample with *sWeights* obtained from the nominal mass fit
1034 (Sec. 4.4) is split in three independent subsamples according to the tagging decision:

- 1035 • candidates tagged exclusively by the OS tagger;

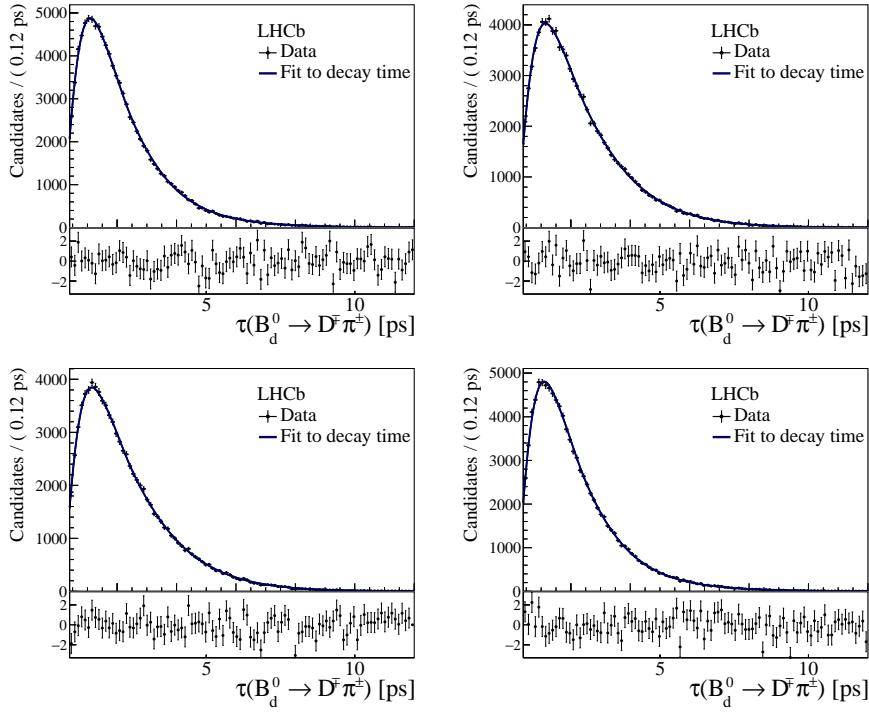


Figure 4.29 – Projections of time PDF onto the decay-time distribution of the *sWeighted* data sample for (top left) $B^0 \rightarrow D^- \pi^+$, (top right) $\bar{B}^0 \rightarrow D^- \pi^+$, (bottom left) $B_d^0 \rightarrow D^+ \pi^-$ and (bottom right) $B_d^0 \rightarrow D^- \pi^+$ for SS inclusively tagged candidates.

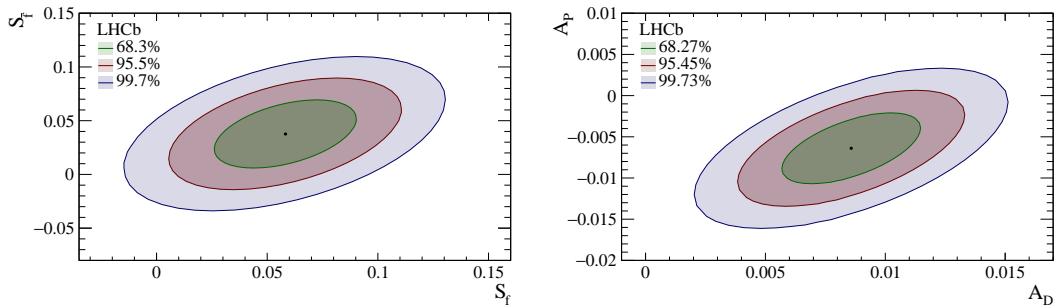


Figure 4.30 – Contour plots for $(S_f, S_{\bar{f}})$ (left) and (A_P, A_D) showing the one, two and three sigma contours. The shown uncertainties include the full statistical uncertainty and the systematic uncertainty due to Gaussian constraints on the mixing frequency Δm and the B^0 decay width Γ . See ?? for the contour plots in wider ranges.

4.7. Decay time fit to data

Table 4.20 – Fitted parameters in the decay time fit. The last column shows existing measurements/independent estimations for a direct comparison, or the Gaussian constraint that is applied to the parameter

Parameter	Value	Comment
S_f	0.058 ± 0.021	Stat. uncert. when fitting w/o Gauss-const. and PIDk syst., 0.0198
$S_{\bar{f}}$	0.038 ± 0.021	Stat. uncert. when fitting w/o Gauss-const. and PIDk syst., 0.0199
A_P	-0.0064 ± 0.0028	Compare with -0.0100 ± 0.0047 (Eq. 4.24)
A_D	0.0086 ± 0.0019	Compare with 0.0086 ± 0.0046 [74]
Γ	$0.6587 \pm 0.0017 \text{ ps}^{-1}$	Gaussian-constrained to $0.6588 \pm 0.0017 \text{ ps}^{-1}$
Δm	$0.5054 \pm 0.0022 \text{ ps}^{-1}$	Gaussian-constrained to $0.5050 \pm 0.0023 \text{ ps}^{-1}$
p_0^{OS}	-0.152 ± 0.021	
p_1^{OS}	-0.035 ± 0.024	
p_2^{OS}	-0.0070 ± 0.0089	
p_3^{OS}	-0.32 ± 0.11	
p_4^{OS}	-0.47 ± 0.49	
Δp_0^{OS}	-0.079 ± 0.049	OS tagger calibration parameters
Δp_1^{OS}	0.141 ± 0.036	
Δp_2^{OS}	-0.024 ± 0.013	
Δp_3^{OS}	-0.26 ± 0.16	
Δp_4^{OS}	-0.52 ± 0.71	
p_0^{SS}	-0.041 ± 0.021	
p_1^{SS}	-0.012 ± 0.022	
Δp_0^{SS}	-0.085 ± 0.044	SS tagger calibration parameters
Δp_1^{SS}	0.043 ± 0.033	
ε_{OS}	0.43237 ± 0.00077	fraction of OS tagged candidates (relative to tagged candidates only)
ε_{SS}	0.93046 ± 0.00040	fraction of SS tagged candidates (relative to tagged candidates only)
v_1	0.3192 ± 0.0062	
v_2	0.494 ± 0.010	
v_3	0.793 ± 0.016	
v_4	0.994 ± 0.019	
v_5	1.093 ± 0.021	time acceptance coefficients
v_6	1.117 ± 0.021	
v_7	1.140 ± 0.021	
v_8	1.175 ± 0.019	
v_9	1.154 ± 0.026	

- 1036 • candidates tagged exclusively by the SS tagger;
- 1037 • candidates tagged by both OS and SS taggers simultaneously.

- 1038 The values of S_f and $S_{\bar{f}}$ obtained in these subsamples are compared in Fig. 4.32. All
- 1039 values are compatible. More details are given in Appendix 4.16. Given the difference of the
- 1040 tagging algorithms and their calibrations, the stability of the results in this test provides
- 1041 additional confidence on the strategy adopted of floating the calibration parameters in
- 1042 the fit.

- 1043 The stability of the results against the B^0 kinematics and global properties of the event is
- 1044 tested. More specifically, the decay time fit is repeated in bins of the following variables:

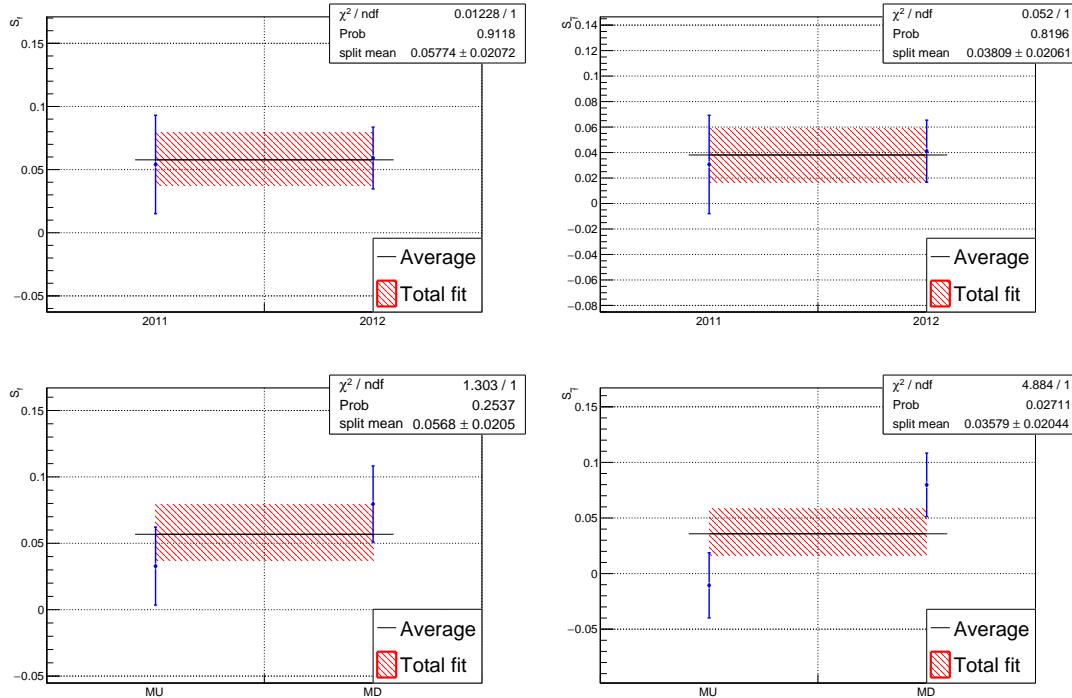


Figure 4.31 – Comparison between the fitted values for S_f (left) and $S_{\bar{f}}$ (right) in the year (top) and polarity (bottom) splits versus the nominal time fit. All the fitted values are shifted by the same, unknown quantity (blinded).

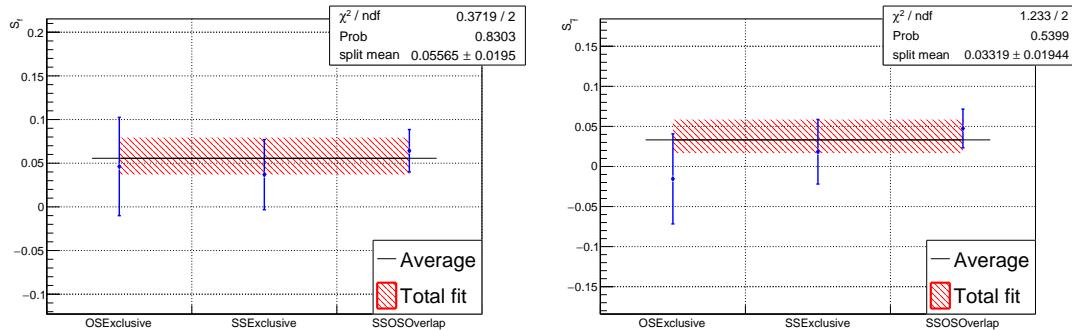


Figure 4.32 – Comparison between the fitted values for S_f (left) and $S_{\bar{f}}$ (right) when candidates tagged exclusively by OS or SS, or both simultaneously are considered. All the fitted values are shifted by the same, unknown quantity (blinded).

-
- 1045 • transverse momentum of the B^0 (4 bins);
 1046 • number of reconstructed primary vertices (3 bins);
 1047 • number of reconstructed tracks (3 bins);
 1048 • difference in pseudorapidity ($\Delta\eta$) between D meson and bachelor pion (4 bins).

1049 The motivation for this test is that flavour tagging calibration parameters depend on the
 1050 above observables; as a consequence, the fitted values for the S_f and $S_{\bar{f}}$ coefficients might
 1051 also show a significant trend in these variables because of the correlation with the flavour
 1052 tagging calibrations. Moreover, the difference in pseudorapidity is sensitive to potential
 1053 misalignments in the detectors which might affect the measured value of CP asymmetries.
 1054 The values of S_f and $S_{\bar{f}}$ obtained in these subsamples are compared in Fig. 4.33, whereas
 1055 more details are given in Appendix 4.16. All values are compatible, and so significant
 1056 dependence of S_f and $S_{\bar{f}}$ over the studied variables is observed.

1057 Finally, the time fit is repeated separately for $B^0 \rightarrow D^- \pi^+$ candidates TOS on L0Hadron
 1058 and all the other candidates. The values of S_f and $S_{\bar{f}}$ obtained in these subsamples are
 1059 compared in Fig. 4.34. All values are compatible, and so significant dependence of S_f
 1060 and $S_{\bar{f}}$ is observed. More details can be found in Appendix 4.16.

1061 **Time fits to bootstrapped Monte Carlo samples**

1062 The fit is also validated using Monte Carlo simulation. The $B^0 \rightarrow D^- \pi^+$ simulated
 1063 sample is *bootstrapped* [75] (*i.e.*, resampled allowing repetition of the same candidate)
 1064 1000 times. Each bootstrapped sample contains the same signal yield as obtained from
 1065 the nominal mass fit (Tab. 4.15), corrected for the *sWeights* dilution factor of Eq. 4.18 to
 1066 have the same effective yield as in the data time fit.

1067 Each sample is then fitted using exactly the same strategy described in Sec. 4.7. The
 1068 Gaussian-constrained parameters, *i.e.* Γ and Δm , are treated in the following way. For
 1069 each fit, a random value of Γ and Δm is drawn from two Gaussian distributions having
 1070 mean and standard deviation equal to the centre and the width of the Gaussian constraints.
 1071 The mean of each Gaussian constraints is taken as the Monte Carlo generation value,
 1072 whereas the standard deviation is the same used for the constraints in the data fit
 1073 (Eq. 4.19). This gives $(1.519\,068 \pm 0.004\,000)$ ps and (0.5100 ± 0.0023) ps $^{-1}$ for τ and
 1074 Δm respectively. Then, these random values are used as the central value of the Gaussian
 1075 constraints in that fit. This allows fluctuations of the Γ and Δm measurements, and
 1076 avoids underestimation of the fitted uncertainties.

1077 The distributions of the fitted value, uncertainty, pull and residual of each parameter⁵.

⁵The residual is defined as fitted value minus generated value, whereas the pull is the residual divided by the fitted uncertainty.

Chapter 4. Measurement of CP violation in $B^0 \rightarrow D^\mp \pi^\pm$ decays

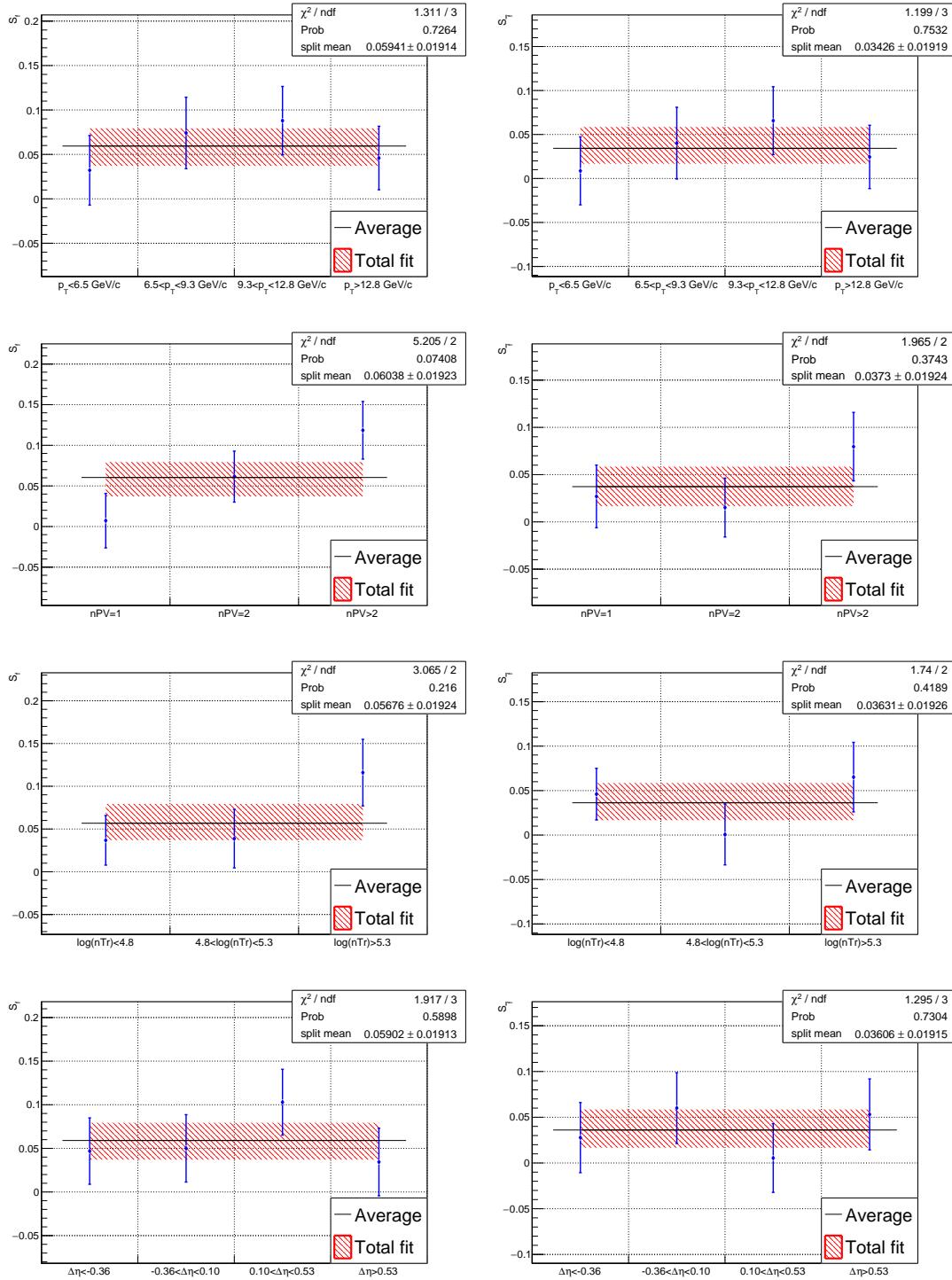


Figure 4.33 – Comparison between the fitted values for S_f (left) and $S_{\bar{f}}$ (right) when the decay time fit is performed in bins of (from top to bottom) the transverse momentum of the B^0 , number of primary vertices, number of tracks and difference in pseudorapidity between the D meson and the bachelor pion. All the fitted values are shifted by the same, unknown quantity (blinded).

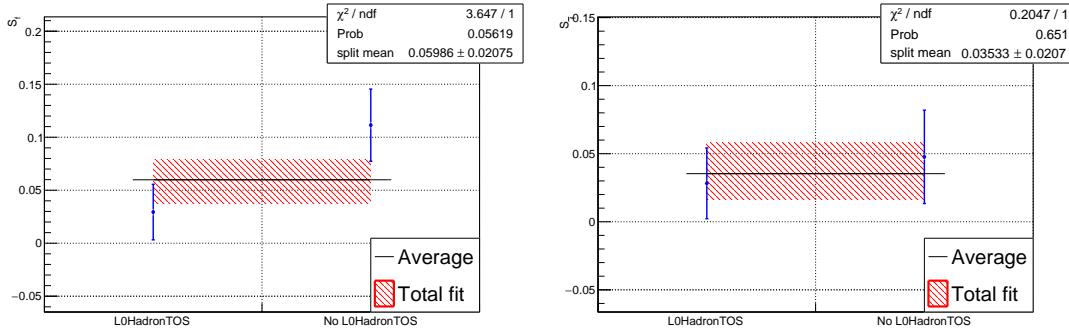


Figure 4.34 – Comparison between the fitted values for S_f (left) and $S_{\bar{f}}$ (right) when the decay time fit is performed separately for candidates TOS on L0Hadron and all the other candidates.

1078 Those of S_f and $S_{\bar{f}}$ are shown in Fig 4.35. Other fitted parameters are reported in
 1079 Appendix 4.17. Pulls and residuals are computed by taking the MC generation value
 1080 as the reference. Each of these distribution is fitted with a Gaussian PDF. The width
 1081 of the fitted pull distributions are close to unity, meaning that the uncertainty coming
 1082 from the maximum likelihood optimisation is correctly estimated. The mean value of the
 1083 distribution of the uncertainties of each parameter is close to the value of the uncertainty
 1084 found in the fit to data. The on-average better precision found in the fit to MC is due to
 1085 the higher tagging performances of the simulation.

1086 The distribution of the residuals of the S_f parameter shows a mean of 0.0071 ± 0.0006 ,
 1087 corresponding to one third of the statistical uncertainty of the fit to data; for $S_{\bar{f}}$, the
 1088 mean is -0.0013 ± 0.0006 , which corresponds to about 6% of the statistical uncertainty
 1089 of the fit to data.

1090 Several configurations are implemented to test the bootstrap study and its results, and
 1091 to try to address the origin of these biases. The fits to the bootstrapped samples are
 1092 repeated in the following different configurations:

- 1093 • using the true flavour of the B candidate instead of the tagging decision and mistag
 1094 probability. No biases are found on S_f and $S_{\bar{f}}$;
- 1095 • using the true flavour of the B candidate instead of the tagging decision and mistag
 1096 probability, but reducing the signal yield to that equivalent to a sample of perfectly
 1097 tagged candidates given the tagging power on data. No biases are found on S_f and
 1098 $S_{\bar{f}}$;
- 1099 • using a *toy* (or *cheated*) tagger, as explained in Appendix 4.18. No biases are found
 1100 on S_f and $S_{\bar{f}}$;

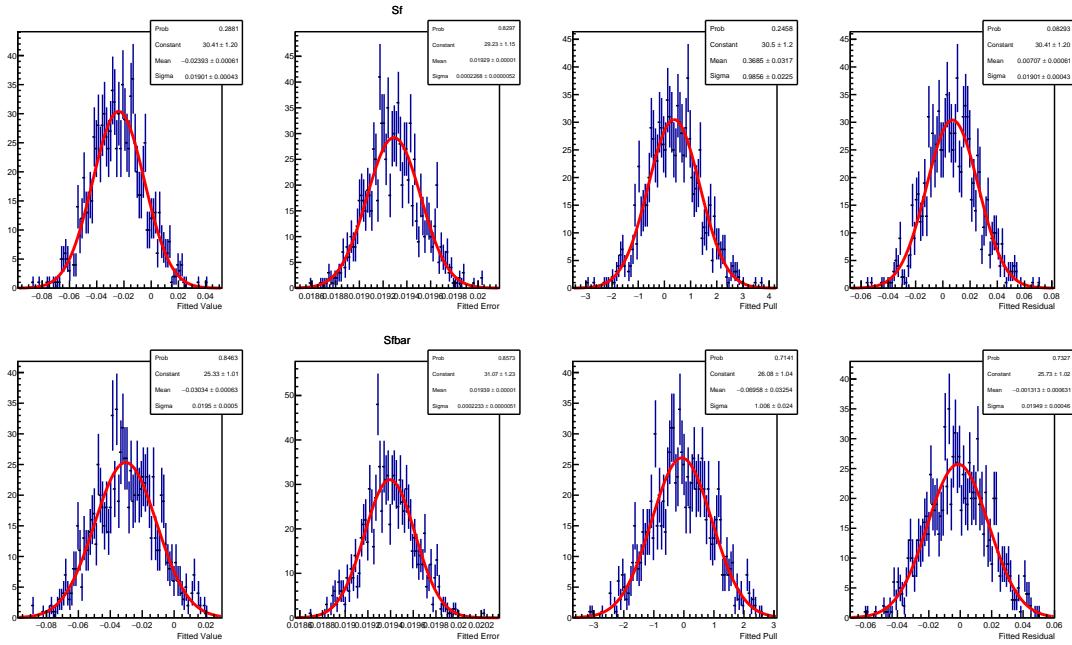


Figure 4.35 – Distributions of fitted value, error, pull and residual for S_f (top) and $S_{\bar{f}}$ (bottom) obtained from fits to bootstrapped Monte Carlo samples. Each distribution is fitted with a Gaussian PDF.

- 1101 • using the calibration parameters obtained in the signal MC sample using the true
1102 flavour information (see Sec. 3.2.1 and Sec. 3.2.2). No biases are found on S_f and
1103 $S_{\bar{f}}$;
- 1104 • fixing the calibration parameters to the values obtained from the MC samples of
1105 the control channels. Biases of the order 1σ on S_f and $S_{\bar{f}}$ are found;
- 1106 • Gaussian-constraints on the calibration parameters using the values obtained from
1107 the MC samples of the control channels. Biases of the order of half the statistical
1108 uncertainty of S_f and $S_{\bar{f}}$ are found;
- 1109 This study confirms that the strategy of floating the calibration parameters in the fit is
1110 the optimal choice. Other than the biases related to the flavour-tagging calibrations, the
1111 origin of the small bias observed on the S_f parameter in the nominal configuration could
1112 not be clarified. To confirm this bias, the study is repeated by fitting additional 1000
1113 bootstrapped samples using an independent fitter. The mean of the distribution of the
1114 residuals in this second study is confirmed to be of the same size, namely 0.0064 ± 0.0007
1115 for S_f and -0.0024 ± 0.0007 for $S_{\bar{f}}$. Hence, the weighted average of the small residuals
1116 on S_f (0.0068 ± 0.0005) and $S_{\bar{f}}$ (-0.0018 ± 0.0005) of both studies are considered as
1117 systematic uncertainties.

4.8 Systematics

The systematic uncertainties are reported in Tab. 4.21 listed by decreasing order of their size. They total 0.011 and 0.007 for S_f and $S_{\bar{f}}$, respectively. A description of each systematic effect is given in the following subsections. The “fit biases” are the residuals observed in the Monte Carlo bootstrap study discussed in section 4.7.1.

Table 4.21 – Systematic uncertainties on the CP parameters S_f and $S_{\bar{f}}$.

Source	S_f	$S_{\bar{f}}$
uncertainty of Δm	0.0073	0.0061
fit biases	0.0068	0.0018
background subtraction	0.0042	0.0023
flavour tagging models	0.0011	0.0015
flavour tagging efficiency asymmetries	0.0012	0.0015
PIDK efficiencies	0.0008	0.0008
acceptance model	0.0007	0.0007
assumption on $\Delta \Gamma$	0.0007	0.0007
assumption on C	0.0006	0.0006
resolution model	0.0012	0.0008
overall	0.0111	0.0073
statistical uncertainties	0.0198	0.0199

4.8.1 Systematic uncertainties from Gaussian constraints

Systematic uncertainties due to external measurements used in the PDF are accounted for through Gaussian constraints in the likelihood. These parameters are the mixing frequency, Δm , and the B^0 lifetime, τ . The fit has been repeated by fixing the Gaussian-constrained parameters to their central values, in order to not propagate the uncertainty of these parameters to the statistical uncertainties of the fit. The statistical uncertainties of S_f and $S_{\bar{f}}$ with Δm fixed are 0.0198 and 0.0199, respectively. Considering the difference in quadrature between the uncertainty from the nominal fit and that from this fit, the systematic uncertainty due to Δm are 0.0073 and 0.0061 for S_f and $S_{\bar{f}}$, respectively. The fit with τ fixed shows that the systematics uncertainty due to this parameter is negligible.

Systematic uncertainty associated to the PIDK efficiencies (Tab. 4.10) are taken into account in the mass fit by means of the Gaussian constraints on these parameters (Sec. 4.4.2). The mass fit is repeated by neglecting these uncertainty in the Gaussian constraints. Then, the time fit is performed with this set of *sWeights*. The difference in quadrature between the uncertainty from this fit and that from the nominal fit gives the systematic due to the binning scheme in the PIDK resampling, which is 0.0008 for both S_f and $S_{\bar{f}}$.

4.8.2 Systematic uncertainties estimated with pseudoexperiments

When computing the systematic uncertainties with pseudoexperiments (or *toys*), a sample with the same size as the data is generated by sampling the PDF with parameters fixed to the value found on the data fit. The values of S_f and $S_{\bar{f}}$ are fixed to those used in the generation of the Monte Carlo sample (Appendix 4.19). In the generation of the samples the PDF is modified to consider alternative models according to the source of systematic uncertainty under investigation. The generated sample is then fitted with the nominal model. For each parameter, the mean of the distribution of the residuals from 1000 toys is taken as the systematic uncertainty. If the mean is consistent with zero, the error on the mean is taken instead. The systematic uncertainties estimated with this toy-based method are the following:

- the flavour tagging calibration model;
- the acceptance model;
- the resolution model;
- the assumption on the C coefficient;
- the assumption on $\Delta\Gamma$.

Flavour tagging calibration model

Toys are generated using for the SS calibration the nominal model with a first order polynomial, and for the OS the model is reduced by one degree compared to the nominal one. In the fit, the calibration models of both taggers are increased by one degree compared to what was used in the generation step. The distribution of the residuals of S_f and $S_{\bar{f}}$ are shown in fig. 4.36. The residuals are not compatible with zero and therefore they are assigned as systematic uncertainties.

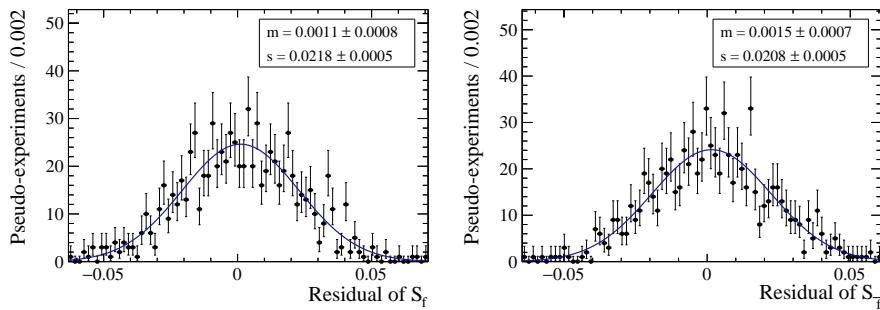


Figure 4.36 – Distribution of residuals for the determination of the systematic uncertainty due to the tagging calibration models.

¹¹⁶³ **Flavour tagging efficiency asymmetries**

¹¹⁶⁴ Toys are generated with the flavour tagging asymmetries set to their estimation from
¹¹⁶⁵ simulation minus their uncertainty, namely -0.14% and -0.13% for the OS and SS
¹¹⁶⁶ tagger, respectively. The distributions of the residuals of S_f and $S_{\bar{f}}$ are shown fig. 4.37.
¹¹⁶⁷ The residuals are not compatible with zero and therefore they are assigned as systematic
 uncertainties.

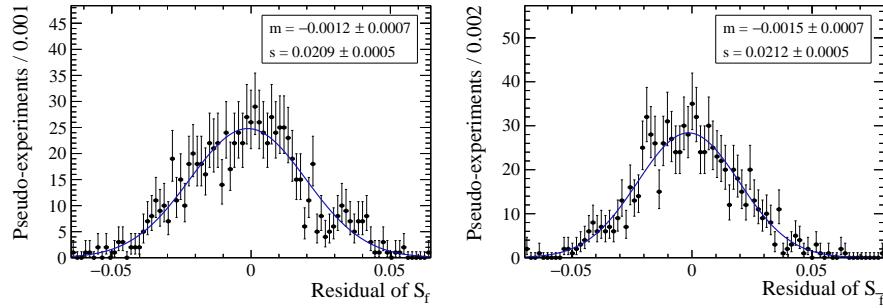


Figure 4.37 – Distribution of residuals for the determination of the systematic uncertainty due to the assumption on the flavour tagging efficiency asymmetry.

¹¹⁶⁸

¹¹⁶⁹ **Acceptance model**

¹¹⁷⁰ The acceptance model is modified in generation by considering new knots for the spline
¹¹⁷¹ function, namely [0.4, 0.45, 0.8, 1.3, 2.5, 6.0, 12.0] ps. The distribution of the residuals of
¹¹⁷² S_f and $S_{\bar{f}}$ are shown in fig. 4.38. Residuals consistent with zero are found and therefore
 the uncertainty on the residuals is assigned as systematic uncertainty.

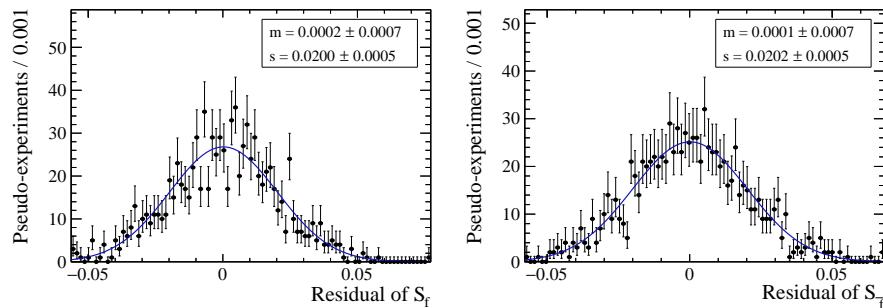


Figure 4.38 – Distribution of residuals for the determination of the systematic uncertainty due to the acceptance model.

¹¹⁷³

¹¹⁷⁴ Decay time resolution

¹¹⁷⁵ Toys are generated with time resolutions 20 fs larger and 20 fs smaller than the nominal
¹¹⁷⁶ value of 55 fs. The distributions of the fitted value of S_f and $S_{\bar{f}}$ are shown in Figs 4.39.
 The largest residual is considered as overall systematic uncertainty.

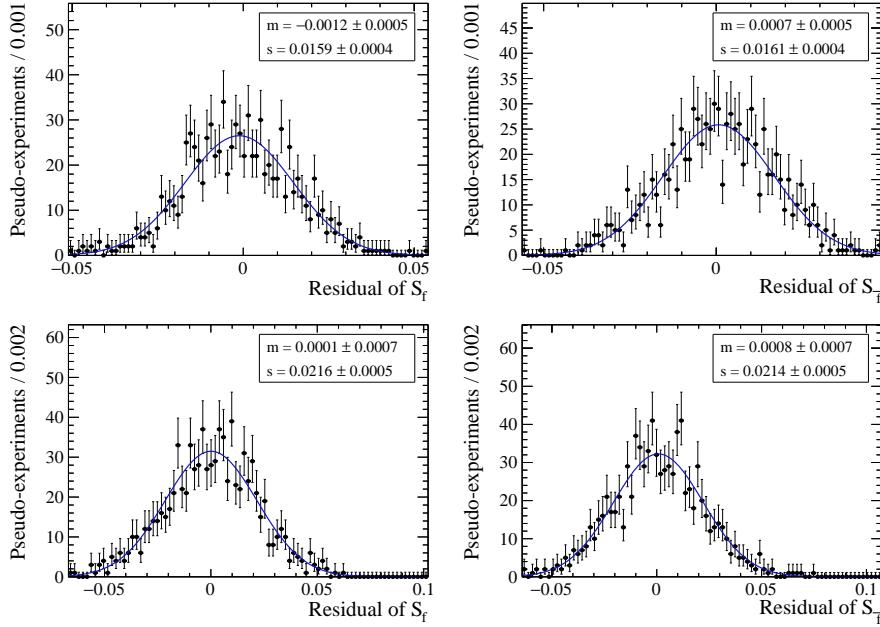


Figure 4.39 – Distribution of fitted values for the determination of the systematic uncertainty due to the resolution model. Top row: 75 fs resolution model; bottom row: 35 fs resolution model.

¹¹⁷⁷

¹¹⁷⁸ Fixed C_f

¹¹⁷⁹ Toys are generated with C_f set to the average of the measurements by Belle and BaBar minus
¹¹⁸⁰ the largest uncertainty between the two measurements, namely 0.993 [76, 77]. The distri-
¹¹⁸¹ butions of the residuals of S_f and $S_{\bar{f}}$ are shown in fig. 4.40. Residuals consistent with zero
¹¹⁸² are found, therefore the uncertainty on the residuals is assigned as systematic uncertainty.

¹¹⁸³

¹¹⁸⁴ Fixed $\Delta\Gamma$

¹¹⁸⁵ Toys are generated with $\Delta\Gamma$ set to the world average value plus its uncertainty, namely
¹¹⁸⁶ 0.0079 ps^{-1} [78]. Moreover, the D_f and $D_{\bar{f}}$ coefficients (defined in Eq. 1.36-1.39) have
¹¹⁸⁷ been fixed to their expected values of -0.0103 and -0.0155 , the same used in the Monte
¹¹⁸⁸ Carlo production of the $B^0 \rightarrow D^- \pi^+$ sample. The distribution of the residuals of

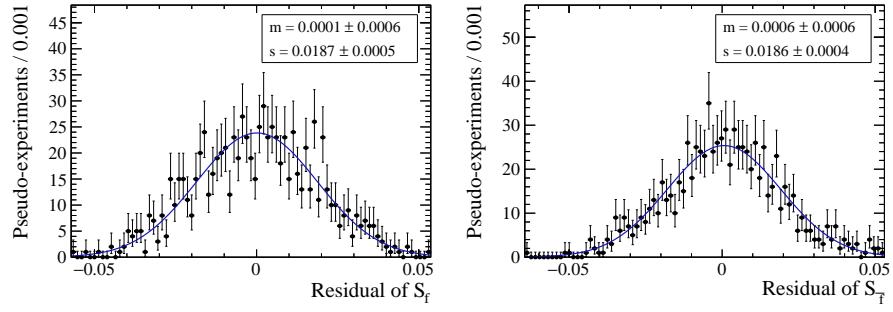


Figure 4.40 – Distribution of residuals for the determination of the systematic uncertainty due to the assumption on the C coefficient.

- ¹¹⁸⁹ S_f and $S_{\bar{f}}$ are shown in fig. 4.41. Residuals consistent with zero are found, therefore the uncertainty on the residuals is assigned as systematic uncertainty.

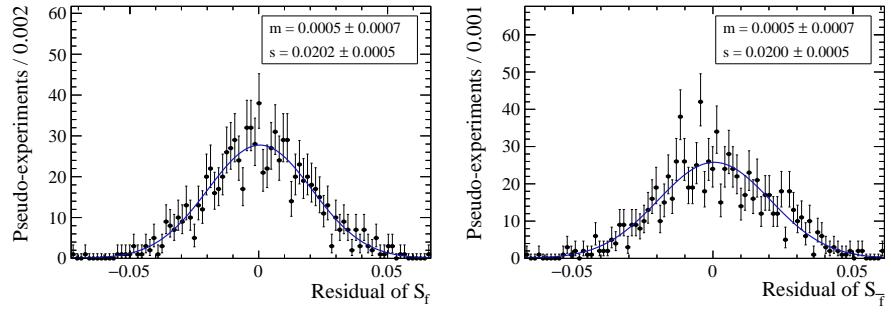


Figure 4.41 – Distribution of fitted values for the determination of the systematic uncertainty due to the assumption on $\Delta\Gamma$.

¹¹⁹⁰

Appendix

1192 **4.9 Opposite side tagging studies**

1193 **4.9.1 Mass fit of $B^+ \rightarrow D^0\pi^+$**

1194 A fit to the mass distribution of B^+ candidates is done to calculate *sWeights*, used in the
1195 subsequent steps of the analysis to subtract the backgrounds surviving the selection. A
1196 two-steps procedure similar to that adopted for the $B^0 \rightarrow D^\mp\pi^\pm$ analysis (“Fit A” in a
1197 wide mass window to account for all backgrounds, and “Fit B” in a subset to calculate
1198 the weights, as described in Sec. 4.4) is used for the fit of the $B^+ \rightarrow D^0\pi^+$ candidates.
1199 The projection of the total PDF on the π sample and the K sample (“Fit A”) is shown
1200 in Fig 4.42, as well as the projection of the total PDF in the reduced sample (“Fit B”).
1201 The π sample and K sample are defined by the PID requirement on the companion track,
1202 $\text{PIDK} < 5$ and $\text{PIDK} > 5$, respectively.

1203 The background components expected in the π sample for the $B^+ \rightarrow D^0\pi^+$ mass fit are
1204 listed below, together with the PDF used for each component:

- 1205 • $B^+ \rightarrow D^0\pi^+$: double-sided Hypatia function.
- 1206 • $B^+ \rightarrow D^0K^+$: double-sided Hypatia function.
- 1207 • $B^0 \rightarrow D^0\pi^+\pi^-$: Crystal ball plus Gaussian function.
- 1208 • $B^+ \rightarrow D^{*0}\pi^+$: JohnsonSU plus Gaussian function.
- 1209 • Combinatorial: single exponential function.

1210 The list for the components expected in the K sample is the following:

- 1211 • $B^+ \rightarrow D^0\pi^+$: double-sided Hypatia function.
- 1212 • $B^+ \rightarrow D^0K^+$: single-sided Hypatia function.
- 1213 • $B^+ \rightarrow D^{*0}\pi^+$: Crystal ball plus exponential function.
- 1214 • $B^+ \rightarrow D^0K^{*+}$: Gaussian function.
- 1215 • Combinatorial: single exponential function.

1216 All the PDFs listed above are defined in details in Appendix 4.13.

1217 The values for the fitted parameters floated in the fit are reported in Tab. 4.22 for Fit A
1218 and Tab. 4.23 for Fit B. The naming convention for each parameter is similar to the one
1219 used in Sec. 4.4.2

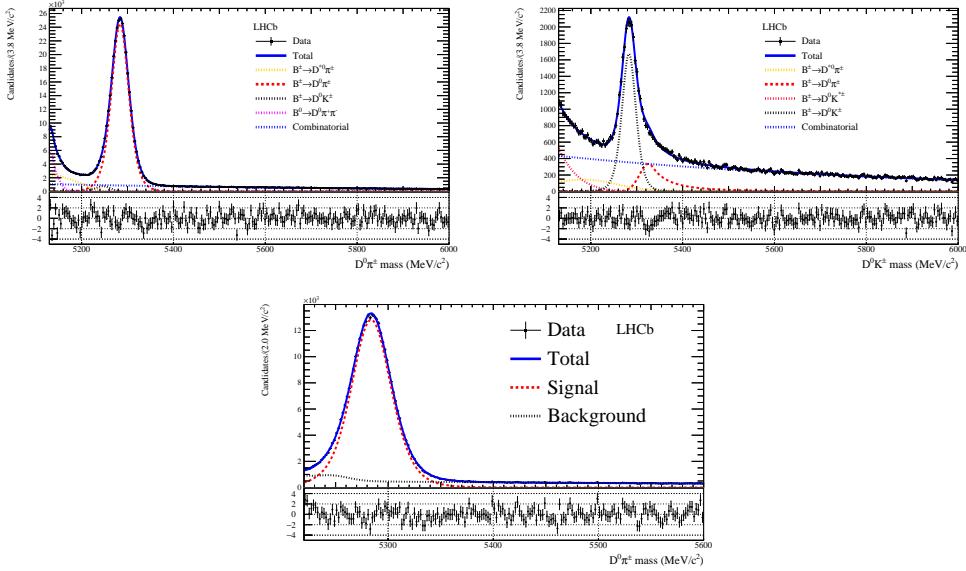


Figure 4.42 – Top: projection of the total PDF in the π and K samples (“Fit A”, simultaneous binned fit in the two samples). Bottom: projection over the π sample (reduced mass range) of the total PDF (“Fit B”, unbinned fit for the $sWeights$ calculation).

1220 4.9.2 Reweighting of $B^+ \rightarrow D^0\pi^+$ to $B^0 \rightarrow D^\pm\pi^\mp$

1221 In order to improve the OS calibration portability, a multi-dimensional reweighting of the
 1222 $sWeighted$ $B^+ \rightarrow D^0\pi^+$ distributions is made for matching the $B^0 \rightarrow D^\pm\pi^\mp$ kinematics.

1223 The reweighting is made in two steps. In the first step, the variables considered in the
 1224 reweighting are the transverse momentum , the pseudo-rapidity η and the decay time τ_B
 1225 of the B candidate, as well as the number of tracks (nTracks) and the number of primary
 1226 vertices (nPv) of the events. A BDT-based approach is followed in order to cope with
 1227 the high dimensionality of the space as well as with the correlations among variables [79].
 1228 A comparison between weighted and unweighted distributions is provided in Figs. 4.43
 1229 and 4.44.

1230 In the second step, a new weight is computed by comparing the two-dimensional distri-
 1231 butions of D meson decay time and HLT2 trigger composition between $B^+ \rightarrow D^0\pi^+$
 1232 and $B_s^0 \rightarrow D^-\pi^+$ after $sWeights$ and the weight from the first step are applied. The
 1233 HLT2 trigger composition observable is a categorical variable which describes which HLT2
 1234 trigger line has been fired by the B candidate:

- 1235 • `Hlt2Topo2BodyBBDTDecision` only (value 0);
- 1236 • `Hlt2Topo3BodyBBDTDecision` or `Hlt2Topo4BodyBBDTDecision` only (value 1);

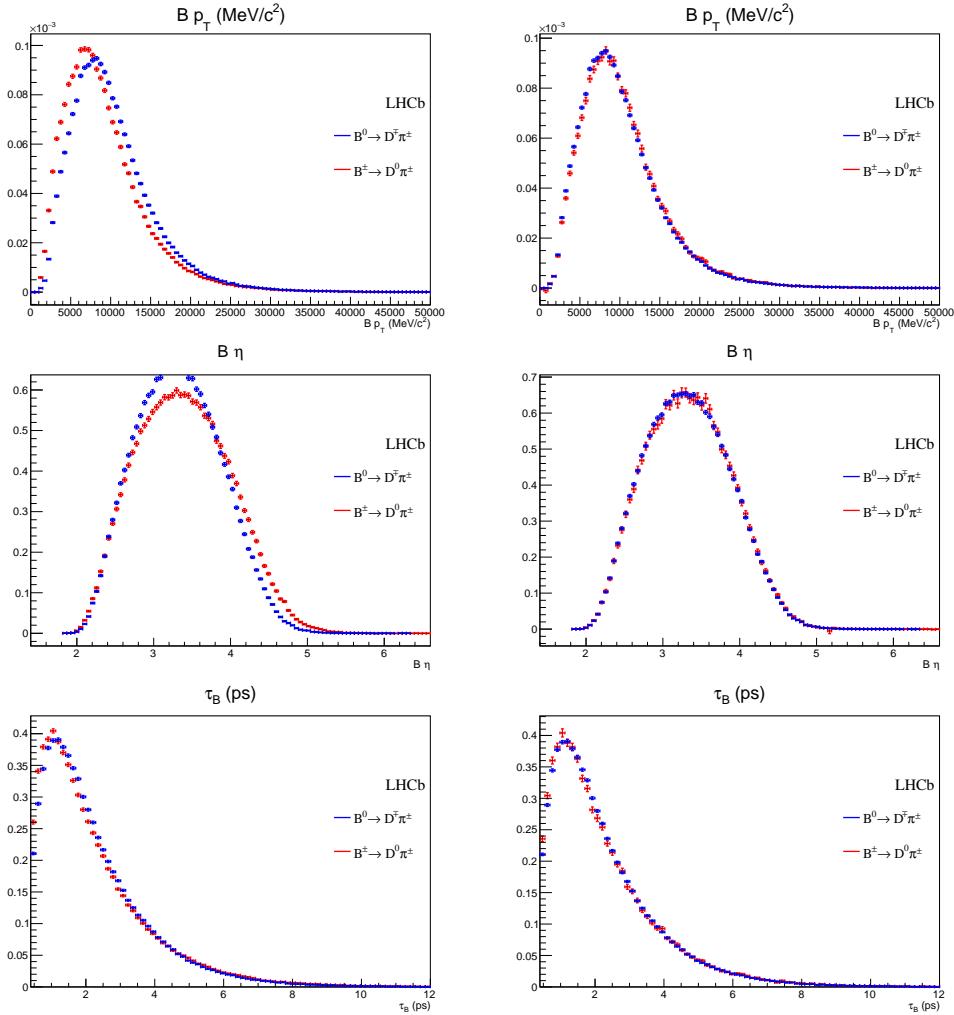


Figure 4.43 – Normalised *sWeighted* distributions of the transverse momentum, the pseudo-rapidity η and the decay time τ_B of the B meson. Left the unweighted distributions, right the distributions after reweighting $B^+ \rightarrow D^0 \pi^+$ are shown.

4.9. Opposite side tagging studies

Table 4.22 – Floating parameters of the total PDF used for Fit A on $B^+ \rightarrow D^0\pi^+$ data.

Parameter number	Parameter name	Fitted value
0	$\mu_{B^0 \rightarrow D\pi\pi}^\pi$	5132.61 ± 0.23
1	$s\sigma_{B^0 \rightarrow D\pi\pi}^\pi$	0.780 ± 0.015
2	$\sigma_{B^+ \rightarrow D^0 K}^K$	19.47 ± 0.30
3	$\sigma_{B^+ \rightarrow D^0 K}^\pi$	16.62 ± 0.69
4	$\mu_{B^+ \rightarrow D^0 K^*}^K$	$(4.96 \pm 0.15)e + 03$
5	$\sigma_{B^+ \rightarrow D^0 K^*}^K$	88 ± 37
6	c_{comb}^K	-0.001305 ± 0.000035
7	c_{comb}^{π}	-0.001279 ± 0.000022
8	$\mu_{B^+ \rightarrow D^0 \pi}^K$	5283.18 ± 0.22
9	$\mu_{B^+ \rightarrow D^0 \pi}^\pi$	5283.880 ± 0.046
10	$sa_{B^+ \rightarrow D^0 \pi}^\pi$	0.804 ± 0.016
11	$\mu_{B^+ \rightarrow D^0 \pi}^R$	5325.4 ± 1.2
12	$sn_{B^+ \rightarrow D^0 \pi}^\pi$	2.70 ± 0.94
13	$\sigma_{B^+ \rightarrow D^0 \pi}^\pi$	22.850 ± 0.054
14	$N_{B^0 \rightarrow D^0 \pi\pi}^\pi$	27245 ± 430
15	$N_{B^+ \rightarrow D^0 K}^K$	18030 ± 296
16	N_{LM}^K	5154 ± 944
17	$N_{B^+ \rightarrow D^{*0} \pi}^K$	5704 ± 1350
18	$N_{B^+ \rightarrow D^{*0} \pi}^\pi$	41871 ± 578
19	N_{comb}^K	58761 ± 555
20	N_{comb}^π	146824 ± 793
21	$N_{B^+ \rightarrow D\pi}^\pi$	322597 ± 812

1237 • overlap of the first two categories (value 2).

1238 This reweighting is done separately from the first one in order to avoid a too fine partition
 1239 of the samples, which would give a very low statistics in less populated bins. The reason
 1240 why HLT2 trigger and D decay time are reweighted simultaneously is that these two
 1241 observables are correlated. The result of this second reweighting is shown in Fig. 4.45.

Table 4.23 – Floating parameters of the total PDF used for Fit B on $B^+ \rightarrow D^0\pi^+$ data.

Parameter number	Parameter name	Fitted value
0	N_{bkg}^π	85687 ± 377
1	$N_{B^+ \rightarrow D\pi}^\pi$	319974 ± 612

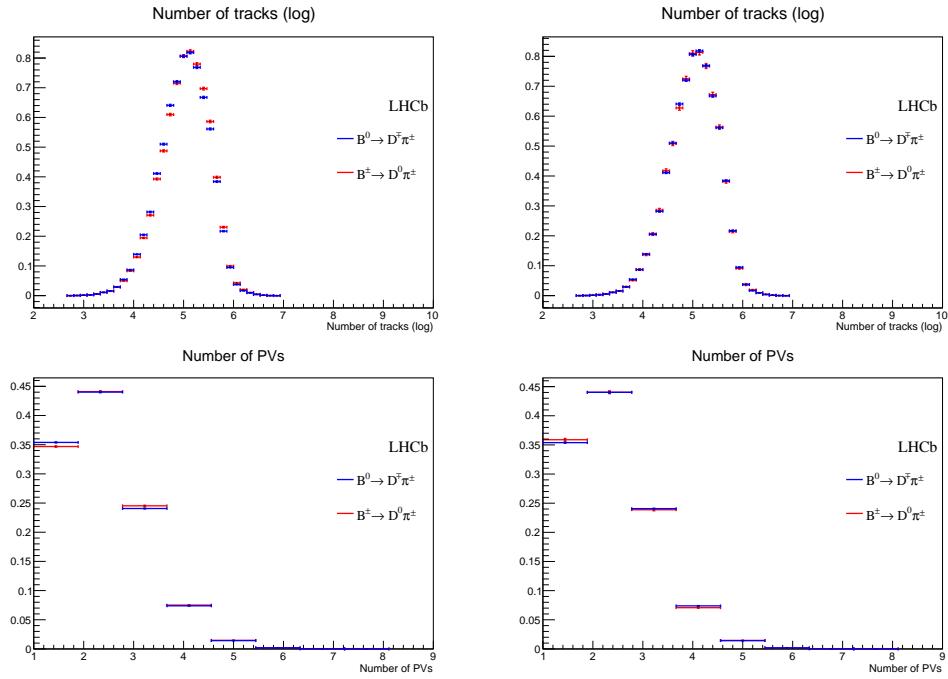


Figure 4.44 – Normalised s Weighted distributions of the number of tracks and PVs in an event. Left the unweighted distributions, right the distributions after reweighting $B^+ \rightarrow D^0\pi^+$ are shown.

4.9.3 GOF tests for OS calibration on $B^+ \rightarrow D^0\pi^+$ data

The number of free parameters (10) used in the GLM model for the OS calibration (Sec. 3.2.1) is the minimum number that allows to obtain satisfactory GOF metrics. The GOF tests are performed automatically by the EPM; the metrics include the Pearson χ^2 , the deviance G^2 , the Cressie-Read (*CR*) metric and the le Cessie-van Houwelingen-Copas-Hosmer metric (*S*), all described in Ref. [43].

All these tests returns a normally distributed score: this means that the score is equal to the distance (measured in standard deviations) from the perfect case, which is a null score. A comparison between the GOF scores obtained for the nominal calibration (10 free parameters) and a simplified model (8 free parameters) is shown in Tab. 4.24. In a simplified model, all scores are order ~ 3 standard deviations or more away from a perfect fit, whereas the scores for the nominal model are ~ 2 at most. For this reason, 10 free parameters are chosen, and the model cannot be simplified further.

4.9. Opposite side tagging studies

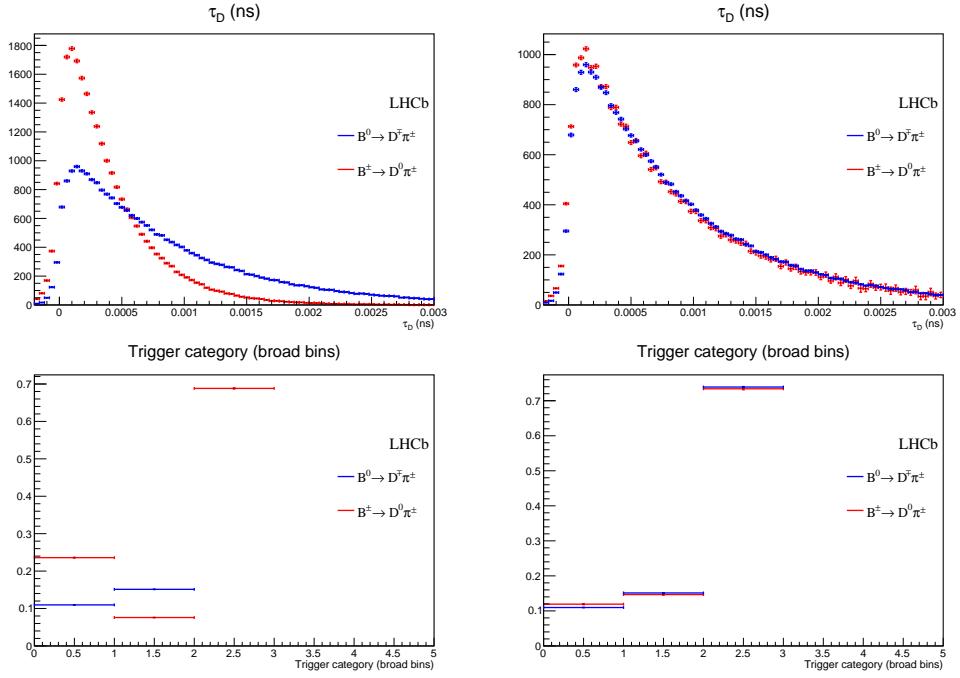


Figure 4.45 – Normalised *sWeighted* distributions of the D meson decay time and HLT2 trigger composition, where the weight obtained from the first reweighting step is also applied. Left the unweighted distributions, right the distributions after reweighting $B^+ \rightarrow D^0\pi^+$ are shown.

Table 4.24 – Fitted OS calibration parameters on the $B^+ \rightarrow D^0\pi^+$ reweighted dataset.

GOF metric	Score (10 parameters)	Score (8 parameters)
χ^2	-2.197	4.1312
G^2	0.71992	-3.8699
CR	-1.6896	2.9273
S	1.847	-4.2701

1255 **4.10 Selection studies**

1256 **4.10.1 BDT input features**

1257 In the following section, the distributions of the input features for the BDT are shown.
 1258 Fig. 4.46 shows the cosine of the direction angle of the B^0 , the χ^2 of the B^0 vertex, the
 1259 χ^2/ndof of the D vertex, the D radial flight distance, the D flight distance χ^2 w.r.t. the
 1260 B^0 vertex and the transverse momentum of the D . In Fig. 4.47, the D IP χ^2 with respect
 1261 to the associated PV and the B^0 vertex, the cosine of the direction angle of the D , the
 1262 IP χ^2 with respect to the associated PV of the bachelor pion, the track χ^2/ndof of the
 1263 bachelor pion and the transverse momentum of the bachelor pion are shown. In Fig. 4.48,
 1264 the IP χ^2 of the associated primary vertex of the D daughters and the χ^2 of the decay
 tree fit with PV constraint are presented.

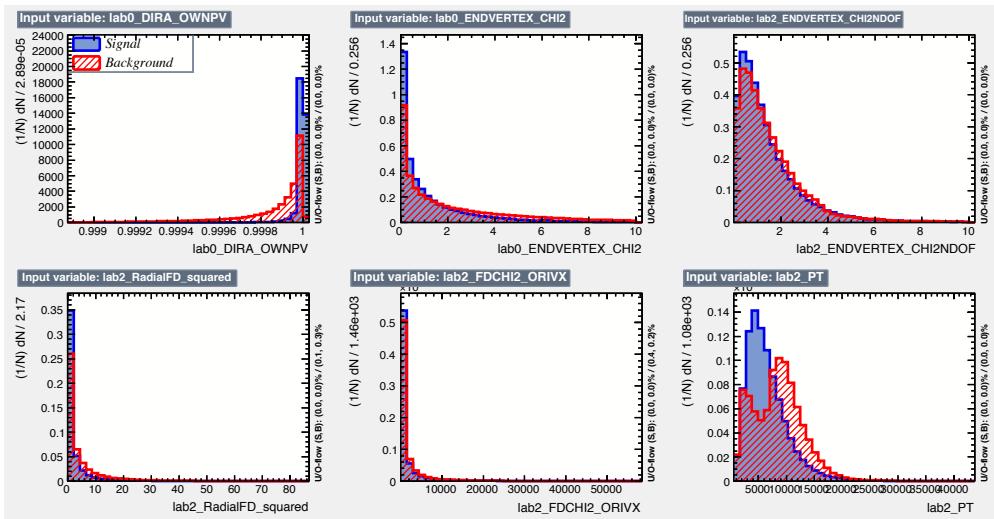


Figure 4.46 – Input features used in the BDT training. From top left to bottom right the cosine of the direction angle of the B^0 , the χ^2 of the B^0 vertex, the χ^2/ndof of the D vertex, the D radial flight distance, the D flight distance χ^2 with respect to the B^0 vertex and the transverse momentum of the D are shown.

1265

1266 **4.10.2 Multiple candidates**

1267 In Tab. 4.25 a summary of multiple candidates left after the stripping selection is given,
 1268 while Tab. 4.26 reports the number of multiple candidates after both stripping and offline
 1269 selection.

4.10. Selection studies

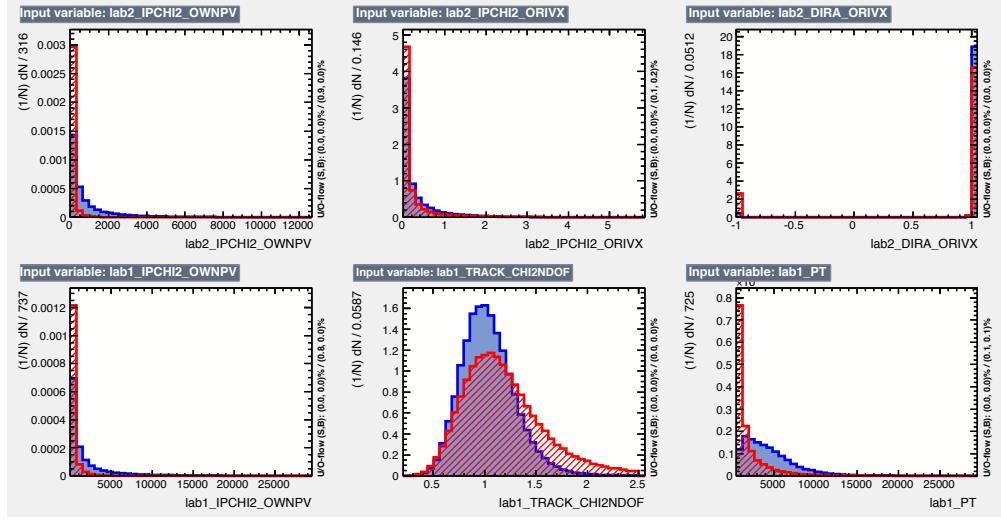


Figure 4.47 – Input features used in the BDT training. From top left to bottom right the D IP χ^2 with respect to the associated PV and the B^0 vertex, the cosine of the direction angle of the D , the IP χ^2 with respect to the associated PV of the bachelor pion, the track χ^2/ndof of the bachelor pion and the transverse momentum of the bachelor pion are shown.

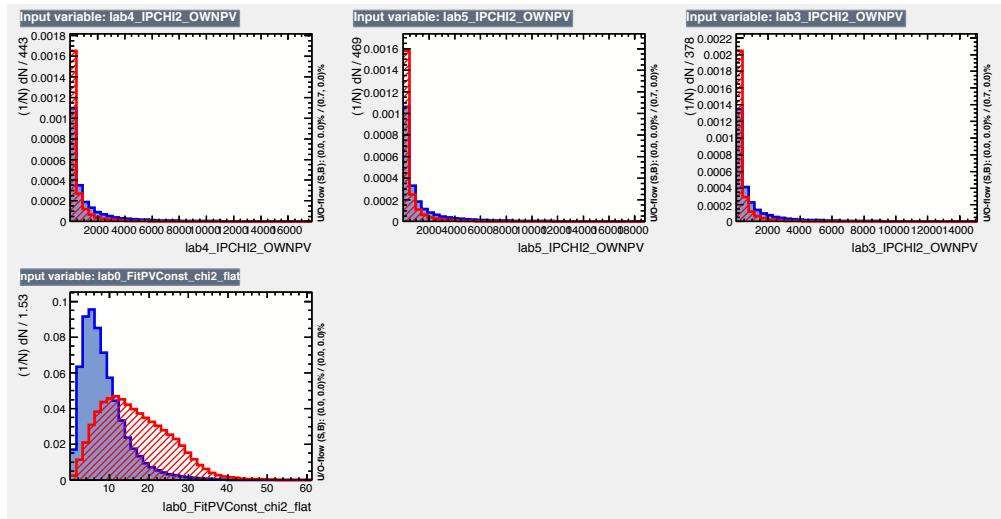


Figure 4.48 – Input features used in the BDT training. From top left to bottom right the IP χ^2 of the associated primary vertex of the D daughters and the χ^2 of the decay tree fit with PV constraint are shown.

Chapter 4. Appendix

Table 4.25 – Number of multiple $B^0 \rightarrow D^\mp\pi^\pm$ candidates left after the stripping selection. The numbers in the first row represent the fraction of B^0 candidates that are not unique in a given event. The numbers in the second row show the fraction of events with multiple B^0 candidates. Additionally, a listing of the occurrence of events with different numbers of multiple B^0 candidates is given. The last row shows the fraction of B^0 candidates that need to be discarded to maintain one candidate per event.

	2011	2012	
fraction of multiple B^0 candidates	18.3 %	19.5 %	
fraction of events with multiple B^0 candidates	9.0 %	9.6 %	
	# B^0 cands	#events	# B^0 cands
			#events
1	5940804	1	16407228
2	483991	2	1426286
3	73902	3	226205
4	20093	4	62640
5	6132	5	19213
6	2505	6	8044
7	1087	7	3326
8	528	8	1686
9	251	9	839
10	146	10	461
11	78	11	279
12	40	12	178
13	28	13	109
14	32	14	85
15	10	15	53
16	12	16	24
17	7	17	16
18	4	18	20
19	5	19	9
20	1	20	11
21	3	21	5
22	2	22	3
23	0	23	2
24	2	24	2
25	1	25	1
26	0	26	4
30	0	30	1
33	0	33	1
40	1	40	0
41	0	41	1
fraction of B^0 candidates to discard		10.1 %	
		11.0 %	

4.10. Selection studies

Table 4.26 – Number of multiple candidates left after both stripping and offline selection. The numbers in the first row represent the fraction of B^0 candidates pairs that are not unique in an event. The numbers in the second row show the fraction of events with multiple B^0 candidates. Additionally, a listing of the occurrence of events with different numbers of multiple B^0 candidates is given. After that, the number of B^0 candidates to discard is given.

	2011	2012		
fraction of multiple B^0 candidates	0.8 %			
fraction of events with multiple B^0 candidates	0.4 %			
	$\#B^0$ cands	#events	$\#B^0$ cands	#events
	1	483074	1	1200956
	2	1886	2	4962
	3	38	3	98
	4	4	4	9
	5	1	5	3
fraction of B^0 candidates to discard	0.4 %		0.4 %	

4.11 Particle Identification plots

The p , η distributions and the PIDK efficiency, misidentification rates before the resampling in bins of p , η are shown in Figs. 4.49 and 4.50 for the bachelor particle of the signal, and in Fig. 4.51 for the D^- daughters.

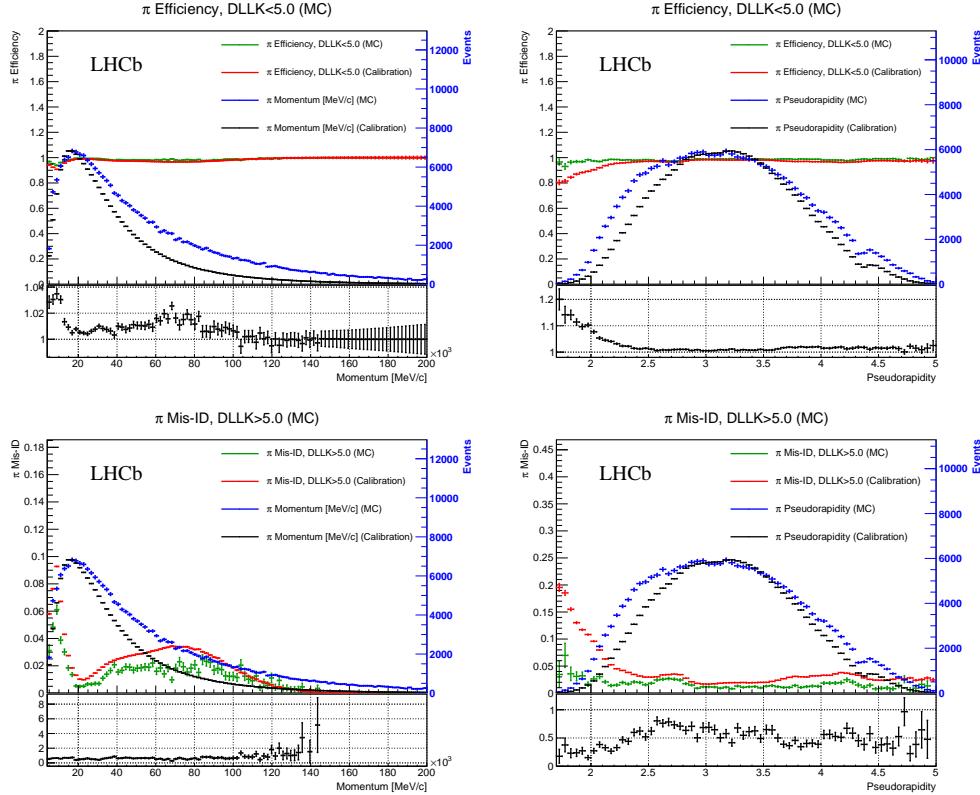


Figure 4.49 – PIDK efficiency or misidentification rates for bachelor pions and p , η distributions before the resampling. Plots from calibration samples and Monte Carlo ($B^0 \rightarrow D\pi$) are superimposed. The ratio of the efficiency or misidentification rate between Monte Carlo and calibration sample is shown in the lower pad (black).

1273

4.11. Particle Identification plots

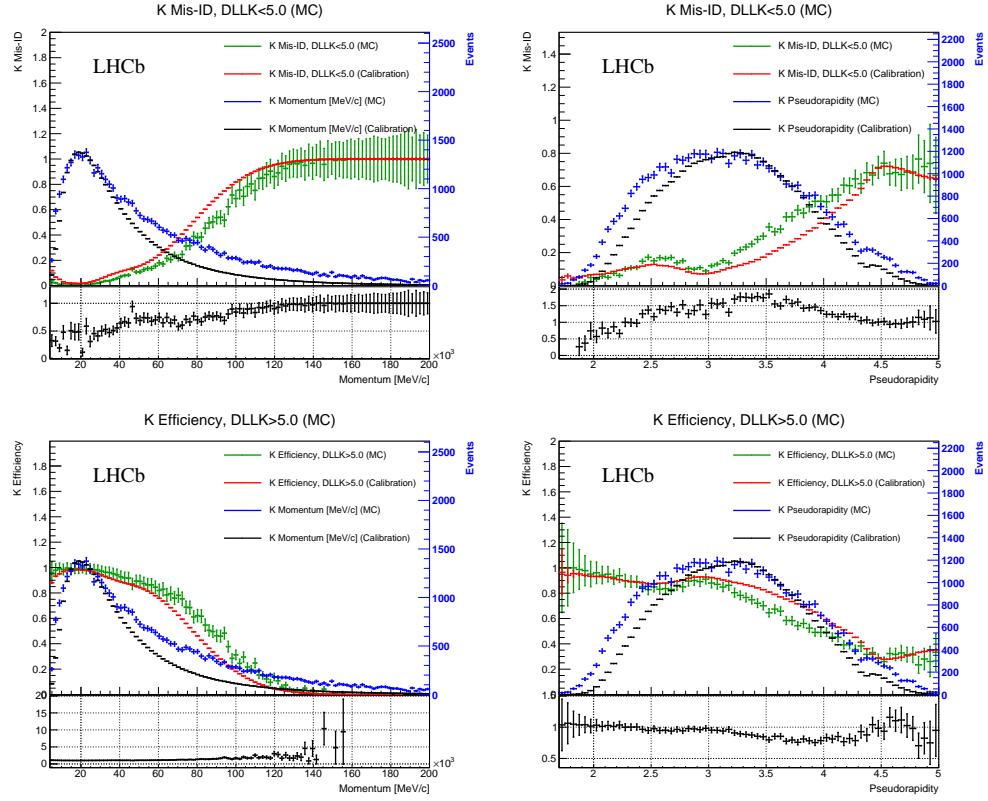


Figure 4.50 – PIDK efficiency or misidentification rates for bachelor kaons and p , η distributions before the resampling. Plots from calibration samples and Monte Carlo ($B^0 \rightarrow DK$) are superimposed. The ratio of the efficiency or misidentification rate between Monte Carlo and calibration sample is shown in the lower pad (black).

Chapter 4. Appendix

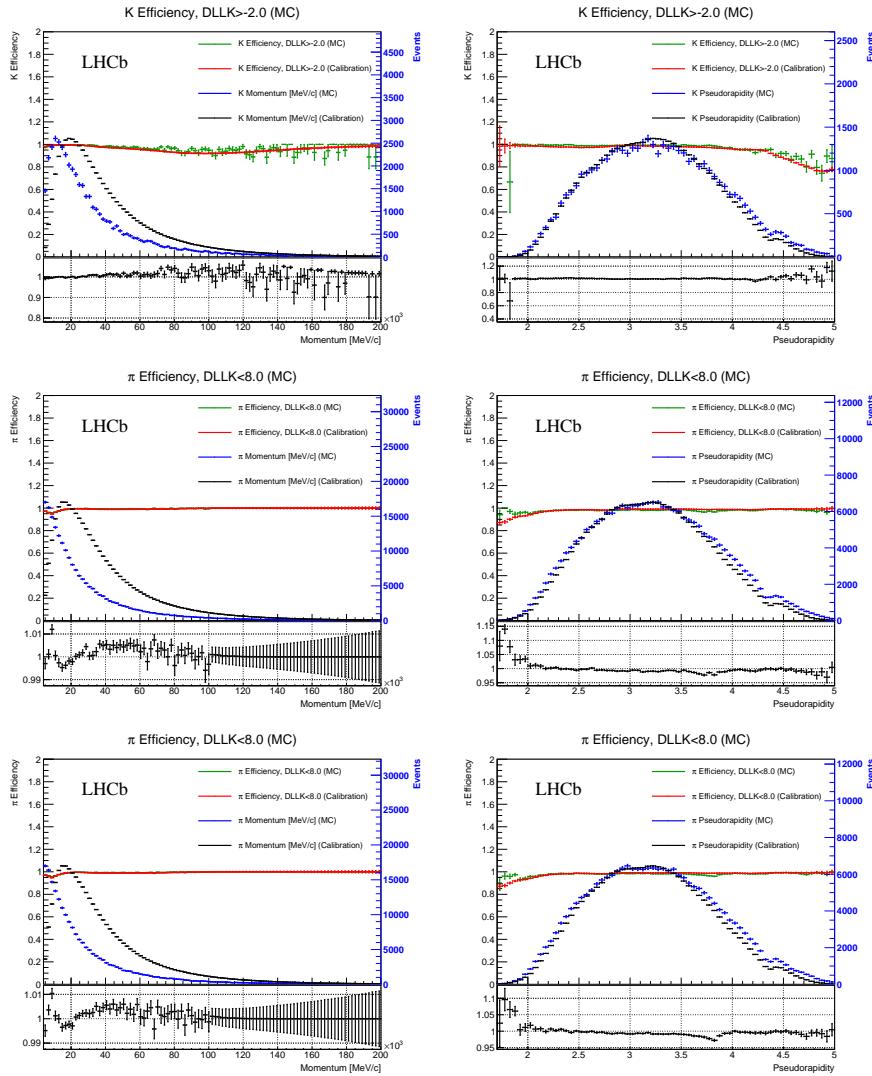


Figure 4.51 – PIDK efficiency, misidentification rates for D^- daughters (pions, kaons) and p, η distributions before the resampling. Plots from calibration samples and Monte Carlo ($B_d \rightarrow DK$ and $B_d \rightarrow D\pi$) are superimposed.

1274 4.12 Invariant mass fit studies

1275 The low correlation between the B^0 invariant mass and decay time is shown by comparing
 1276 the distribution of the decay time in bins of the invariant mass after applying the full
 1277 selection. This is done separately for signal and background. For the signal distribution
 1278 simulated data is used and the decay time is shown in six bins of the invariant mass.
 (fig. 4.52). In order to account for the combinatorial background, the upper mass sideband

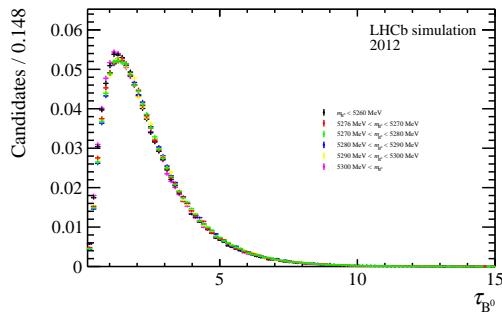


Figure 4.52 – Signal decay time distribution in divided in six bins of the invariant mass. The shapes are shown normalised.

1279
 1280 is chosen as a proxy. Fig. 4.53 shows the decay time in four bins of the invariant mass.
 1281 The physics background contribution in the signal region is considered to be small enough,
 so that even a large correlation does not matter. Given the small differences for all

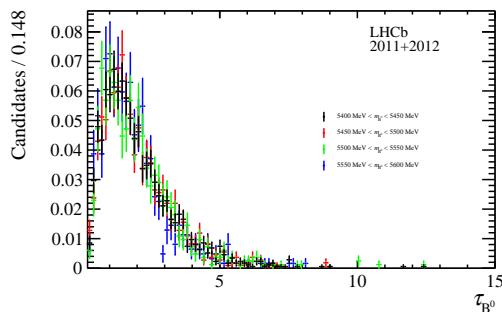


Figure 4.53 – Upper mass side band decay time distribution in divided in four bins of the invariant mass. The shapes are shown normalised.

1282 distributions, the correlations between decay time and invariant mass is assumed to be
 1283 small enough that it's possible to use the invariant mass in the *sPlot* [64] technique for
 1284 disentangling signal from background.

₁₂₈₆ 4.13 PDFs definitions

₁₂₈₇ Throughout the text, each parameter p defined inside a PDF used for the mass fit is
₁₂₈₈ labelled as p_c^s , where $s = \pi, K$ indicates the sample and $c = B^0 \rightarrow D\pi, B^0 \rightarrow D\rho \dots$
₁₂₈₉ indicates the component. For sake of clarity, the s and c labels are dropped in the
₁₂₉₀ equations that follow. The *observable* is always indicated as m . The \propto symbol indicates
₁₂₉₁ that all PDFs are defined up to a normalisation constant, which depends on the interval
₁₂₉₂ chosen for m .

- Exponential function

$$E(m, c) \propto e^{-cm}.$$

- Gaussian function

$$G(m, \mu, \sigma) \propto e^{-\frac{(m-\mu)^2}{2\sigma^2}}.$$

- Double Gaussian function

$$DG(m, \mu, \sigma_1, \sigma_2, f) \propto fe^{-\frac{(m-\mu)^2}{2\sigma_1^2}} + (1-f)e^{-\frac{(m-\mu)^2}{2\sigma_2^2}}.$$

- Single-sided Crystal ball function

Having defined the following parameters:

$$A = \left(\frac{n}{|\alpha|} \right)^n e^{-\frac{|\alpha|^2}{2}}, \quad B = \frac{n}{|\alpha|} - |\alpha|,$$

the single-sided Crystal Ball function is expressed as follows:

$$CB(m, \mu, \sigma, \alpha, n) \propto \begin{cases} e^{-\frac{(m-\mu)^2}{2\sigma^2}}, & \text{if } \frac{m-\mu}{\sigma} > -\alpha, \\ A \left(B - \frac{m-\mu}{\sigma} \right)^{-n}, & \text{if } \frac{m-\mu}{\sigma} \leq -\alpha. \end{cases}$$

- Double-sided Hypatia function

Having defined:

$$h(m, \mu, \sigma, \lambda, \zeta, \beta) \propto \left((m - \mu)^2 + A_\lambda(\zeta) \sigma^2 \right)^{\frac{1}{2}\lambda - \frac{1}{4}} e^{\beta(m-\mu)} K_{\lambda - \frac{1}{2}} \left(\zeta \sqrt{1 + \left(\frac{m - \mu}{A_\lambda(\zeta) \sigma} \right)^2} \right),$$

and its first derivative with respect to m , h' , then the double-sided Hypatia function H is expressed as follows:

$$H(m, \mu, \sigma, \lambda, \zeta, \beta, a_1, n_1, a_2, n_2) \propto$$

$$\begin{cases} h(m, \mu, \sigma, \lambda, \zeta, \beta), & \text{if } \frac{m-\mu}{\sigma} > -a_1 \text{ or } \frac{m-\mu}{\sigma} < a_2, \\ \frac{h(\mu-a_1\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{\left(1-m/\left(n^{\frac{h(\mu-a_1\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{h'(\mu-a_1\sigma, \mu, \sigma, \lambda, \zeta, \beta)}-a_1\sigma}\right)\right)^{n_1}}, & \text{if } \frac{m-\mu}{\sigma} \leq -a_1, \\ \frac{h(\mu-a_2\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{\left(1-m/\left(n^{\frac{h(\mu-a_2\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{h'(\mu-a_2\sigma, \mu, \sigma, \lambda, \zeta, \beta)}-a_2\sigma}\right)\right)^{n_2}}, & \text{if } \frac{m-\mu}{\sigma} \geq a_2. \end{cases}$$

The K_λ functions are special Bessel functions of third kind, whereas A_λ is defined as:

$$A_\lambda^2 = \frac{\zeta K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)}.$$

¹²⁹³ • **Single-sided Hypatia function**

¹²⁹⁴ A single-sided Hypatia function is obtained from a double-sided Hypatia function
¹²⁹⁵ in the limit $a_2 \rightarrow +\infty$, $n_2 = 0$ (and by labelling a_1 and n_1 as a and n respectively).

• **Johnson SU function**

Having defined the following parameters:

$$w = e^{\tau^2}$$

$$\omega = -\nu\tau,$$

$$c = \frac{1}{\sqrt{\frac{1}{2}(w-1)(w \cosh 2\omega + 1)}},$$

$$z = \frac{m - (\mu + c + \sigma\sqrt{w} \sinh \omega)}{c\sigma},$$

$$r = -\nu + \frac{\sinh^{-1} z}{\tau},$$

the Johnson SU function is expressed as follows:

$$J(m, \mu, \sigma, \nu, \tau) \propto \frac{1}{2\pi c(\nu, \tau)\sigma} e^{-\frac{1}{2}r(m, \mu, \sigma, \nu, \tau)^2} \frac{1}{\tau\sqrt{z(m, \mu, \sigma, \nu, \tau)^2 + 1}}.$$

1296 4.14 PDF for the decay-time fit

1297 The PDF describing the B^0 decay time distribution can be written in the most general
 1298 case as follows:

$$P(t|\eta_{OS}, \eta_{SS}, d_{OS}, d_{SS}) = \frac{1}{4\tau} e^{-\frac{t}{\tau}} \left[C_{\cosh}^{\text{eff}} \cosh \left(\frac{\Delta\Gamma t}{2} \right) + C_{\sinh}^{\text{eff}} \sinh \left(\frac{\Delta\Gamma t}{2} \right) \right. \\ \left. + C_{\cos}^{\text{eff}} \cos(\Delta m t) + C_{\sin}^{\text{eff}} \sin(\Delta m t) \right]. \quad (4.25)$$

1299 The four “effective” CP coefficients inside Eq. 4.25 depend on the final state ($f = D^+ \pi^-$
 1300 or $f = D^- \pi^+$), the tagging decision ($d_i = \pm 1$, $d_i = 0$, for $i = OS, SS$), the mistag
 1301 and the tagging efficiency of the OS and SS taggers, and the asymmetries (production,
 1302 detection and tagging efficiency). In the ideal case (no asymmetries, zero mistag, 100%
 1303 tagging efficiency), these effective coefficients become the “physical” coefficients already
 1304 introduced in Sec. 4.1.

1305 The OS and SS taggers are combined on the fly during the time fit. All the steps to build
 1306 the final PDF are described in details below.

The tagging efficiency ϵ_{tag}^i of the i -th tagger ($i = OS, SS$) is corrected for the tagging
 efficiency asymmetry A_{eff}^i as follows:

$$\epsilon_b^i = \epsilon_{\text{tag}}^i - \frac{1}{2} A_{\text{eff}}^i, \quad (4.26)$$

$$\epsilon_{\bar{b}}^i = \epsilon_{\text{tag}}^i + \frac{1}{2} A_{\text{eff}}^i, \quad (4.27)$$

1307 where ϵ_b^i and $\epsilon_{\bar{b}}^i$ are the two different tagging efficiencies for B^0 and \bar{B}^0 . Moreover, the
 1308 OS and SS mistag are calibrated taking into account asymmetries between B^0 and \bar{B}^0
 1309 according to Eq. 3.10. We will refer to them here as ω_i^b and $\omega_i^{\bar{b}}$ (for $i = OS, SS$). It’s now
 1310 convenient to define the quantity $\Delta^\pm(d_{OS}, d_{SS})$ in the following way according to the OS
 1311 and SS tagging decisions.

- If $d_{OS} = d_{SS} = 0$ (un-tagged B^0 candidate):

$$\Delta^\pm = (1 - \epsilon_b^{OS} - \epsilon_b^{SS} + \epsilon_b^{OS} * \epsilon_b^{SS}) \pm (1 - \epsilon_{\bar{b}}^{OS} - \epsilon_{\bar{b}}^{SS} + \epsilon_{\bar{b}}^{OS} * \epsilon_{\bar{b}}^{SS}). \quad (4.28)$$

- If $d_i = 0$, $d_j \neq 0$ and $i \neq j$ (B^0 candidate tagged by one tagger):

$$\begin{aligned}\Delta^\pm &= \frac{1}{2}\epsilon_b^j \left[1 - \epsilon_b^i + d_j \left(1 - \epsilon_b^i - 2\omega_j^b + 2\omega_j^b \epsilon_b^i \right) \right] \\ &\pm \frac{1}{2}\epsilon_b^j \left[1 - \epsilon_b^i + d_j \left(1 - \epsilon_b^i - 2\omega_j^b + 2\omega_j^b \epsilon_b^i \right) \right]\end{aligned}\quad (4.29)$$

- If $d_{OS} = d_{SS} = 1$ (B^0 candidates tagged by both taggers):

$$\begin{aligned}\Delta^\pm &= \frac{1}{4}\epsilon_b^{SS}\epsilon_b^{OS} \left[1 + d_{SS} \left(1 - 2\omega_{SS}^b \right) + d_{OS} \left(1 - 2\omega_{OS}^b \right) \right. \\ &\quad \left. + d_{OS}d_{SS} \left(1 - 2\omega_{SS}^b - 2\omega_{OS}^b + 4\omega_{SS}^b\omega_{OS}^b \right) \right] \\ &\pm \frac{1}{4}\epsilon_b^{SS}\epsilon_b^{OS} \left[1 + d_{SS} \left(1 - 2\omega_{SS}^b \right) + d_{OS} \left(1 - 2\omega_{OS}^b \right) \right. \\ &\quad \left. + d_{OS}d_{SS} \left(1 - 2\omega_{SS}^b - 2\omega_{OS}^b + 4\omega_{SS}^b\omega_{OS}^b \right) \right].\end{aligned}\quad (4.30)$$

Finally, the effective CP coefficients can be written as follows:

$$C_{\sin, \cos}^{\text{eff}} = \begin{cases} (1 + A_D)C_{\sin, \cos}^{\text{phys}}(\Delta^- - A_P\Delta^+), & \text{for } f = D^-\pi^+ \\ -(1 - A_D)C_{\sin, \cos}^{\text{phys}}(\Delta^- - A_P\Delta^+), & \text{for } f = D^+\pi^- \end{cases} \quad (4.31)$$

$$C_{\sinh, \cosh}^{\text{eff}} = \begin{cases} (1 + A_D)C_{\sinh}^{\text{phys}}(\Delta^+ - A_P\Delta^-), & \text{for } f = D^-\pi^+ \\ (1 - A_D)C_{\sinh}^{\text{phys}}(\Delta^+ - A_P\Delta^-), & \text{for } f = D^+\pi^- \end{cases} \quad (4.32)$$

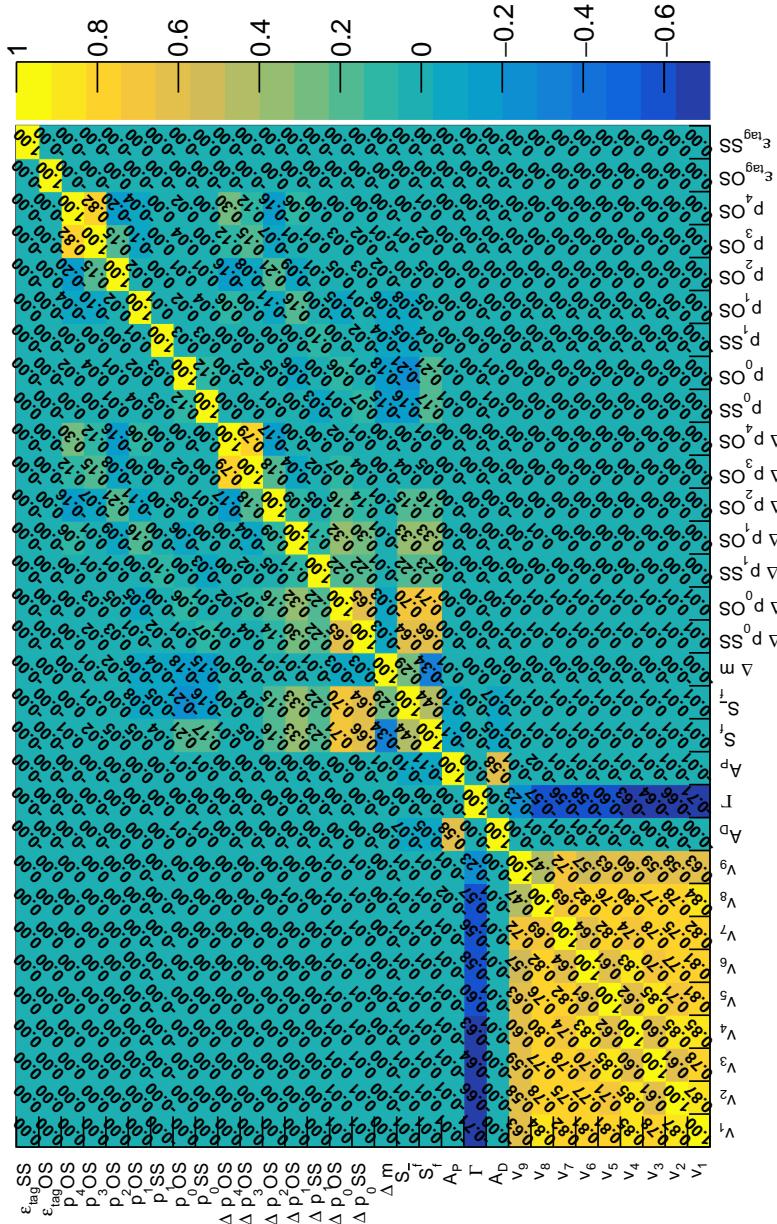
¹³¹² 4.15 Correlation matrix of decay-time fit to $B^0 \rightarrow D^\mp \pi^\pm$


Figure 4.54 – Correlation matrix between the floating parameters in the fit.

4.16. Decay-time fits in $B^0 \rightarrow D^\mp\pi^\pm$ data subsamples

₁₃₁₃ 4.16 Decay-time fits in $B^0 \rightarrow D^\mp\pi^\pm$ data subsamples

Chapter 4. Appendix

Table 4.27 – Fitted parameters in the blinded time fits to the 2011, 2012, magnet up and down subsamples. The reported values for S_f and $S_{\bar{f}}$ are all shifted by the same, unknown amount (blinded).

Parameter	2011	2012	Magnet Up	Magnet Down
v_1	0.350 ± 0.011	0.3082 ± 0.0066	0.3199 ± 0.0078	0.3186 ± 0.0075
v_2	0.535 ± 0.018	0.479 ± 0.011	0.506 ± 0.014	0.483 ± 0.013
v_3	0.886 ± 0.029	0.760 ± 0.018	0.786 ± 0.021	0.800 ± 0.020
v_4	1.055 ± 0.034	0.973 ± 0.021	1.013 ± 0.026	0.978 ± 0.024
v_5	1.186 ± 0.038	1.060 ± 0.023	1.096 ± 0.028	1.089 ± 0.026
v_6	1.176 ± 0.037	1.095 ± 0.023	1.140 ± 0.028	1.096 ± 0.026
v_7	1.231 ± 0.039	1.108 ± 0.023	1.156 ± 0.028	1.126 ± 0.027
v_8	1.267 ± 0.035	1.143 ± 0.021	1.189 ± 0.026	1.163 ± 0.024
v_9	1.203 ± 0.050	1.135 ± 0.029	1.174 ± 0.037	1.136 ± 0.035
A_D	0.0137 ± 0.0036	0.0065 ± 0.0022	0.0075 ± 0.0027	0.0096 ± 0.0026
Γ	0.6588 ± 0.0017	0.6587 ± 0.0017	0.6587 ± 0.0017	0.6588 ± 0.0017
A_P	-0.0022 ± 0.0054	-0.0079 ± 0.0033	-0.0056 ± 0.0041	-0.0072 ± 0.0039
S_f	0.054 ± 0.039	0.059 ± 0.024	0.032 ± 0.029	0.079 ± 0.029
$S_{\bar{f}}$	0.031 ± 0.039	0.041 ± 0.024	-0.010 ± 0.029	0.080 ± 0.029
Δm	0.5047 ± 0.0023	0.5058 ± 0.0023	0.5054 ± 0.0023	0.5050 ± 0.0023
$\Delta p_0 SS$	-0.142 ± 0.083	-0.061 ± 0.052	-0.102 ± 0.063	-0.076 ± 0.061
$\Delta p_0 OS$	-0.117 ± 0.092	-0.065 ± 0.057	-0.103 ± 0.070	-0.063 ± 0.068
$\Delta p_1 SS$	0.068 ± 0.062	0.030 ± 0.039	0.048 ± 0.046	0.036 ± 0.046
$\Delta p_1 OS$	0.064 ± 0.067	0.172 ± 0.042	0.187 ± 0.051	0.101 ± 0.050
$\Delta p_2 OS$	-0.036 ± 0.025	-0.018 ± 0.015	-0.020 ± 0.019	-0.031 ± 0.017
$\Delta p_3 OS$	-0.08 ± 0.30	-0.29 ± 0.20	-0.17 ± 0.26	-0.34 ± 0.21
$\Delta p_4 OS$	0.8 ± 1.4	-1.5 ± 1.2	-0.2 ± 1.5	-0.72 ± 0.89
$p_0 SS$	-0.046 ± 0.039	-0.041 ± 0.025	0.020 ± 0.030	-0.097 ± 0.030
$p_0 OS$	-0.126 ± 0.038	-0.166 ± 0.025	-0.143 ± 0.030	-0.160 ± 0.029
$p_1 SS$	0.005 ± 0.042	-0.019 ± 0.026	-0.022 ± 0.031	-0.004 ± 0.032
$p_1 OS$	-0.095 ± 0.044	-0.011 ± 0.028	-0.054 ± 0.034	-0.014 ± 0.033
$p_2 OS$	-0.025 ± 0.018	-0.000 ± 0.010	-0.013 ± 0.013	-0.004 ± 0.012
$p_3 OS$	-0.03 ± 0.21	-0.43 ± 0.14	-0.20 ± 0.18	-0.40 ± 0.15
$p_4 OS$	1.08 ± 0.94	-1.41 ± 0.70	-0.11 ± 0.92	-0.63 ± 0.64
$\varepsilon_{tag} OS$	0.4270 ± 0.0014	0.43458 ± 0.00091	0.4321 ± 0.0011	0.4326 ± 0.0011
$\varepsilon_{tag} SS$	0.92753 ± 0.00076	0.93163 ± 0.00046	0.92978 ± 0.00058	0.93107 ± 0.00054

4.16. Decay-time fits in $B^0 \rightarrow D^\mp \pi^\pm$ data subsamples

Table 4.28 – Fitted parameters in the blinded time fits when only OS tagged, SS tagged or OS and SS candidates are considered exclusively. The reported values for S_f and $S_{\bar{f}}$ are all shifted by the same, unknown amount (blinded).

Parameter	OS exclusive	SS exclusive	OS and SS exclusive
v_1	0.276 ± 0.014	0.3155 ± 0.0070	0.3344 ± 0.0087
v_2	0.403 ± 0.024	0.500 ± 0.012	0.504 ± 0.015
v_3	0.668 ± 0.040	0.793 ± 0.019	0.821 ± 0.023
v_4	0.903 ± 0.050	0.992 ± 0.023	1.017 ± 0.028
v_5	1.070 ± 0.059	1.083 ± 0.024	1.112 ± 0.030
v_6	1.045 ± 0.058	1.100 ± 0.024	1.159 ± 0.031
v_7	1.154 ± 0.062	1.137 ± 0.025	1.142 ± 0.030
v_8	1.138 ± 0.054	1.165 ± 0.022	1.199 ± 0.028
v_9	1.117 ± 0.086	1.148 ± 0.032	1.171 ± 0.040
A_D	0.0100 ± 0.0066	0.0086 ± 0.0024	0.0083 ± 0.0030
Γ	0.6588 ± 0.0017	0.6587 ± 0.0017	0.6587 ± 0.0017
A_P	0.009 ± 0.010	-0.0110 ± 0.0036	-0.0022 ± 0.0044
S_f	0.046 ± 0.056	0.037 ± 0.040	0.064 ± 0.024
$S_{\bar{f}}$	-0.015 ± 0.056	0.019 ± 0.040	0.048 ± 0.024
Δm	0.5051 ± 0.0023	0.5055 ± 0.0023	0.5049 ± 0.0023
$\Delta p_0 SS$		-0.102 ± 0.075	-0.112 ± 0.058
$\Delta p_0 OS$	-0.10 ± 0.13		-0.059 ± 0.055
$\Delta p_1 SS$		0.014 ± 0.041	0.075 ± 0.050
$\Delta p_1 OS$	0.120 ± 0.089		0.154 ± 0.038
$\Delta p_2 OS$	-0.071 ± 0.033		-0.015 ± 0.013
$\Delta p_3 OS$	-0.34 ± 0.37		-0.24 ± 0.16
$\Delta p_4 OS$	0.9 ± 2.1		-0.59 ± 0.71
$p_0 SS$		-0.023 ± 0.028	-0.072 ± 0.031
$p_0 OS$	-0.164 ± 0.051		-0.148 ± 0.022
$p_1 SS$		-0.027 ± 0.027	0.008 ± 0.034
$p_1 OS$	-0.017 ± 0.058		-0.036 ± 0.024
$p_2 OS$	-0.007 ± 0.023		-0.0077 ± 0.0091
$p_3 OS$	-0.16 ± 0.26		-0.34 ± 0.11
$p_4 OS$	0.2 ± 1.4		-0.55 ± 0.50

Chapter 4. Appendix

Table 4.29 – Fitted parameters in the blinded time fits in bins of B^0 transverse momentum (GeV/c). The reported values for S_f and $S_{\bar{f}}$ are all shifted by the same, unknown amount (blinded).

Parameter	[0, 6.5)	[6.5, 9.3)	[9.3, 12.8)	[12.8, $+\infty$)
v_1	0.0492 ± 0.0018	0.2480 ± 0.0073	0.463 ± 0.014	0.779 ± 0.025
v_2	0.1206 ± 0.0048	0.433 ± 0.014	0.702 ± 0.023	1.030 ± 0.036
v_3	0.493 ± 0.015	0.757 ± 0.024	0.937 ± 0.032	1.307 ± 0.047
v_4	0.749 ± 0.021	0.955 ± 0.029	1.127 ± 0.037	1.344 ± 0.048
v_5	0.902 ± 0.025	1.079 ± 0.033	1.162 ± 0.038	1.432 ± 0.051
v_6	0.981 ± 0.027	1.033 ± 0.032	1.211 ± 0.039	1.389 ± 0.049
v_7	0.993 ± 0.027	1.114 ± 0.033	1.203 ± 0.039	1.406 ± 0.050
v_8	1.062 ± 0.025	1.142 ± 0.030	1.215 ± 0.035	1.403 ± 0.045
v_9	1.033 ± 0.038	1.112 ± 0.045	1.245 ± 0.053	1.365 ± 0.063
A_D	0.0092 ± 0.0033	0.0140 ± 0.0035	0.0008 ± 0.0037	0.0098 ± 0.0040
Γ	0.6587 ± 0.0017	0.6587 ± 0.0017	0.6588 ± 0.0017	0.6588 ± 0.0017
A_P	-0.0052 ± 0.0053	-0.0041 ± 0.0053	-0.0158 ± 0.0054	-0.0013 ± 0.0056
S_f	0.032 ± 0.039	0.074 ± 0.040	0.088 ± 0.038	0.046 ± 0.036
$S_{\bar{f}}$	0.008 ± 0.039	0.041 ± 0.041	0.066 ± 0.039	0.025 ± 0.036
Δm	0.5048 ± 0.0023	0.5054 ± 0.0023	0.5056 ± 0.0023	0.5045 ± 0.0023
$\Delta p_0 SS$	-0.24 ± 0.13	-0.075 ± 0.087	-0.047 ± 0.080	-0.091 ± 0.074
$\Delta p_0 OS$	-0.094 ± 0.099	-0.104 ± 0.099	-0.064 ± 0.090	-0.062 ± 0.082
$\Delta p_1 SS$	-0.03 ± 0.11	-0.041 ± 0.071	0.047 ± 0.059	0.073 ± 0.053
$\Delta p_1 OS$	0.074 ± 0.066	0.183 ± 0.073	0.218 ± 0.071	0.078 ± 0.064
$\Delta p_2 OS$	-0.040 ± 0.022	-0.012 ± 0.027	-0.042 ± 0.032	-0.001 ± 0.023
$\Delta p_3 OS$	-0.20 ± 0.27	-0.58 ± 0.37	-0.08 ± 0.58	-0.00 ± 0.29
$\Delta p_4 OS$	-0.0 ± 1.1	-3.2 ± 1.9	-2.8 ± 2.6	0.4 ± 1.2
$p_0 SS$	-0.198 ± 0.084	-0.223 ± 0.044	-0.100 ± 0.038	0.146 ± 0.032
$p_0 OS$	-0.088 ± 0.041	-0.123 ± 0.040	-0.149 ± 0.039	-0.237 ± 0.037
$p_1 SS$	-0.142 ± 0.081	-0.045 ± 0.051	-0.034 ± 0.040	-0.014 ± 0.033
$p_1 OS$	-0.008 ± 0.045	0.044 ± 0.048	-0.052 ± 0.046	-0.117 ± 0.042
$p_2 OS$	0.006 ± 0.016	-0.027 ± 0.019	-0.005 ± 0.021	-0.004 ± 0.016
$p_3 OS$	-0.60 ± 0.20	-0.44 ± 0.23	0.06 ± 0.31	-0.19 ± 0.20
$p_4 OS$	-1.61 ± 0.83	-0.9 ± 1.3	-0.8 ± 1.5	0.07 ± 0.81
$\varepsilon_{tag} OS$	0.4461 ± 0.0015	0.4210 ± 0.0014	0.4176 ± 0.0014	0.4452 ± 0.0015
$\varepsilon_{tag} SS$	0.92426 ± 0.00078	0.94508 ± 0.00066	0.93922 ± 0.00070	0.91274 ± 0.00084

4.16. Decay-time fits in $B^0 \rightarrow D^\mp \pi^\pm$ data subsamples

Table 4.30 – Fitted parameters in the blinded time fits in bins of number of primary vertices. The reported values for S_f and $S_{\bar{f}}$ are all shifted by the same, unknown amount (blinded).

Parameter	1	2	> 2
v_1	0.3144 ± 0.0084	0.3265 ± 0.0082	0.3150 ± 0.0090
v_2	0.484 ± 0.014	0.491 ± 0.014	0.510 ± 0.016
v_3	0.754 ± 0.022	0.802 ± 0.022	0.829 ± 0.026
v_4	0.935 ± 0.026	1.012 ± 0.026	1.040 ± 0.031
v_5	1.027 ± 0.029	1.093 ± 0.029	1.170 ± 0.035
v_6	1.038 ± 0.029	1.121 ± 0.029	1.204 ± 0.035
v_7	1.087 ± 0.030	1.144 ± 0.029	1.197 ± 0.035
v_8	1.103 ± 0.027	1.173 ± 0.026	1.262 ± 0.032
v_9	1.118 ± 0.041	1.150 ± 0.038	1.203 ± 0.046
A_D	0.0064 ± 0.0032	0.0053 ± 0.0029	0.0152 ± 0.0033
Γ	0.6587 ± 0.0017	0.6587 ± 0.0017	0.6588 ± 0.0017
A_P	-0.0055 ± 0.0047	-0.0112 ± 0.0043	-0.0009 ± 0.0049
S_f	0.007 ± 0.034	0.061 ± 0.031	0.118 ± 0.035
$S_{\bar{f}}$	0.027 ± 0.033	0.016 ± 0.031	0.080 ± 0.036
Δm	0.5052 ± 0.0023	0.5051 ± 0.0023	0.5052 ± 0.0023
$\Delta p_0 SS$	-0.102 ± 0.072	-0.175 ± 0.066	0.065 ± 0.078
$\Delta p_0 OS$	-0.189 ± 0.079	-0.122 ± 0.074	0.120 ± 0.087
$\Delta p_1 SS$	-0.009 ± 0.056	0.046 ± 0.049	0.094 ± 0.057
$\Delta p_1 OS$	0.119 ± 0.065	0.088 ± 0.053	0.273 ± 0.065
$\Delta p_2 OS$	-0.055 ± 0.026	-0.020 ± 0.019	0.000 ± 0.023
$\Delta p_3 OS$	0.16 ± 0.28	-0.53 ± 0.26	-0.17 ± 0.30
$\Delta p_4 OS$	1.9 ± 1.3	-2.6 ± 1.4	-1.2 ± 1.6
$p_0 SS$	-0.031 ± 0.035	-0.057 ± 0.032	-0.031 ± 0.037
$p_0 OS$	-0.193 ± 0.034	-0.177 ± 0.032	-0.067 ± 0.037
$p_1 SS$	0.020 ± 0.038	-0.034 ± 0.034	-0.015 ± 0.039
$p_1 OS$	-0.011 ± 0.043	-0.033 ± 0.035	-0.042 ± 0.042
$p_2 OS$	-0.034 ± 0.018	0.011 ± 0.013	-0.002 ± 0.016
$p_3 OS$	-0.37 ± 0.20	-0.35 ± 0.17	-0.11 ± 0.20
$p_4 OS$	-0.05 ± 0.93	-1.25 ± 0.97	-0.4 ± 1.1
$\varepsilon_{tag} OS$	0.4394 ± 0.0013	0.4330 ± 0.0012	0.4240 ± 0.0013
$\varepsilon_{tag} SS$	0.93394 ± 0.00065	0.93134 ± 0.00059	0.92557 ± 0.00071

Chapter 4. Appendix

Table 4.31 – Fitted parameters in the blinded time fits in bins of the logarithm of the number of tracks. The reported values for S_f and $S_{\bar{f}}$ are all shifted by the same, unknown amount (blinded).

Parameter	[0, 4.8)	[4.8, 5.2)	[5.2, $+\infty$)
v_1	0.2904 ± 0.0075	0.3406 ± 0.0091	0.3303 ± 0.0089
v_2	0.450 ± 0.013	0.521 ± 0.015	0.518 ± 0.016
v_3	0.728 ± 0.021	0.814 ± 0.024	0.847 ± 0.025
v_4	0.932 ± 0.025	1.010 ± 0.028	1.051 ± 0.030
v_5	1.008 ± 0.027	1.116 ± 0.031	1.167 ± 0.033
v_6	1.033 ± 0.028	1.144 ± 0.032	1.186 ± 0.033
v_7	1.066 ± 0.028	1.140 ± 0.032	1.226 ± 0.034
v_8	1.083 ± 0.025	1.219 ± 0.029	1.235 ± 0.030
v_9	1.135 ± 0.039	1.123 ± 0.042	1.211 ± 0.044
A_D	0.0092 ± 0.0031	0.0073 ± 0.0031	0.0091 ± 0.0031
Γ	0.6587 ± 0.0017	0.6588 ± 0.0017	0.6588 ± 0.0017
A_P	-0.0031 ± 0.0046	-0.0071 ± 0.0046	-0.0093 ± 0.0047
S_f	0.036 ± 0.029	0.038 ± 0.034	0.116 ± 0.039
$S_{\bar{f}}$	0.046 ± 0.029	0.000 ± 0.034	0.066 ± 0.039
Δm	0.5056 ± 0.0023	0.5049 ± 0.0023	0.5050 ± 0.0023
$\Delta p_0 SS$	-0.100 ± 0.063	-0.057 ± 0.073	-0.089 ± 0.083
$\Delta p_0 OS$	-0.077 ± 0.071	-0.190 ± 0.081	0.056 ± 0.091
$\Delta p_1 SS$	0.003 ± 0.047	0.074 ± 0.054	0.069 ± 0.063
$\Delta p_1 OS$	0.105 ± 0.062	0.099 ± 0.059	0.239 ± 0.065
$\Delta p_2 OS$	-0.030 ± 0.031	-0.042 ± 0.021	-0.010 ± 0.023
$\Delta p_3 OS$	0.26 ± 0.59	-0.11 ± 0.26	-0.71 ± 0.29
$\Delta p_4 OS$	-1.6 ± 2.3	1.8 ± 1.4	-3.8 ± 1.5
$p_0 SS$	0.015 ± 0.031	-0.028 ± 0.035	-0.126 ± 0.039
$p_0 OS$	-0.077 ± 0.031	-0.174 ± 0.035	-0.221 ± 0.038
$p_1 SS$	-0.038 ± 0.032	-0.037 ± 0.037	0.028 ± 0.044
$p_1 OS$	-0.050 ± 0.041	-0.038 ± 0.038	-0.003 ± 0.042
$p_2 OS$	-0.047 ± 0.020	0.013 ± 0.015	0.018 ± 0.016
$p_3 OS$	-0.15 ± 0.28	-0.27 ± 0.19	-0.22 ± 0.20
$p_4 OS$	-0.3 ± 1.4	-1.33 ± 0.94	-1.13 ± 0.97
$\varepsilon_{tag} OS$	0.4082 ± 0.0012	0.4370 ± 0.0013	0.4523 ± 0.0013
$\varepsilon_{tag} SS$	0.90933 ± 0.00073	0.93877 ± 0.00061	0.94361 ± 0.00059

4.16. Decay-time fits in $B^0 \rightarrow D^\mp \pi^\pm$ data subsamples

Table 4.32 – Fitted parameters in the blinded time fits in bins of the difference in pseudorapidity between the D meson and the bachelor pion. The reported values for S_f and $S_{\bar{f}}$ are all shifted by the same, unknown amount (blinded).

Parameter	$(-\infty, -0.36)$	$[-0.36, 0.10)$	$[0.10, 0.53)$	$[0.53, +\infty)$
v_1	0.2796 ± 0.0082	0.377 ± 0.011	0.396 ± 0.012	0.2429 ± 0.0072
v_2	0.439 ± 0.015	0.565 ± 0.019	0.591 ± 0.020	0.404 ± 0.013
v_3	0.768 ± 0.025	0.872 ± 0.029	0.879 ± 0.030	0.679 ± 0.022
v_4	0.882 ± 0.028	1.061 ± 0.034	1.105 ± 0.035	0.952 ± 0.029
v_5	1.078 ± 0.033	1.152 ± 0.037	1.180 ± 0.038	0.982 ± 0.031
v_6	1.009 ± 0.031	1.181 ± 0.037	1.202 ± 0.038	1.095 ± 0.033
v_7	1.112 ± 0.034	1.171 ± 0.037	1.208 ± 0.038	1.083 ± 0.033
v_8	1.095 ± 0.029	1.222 ± 0.034	1.200 ± 0.034	1.191 ± 0.031
v_9	1.130 ± 0.046	1.128 ± 0.048	1.312 ± 0.052	1.065 ± 0.045
A_D	0.0072 ± 0.0036	0.0043 ± 0.0037	0.0118 ± 0.0036	0.0106 ± 0.0035
Γ	0.6587 ± 0.0017	0.6587 ± 0.0017	0.6588 ± 0.0017	0.6588 ± 0.0017
A_P	-0.0092 ± 0.0053	-0.0121 ± 0.0054	0.0004 ± 0.0053	-0.0051 ± 0.0053
S_f	0.047 ± 0.038	0.050 ± 0.039	0.103 ± 0.038	0.034 ± 0.039
$S_{\bar{f}}$	0.028 ± 0.038	0.060 ± 0.039	0.006 ± 0.038	0.054 ± 0.039
Δm	0.5052 ± 0.0023	0.5047 ± 0.0023	0.5046 ± 0.0023	0.5059 ± 0.0023
$\Delta p_0 SS$	-0.143 ± 0.083	-0.121 ± 0.082	0.024 ± 0.081	-0.098 ± 0.084
$\Delta p_0 OS$	-0.105 ± 0.092	-0.082 ± 0.093	0.001 ± 0.089	-0.140 ± 0.093
$\Delta p_1 SS$	0.157 ± 0.066	0.033 ± 0.058	-0.038 ± 0.063	0.017 ± 0.063
$\Delta p_1 OS$	0.185 ± 0.074	0.276 ± 0.069	0.029 ± 0.067	0.082 ± 0.067
$\Delta p_2 OS$	-0.021 ± 0.037	-0.054 ± 0.026	0.007 ± 0.025	-0.032 ± 0.022
$\Delta p_3 OS$	0.16 ± 0.79	-0.31 ± 0.30	-0.10 ± 0.31	-0.56 ± 0.28
$\Delta p_4 OS$	-2.5 ± 2.9	-0.1 ± 1.3	-0.4 ± 1.3	-1.2 ± 1.3
$p_0 SS$	-0.061 ± 0.040	-0.009 ± 0.038	-0.029 ± 0.040	-0.071 ± 0.041
$p_0 OS$	-0.127 ± 0.039	-0.145 ± 0.039	-0.157 ± 0.039	-0.188 ± 0.041
$p_1 SS$	0.018 ± 0.044	-0.060 ± 0.039	0.055 ± 0.043	-0.054 ± 0.044
$p_1 OS$	-0.032 ± 0.049	-0.033 ± 0.044	-0.044 ± 0.045	-0.010 ± 0.045
$p_2 OS$	-0.037 ± 0.023	0.011 ± 0.017	0.003 ± 0.018	-0.016 ± 0.015
$p_3 OS$	-0.01 ± 0.32	-0.12 ± 0.22	0.01 ± 0.23	-0.95 ± 0.20
$p_4 OS$	-0.7 ± 2.0	0.16 ± 0.92	0.66 ± 0.96	-2.93 ± 0.91
$\varepsilon_{tag} OS$	0.4370 ± 0.0015	0.4344 ± 0.0015	0.4346 ± 0.0015	0.4236 ± 0.0014
$\varepsilon_{tag} SS$	0.93563 ± 0.00072	0.92611 ± 0.00077	0.92349 ± 0.00078	0.93669 ± 0.00071

Chapter 4. Appendix

Table 4.33 – Fitted parameters in the blinded time fits performed for candidates TOS on L0HadronTOS and all the other candidates. The reported values for S_f and $S_{\bar{f}}$ are all shifted by the same, unknown amount (blinded).

Parameter	L0HadronTOS	Other candidates
v_1	0.4111 ± 0.0092	0.1889 ± 0.0052
v_2	0.608 ± 0.015	0.331 ± 0.010
v_3	0.909 ± 0.022	0.633 ± 0.019
v_4	1.087 ± 0.025	0.863 ± 0.024
v_5	1.174 ± 0.027	0.981 ± 0.028
v_6	1.188 ± 0.027	1.018 ± 0.028
v_7	1.200 ± 0.027	1.051 ± 0.029
v_8	1.234 ± 0.024	1.095 ± 0.026
v_9	1.218 ± 0.034	1.063 ± 0.039
A_D	0.0079 ± 0.0024	0.0093 ± 0.0032
Γ	0.6588 ± 0.0017	0.6587 ± 0.0017
A_P	-0.0092 ± 0.0035	-0.0002 ± 0.0049
S_f	0.029 ± 0.026	0.111 ± 0.034
$S_{\bar{f}}$	0.029 ± 0.026	0.048 ± 0.034
Δm	0.5053 ± 0.0023	0.5049 ± 0.0023
$\Delta p_0 SS$	-0.108 ± 0.053	-0.061 ± 0.083
$\Delta p_0 OS$	-0.145 ± 0.060	0.044 ± 0.084
$\Delta p_1 SS$	0.012 ± 0.038	0.100 ± 0.070
$\Delta p_1 OS$	0.122 ± 0.046	0.151 ± 0.059
$\Delta p_2 OS$	-0.032 ± 0.017	-0.012 ± 0.020
$\Delta p_3 OS$	-0.34 ± 0.25	0.04 ± 0.25
$\Delta p_4 OS$	-1.8 ± 1.3	0.7 ± 1.1
$p_0 SS$	0.013 ± 0.024	-0.229 ± 0.046
$p_0 OS$	-0.196 ± 0.027	-0.083 ± 0.034
$p_1 SS$	-0.024 ± 0.025	-0.029 ± 0.050
$p_1 OS$	-0.066 ± 0.030	-0.005 ± 0.039
$p_2 OS$	-0.006 ± 0.012	-0.010 ± 0.015
$p_3 OS$	-0.20 ± 0.16	-0.46 ± 0.18
$p_4 OS$	-0.66 ± 0.79	-0.70 ± 0.80
$\varepsilon_{tag} OS$	0.40498 ± 0.00094	0.4834 ± 0.0014
$\varepsilon_{tag} SS$	0.93331 ± 0.00048	0.92502 ± 0.00071

¹³¹⁴

4.17 Decay-time fit validation with bootstrapping

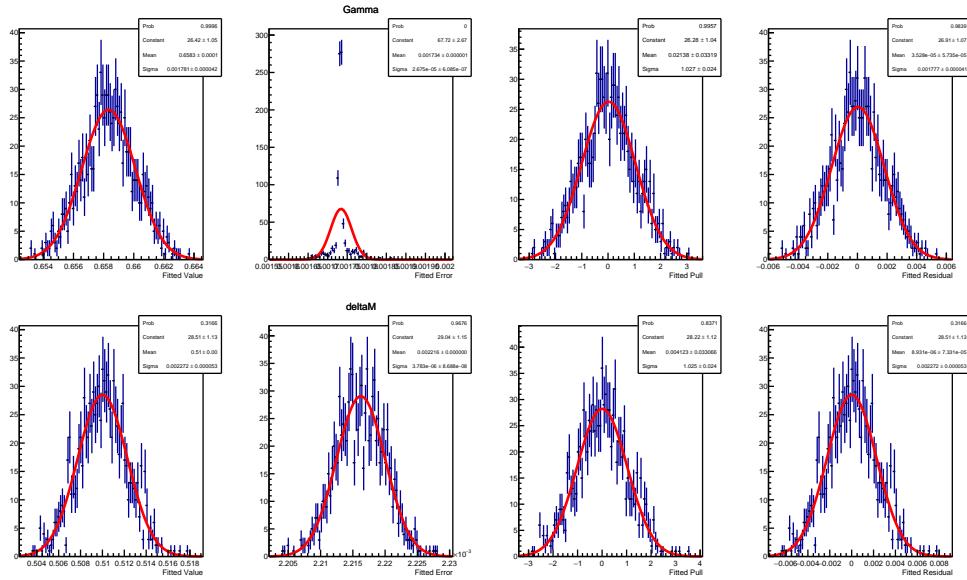
¹³¹⁵ In this appendix, the result of the Monte Carlo time fit validation is shown. The results for
¹³¹⁶ production, detection asymmetry, and acceptance are not shown because no independent
¹³¹⁷ reference value for these parameters exist in Monte Carlo.


Figure 4.55 – Distributions of fitted value, error, pull and residual for Γ (top) and Δm (bottom). Each distribution is fitted with a Gaussian PDF. Pulls and residuals are computed by taking the Monte Carlo generation value as reference.

Chapter 4. Appendix

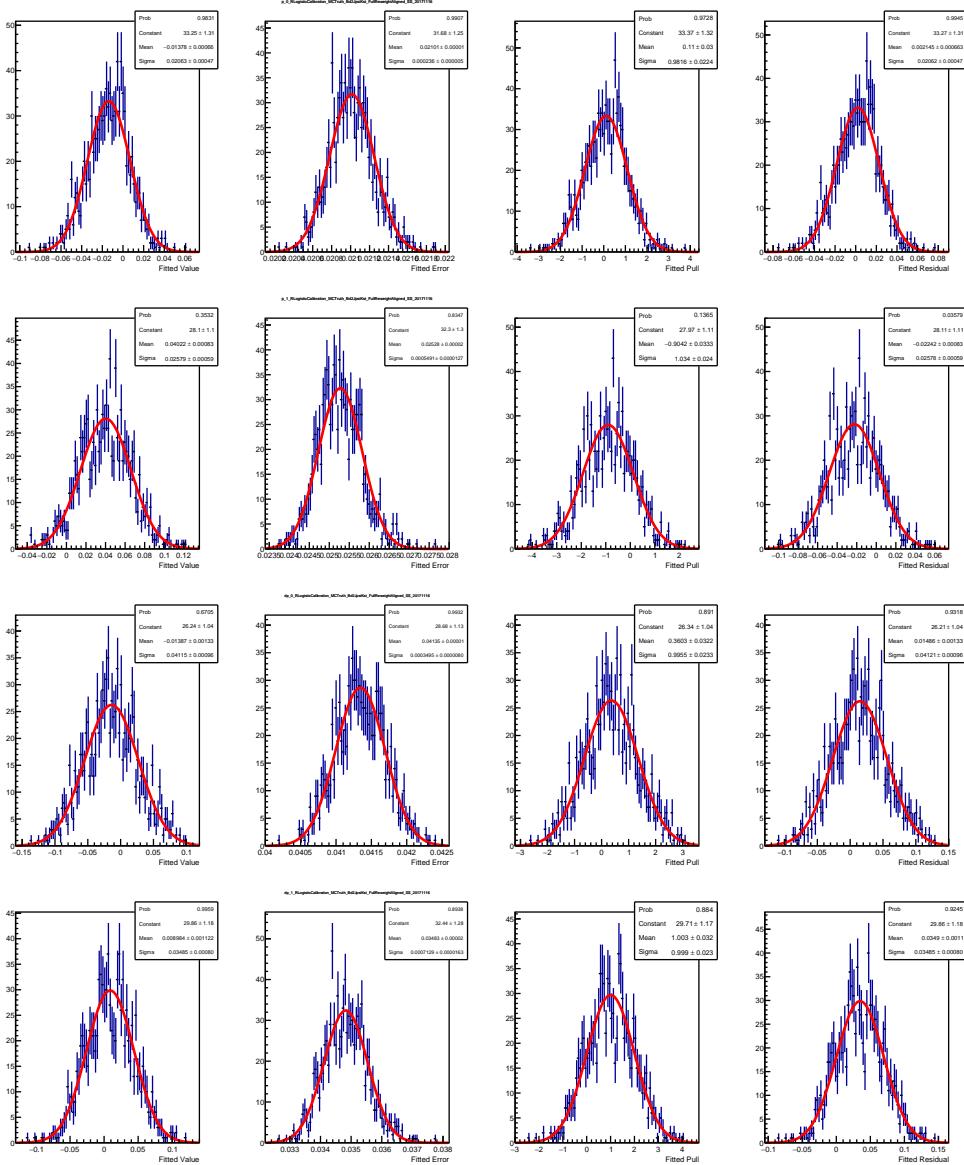


Figure 4.56 – Distributions of fitted value, error, pull and residual for the SS tagger calibration parameters. Each distribution is fitted with a Gaussian PDF. Pulls and residuals are computed by taking the values find on the $B^0 \rightarrow J/\psi K^{*0}$ Monte Carlo calibration as reference.

4.17. Decay-time fit validation with bootstrapping

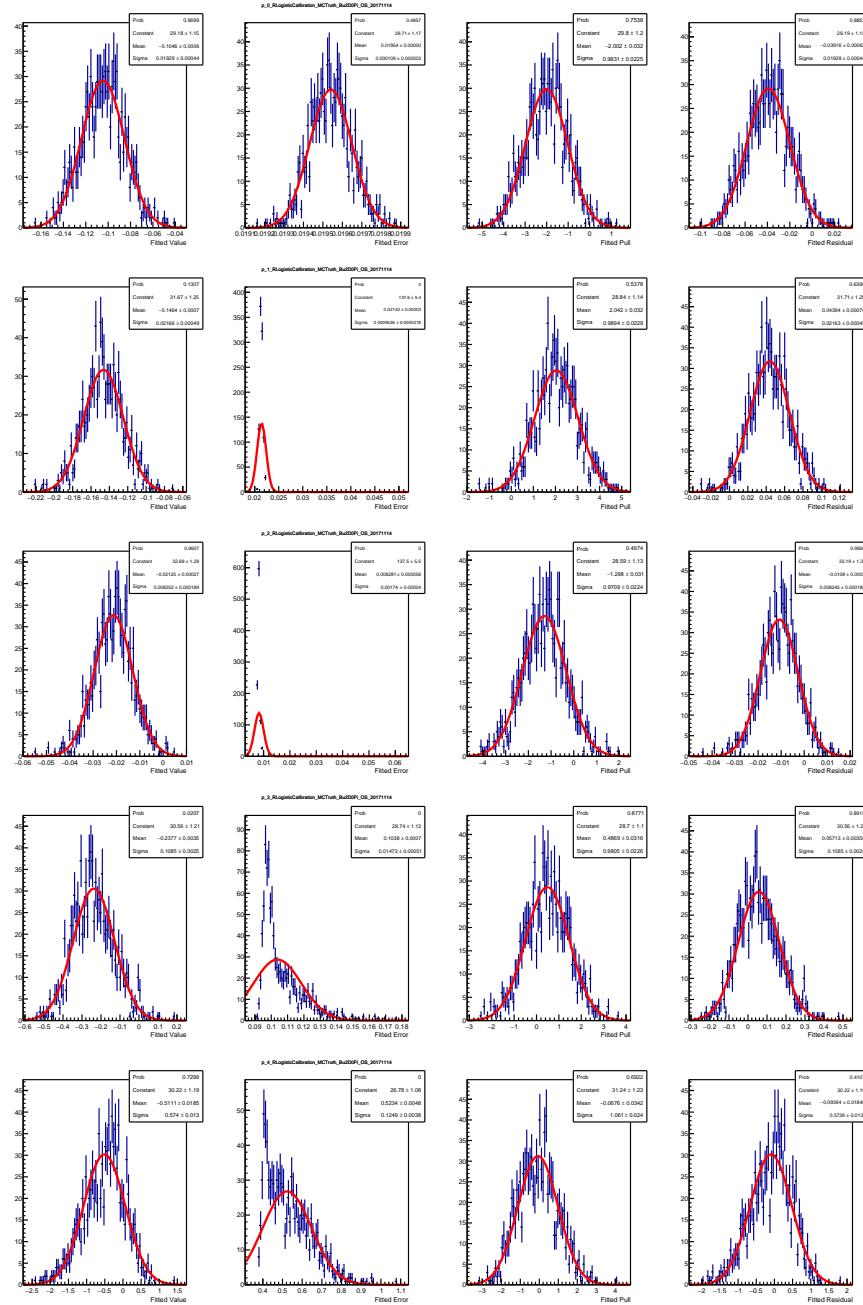


Figure 4.57 – Distributions of fitted value, error, pull and residual for the OS tagger calibration parameters. Each distribution is fitted with a Gaussian PDF. Pulls and residuals are computed by taking the values find on the $B^+ \rightarrow D^0\pi^+$ Monte Carlo calibration as reference.

Chapter 4. Appendix

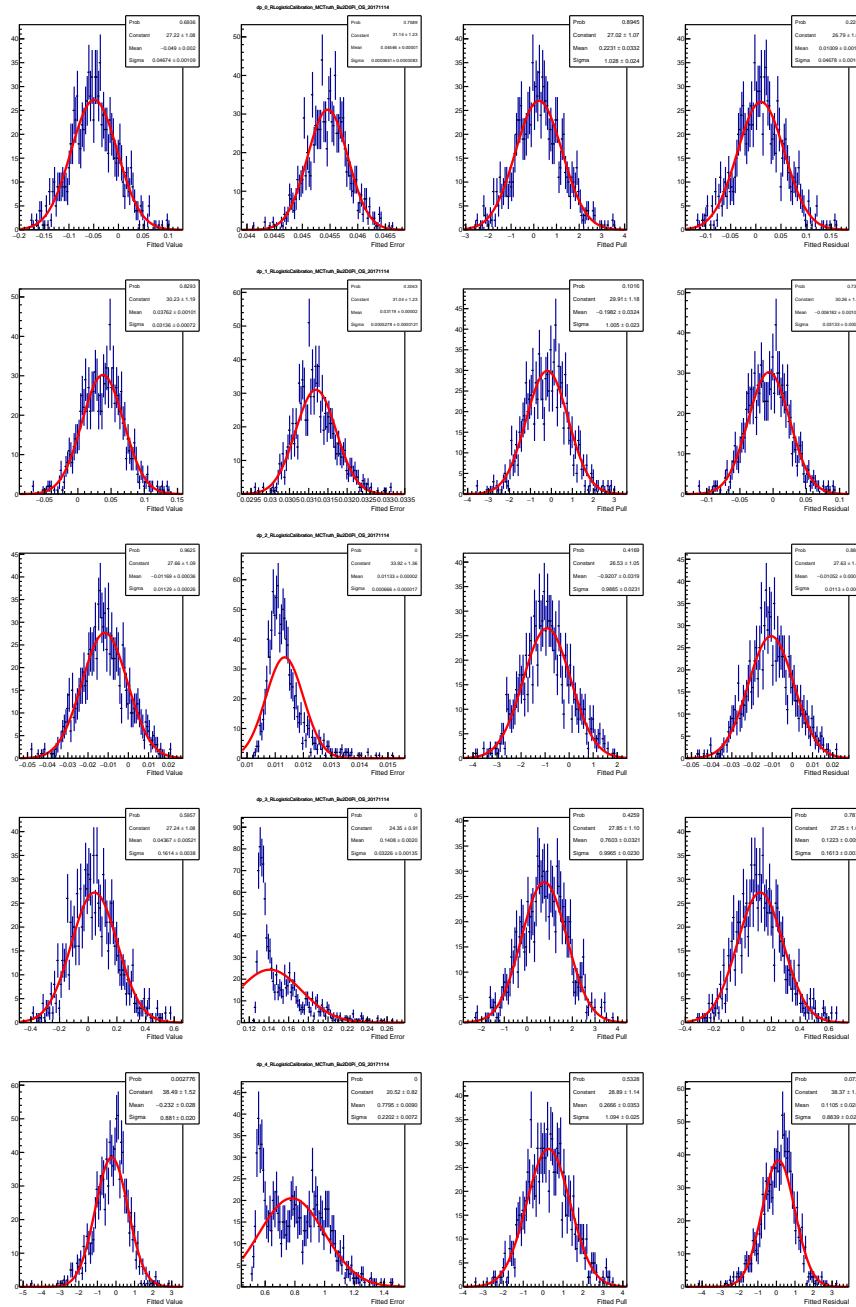


Figure 4.58 – Distributions of fitted value, error, pull and residual for the OS mistag asymmetries. Each distribution is fitted with a Gaussian PDF. Pulls and residuals are computed by taking the values find on the $B^+ \rightarrow D^0\pi^+$ Monte Carlo calibration as reference.

₁₃₁₈ 4.18 Test of the decay-time fit via a toy tagger

The *toy tagger* used to perform the test mentioned in Sec. 4.7.1 is created as follows. First, a mistag (η) PDF is created from the *sWeighted* η distribution of the OS tagger on data. This template is created as a `RooHistPdf`. Then, for each candidate in each bootstrapped Monte Carlo sample, a value of η is drawn from this PDF. The decision of the toy tagger is taken from the true ID of the B^0 meson, which is always correct by definition. In order to emulate wrong tagging decisions, a random number r_i is generated for the i^{th} B^0 candidate between 0 and 1. If η_i is the mistag assigned to this candidate, the tagging decision d_i is flipped (and thus made wrong) according to the following criterion:

$$d_i = \begin{cases} -d_i & \text{if } r_i \leq \eta_i \\ d_i & \text{otherwise} \end{cases}. \quad (4.33)$$

₁₃₁₉ During the time fit, the mistag calibration is simply taken as a linear function (Eq. 3.9)
₁₃₂₀ with $p_0 = \langle \eta \rangle = 0.370029$ (taken from the adopted template) and $p_1 = 1$, which means
₁₃₂₁ $\omega = \eta$ for all candidates. In fact, the per-event mistag η is the true mistag ω probability by
₁₃₂₂ construction. In this way, it is possible to test the time fit with a per-event mistag without
₁₃₂₃ relying on any approximation or uncertainty coming from the calibration procedure.
₁₃₂₄ Moreover, the tagging efficiency is 100% by construction.

₁₃₂₅ The distributions of the fitted value, error, pull and residual for the relevant parameters are
₁₃₂₆ shown in Fig. 4.59. Each distribution is fitted with a Gaussian PDF. All pull distributions
₁₃₂₇ have means compatible with 0 and widths compatible with 1, meaning that the maximum
₁₃₂₈ likelihood estimation of the parameters is unbiased and returns correct uncertainties.

Chapter 4. Appendix

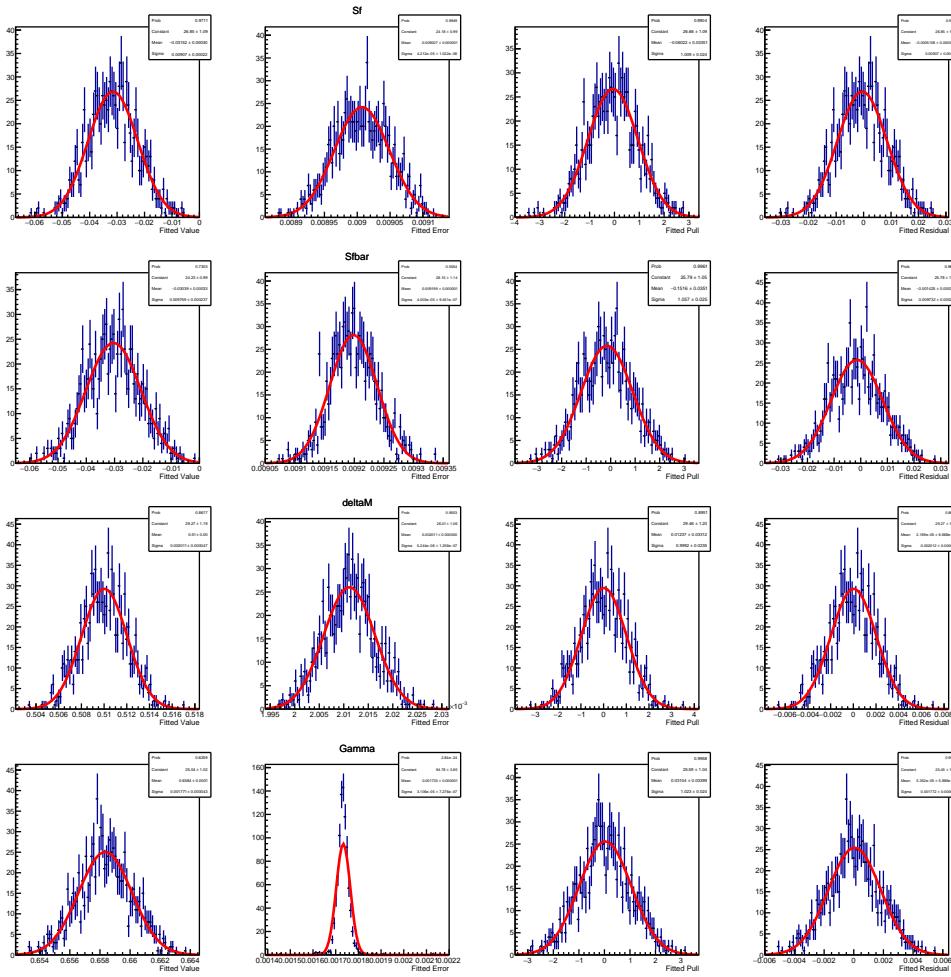


Figure 4.59 – Distributions of fitted value, error, pull and residual for the main parameters fitted on bootstrapped Monte Carlo samples with a toy tagger. Each distribution is fitted with a Gaussian PDF. Pulls and residuals are computed by taking the Monte Carlo generation value as reference.

¹³²⁹ **4.19 Inputs for $B^0 \rightarrow D^\pm \pi^\mp$ simulation**
¹³³⁰ In table 4.34, the input values used for the generation of the Monte Carlo sample are
¹³³¹ reported. The resulting values of the CP coefficients are listed as well.

 Table 4.34 – Parameter values used in the generation of the $B^0 \rightarrow D^\pm \pi^\mp$ MC sample, and resulting values for the CP coefficients.

Parameter	Generation value
Δm	0.51 ps^{-1}
τ	$1.519\,068 \text{ ps}$
$\Delta\Gamma/\Gamma$	0
$ q/p $	1
$\arg(q/p)$	-0.764
$ A_f $	0.0849
$\arg(A_f)$	0.002 278
$ \bar{A}_f $	0.001 37
$\arg(\bar{A}_f)$	-1.128 958
$ A_{\bar{f}} $	0.001 37
$\arg(A_{\bar{f}})$	1.3145
$ \bar{A}_{\bar{f}} $	0.0849
$\arg(\bar{A}_{\bar{f}})$	0.002 278
S_f	-0.0305
$S_{\bar{f}}$	-0.0282
C_f	0.9995
$C_{\bar{f}}$	-0.9995
D_f	-0.0103
$D_{\bar{f}}$	-0.0155

Bibliography

- 1333 [1] A. Salam, *Weak and electromagnetic interactions*, Proceedings of the 8th Nobel Symposium on Elementary particle theory, relativistic groups and analyticity, Stockholm, Sweden (1968).
- 1336 [2] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1968) 1264.
- 1337 [3] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.
- 1338 [4] <http://www-f9.ijs.si/%7Elubej/SM.pdf>.
- 1339 [5] E. Fermi, *Sulla quantizzazione del gas perfetto monoatomico*, Rendiconti Lincei **3** (1926) 145.
- 1341 [6] P. A. M. Dirac, *On the Theory of Quantum Mechanics*, Proceedings of the Royal Society A. **112 (762)** (1926) 661.
- 1343 [7] S. Bose, *Plancks Gesetz und Lichtquantenhypothese*, Zeitschrift für Physik **26** (1924) 178.
- 1345 [8] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the Z(4430)⁻ state*, Phys. Rev. Lett. **112** (2014) 222002, [arXiv:1404.1903](https://arxiv.org/abs/1404.1903).
- 1347 [9] LHCb collaboration, R. Aaij *et al.*, *Observation of J/ψ p resonances consistent with pentaquark states in Λ_b⁰ → J/ψ p K⁻ decays*, Phys. Rev. Lett. **115** (2015) 072001, [arXiv:1507.03414](https://arxiv.org/abs/1507.03414).
- 1350 [10] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Bosons*, Phys. Rev. Lett. **13** (1964) 321.
- 1352 [11] P. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.
- 1354 [12] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.
- 1356 [13] The ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1.

Bibliography

- 1359 [14] The CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the*
1360 *CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30.
- 1361 [15] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **178** (1963)
1362 531.
- 1363 [16] M. Kobayashi and T. Maskawa, *CP violation in the Renormalizable Theory of Weak*
1364 *Interaction*, Progr. Theor. Phys. **49** (1973) 652.
- 1365 [17] L. L. Chau and W. Y. Keung, *Comments on the Parametrization of the Kobayashi-*
1366 *Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) 1802.
- 1367 [18] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Progr. Theor.
1368 Phys. **49** (1973) 652.
- 1369 [19] CKMfitter group, J. Charles *et al.*, *CP Violation and the CKM Matrix: Assessing*
1370 *the Impact of the Asymmetric B Factories*, Eur. Phys. J. **C41** (2005) 1,
1371 updated results and plots available at <http://ckmfitter.in2p3.fr>.
- 1372 [20] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. C.
1373 **C40** (2016) 100001.
- 1374 [21] I. Bigi and A. I. Sanda, *CP Violation*, Cambridge Monographs on Particle Physics,
1375 Nuclear Physics and Cosmology, 2009.
- 1376 [22] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys.
1377 **A30** (2015) 1530022.
- 1378 [23] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007.
- 1379 [24] LHCb collaboration, R. Antunes Nobrega *et al.*, *LHCb reoptimized detector design and*
1380 *performance: Technical Design Report*, CERN-LHCC-2003-030. LHCb-TDR-009.
- 1381 [25] LHCb collaboration, A. Franca Barbosa *et al.*, *LHCb inner tracker: Technical Design*
1382 *Report*, CERN-LHCC-2002-029. LHCb-TDR-008.
- 1383 [26] LHCb collaboration, P. R. Barbosa Marinho *et al.*, *LHCb outer tracker: Technical*
1384 *Design Report*, CERN-LHCC-2001-024. LHCb-TDR-006.
- 1385 [27] V. Battista, F. Blanc, M. Martinelli, and M. Tobin, *A study of spillover clusters and*
1386 *ghost tracks in the Silicon Tracker with 25 ns bunch spacing*, CERN-LHCB-INT-
1387 2016-010.
- 1388 [28] A. Powell *et al.*, *Particle identification at LHCb*, Proceedings of ICHEP2010 **020**
1389 (2010), LHCb-PROC-2011-008.
- 1390 [29] LHCb collaboration, S. Amato *et al.*, *LHCb calorimeters: Technical Design Report*,
1391 CERN-LHCC-2000-036.

- 1392 [30] A. A. Alves Jr *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013)
1393 P02022.
- 1394 [31] LHCb collaboration, R. Antunes Nobrega *et al.*, *LHCb trigger system: Technical*
- 1395 *Design Report*, CERN-LHCC-2003-031.
- 1396 [32] S. Torbjörn *et al.*, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**
1397 (2015) 159.
- 1398 [33] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth.
1399 **A462** (2001) 152.
- 1400 [34] P. Gononka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections*
1401 *in Z and W decays*, Eur. Phys. J. **C45** (2006) 97.
- 1402 [35] S. Agostinelli *et al.*, *Geant4: a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003)
1403 250.
- 1404 [36] D. Fazzini, *Development of "same side" flavour tagging algorithms for measurements*
1405 *of flavour oscillations and CP violation in the B⁰ mesons system*, CERN-THESIS-
1406 2015-040.
- 1407 [37] LHCb collaboration, R. Aaij *et al.*, *New algorithms to tag the flavour of B⁰ mesons*
1408 *using pions and protons*, Eur. Phys. J. (2016), no. C77 [arXiv:1610.06019](#).
- 1409 [38] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the*
1410 *LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, [arXiv:1202.4979](#).
- 1411 [39] LHCb collaboration, R. Aaij *et al.*, *B flavour tagging using charm decays at the*
1412 *LHCb experiment*, JINST **10** (2015) P10005, [arXiv:1507.07892](#).
- 1413 [40] M. Grabalosa, *Flavour Tagging developments within the LHCb experiment*, PhD
1414 thesis, Barcelona U., 2012-03-28.
- 1415 [41] LHCb collaboration, *Optimization and calibration of the LHCb flavour tagging*
1416 *performance using 2010 data*, LHCb-CONF-2011-003.
- 1417 [42] V. Battista, *b-flavour tagging in pp collisions at LHCb*, Il Nuovo Cimento **39C** (2016)
1418 335.
- 1419 [43] J. Wimberley, *Calibration flavor tagging algorithms with binomial regression*, LHCb-
1420 INT-2017-002.
- 1421 [44] I. Dunietz and R. G. Sachs, *Asymmetry between inclusive charmed and anticharmed*
1422 *modes in B⁰, \bar{B}^0 decay as a measure of CP violation*, Phys. Rev. **D37** (1988) 3186.
- 1423 [45] R. Aleksan, I. Dunietz, and B. Kayser, *Determining the CP violating phase gamma*,
1424 Z Phys. **C54** (1992) 653.

Bibliography

- [46] R. Fleischer, *New strategies to obtain insights into CP violation through $B_{(s)} \rightarrow D_{(s)}^\pm K^\mp$, $D_{(s)}^{*\pm} K^\mp$, ... and $B_{(d)} \rightarrow D^\pm \pi^\mp$, $D^{*\pm} \pi^\mp$, ... decays*, Nucl. Phys. **B671** (2003) 459.
- [47] LHCb collaboration, R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetries in $B_s^0 \rightarrow J/\psi K_s^0$* , JHEP **06** (2015) 131, [arXiv:1503.07055](#).
- [48] BaBar collaboration, B. Aubert *et al.*, *Measurement of the Branching Fractions of the Rare Decays $B^0 \rightarrow D_s^{*+} \pi^-$, $B^0 \rightarrow D_s^{*+} \rho^-$, and $B^0 \rightarrow D_s^{*-} K^{*+}$* , Phys. Rev. **D78** (2008) 032005.
- [49] Belle collaboration, A. Das *et al.*, *Measurements of Branching Fractions for $B^0 \rightarrow D_s^+ \pi^-$ and $\bar{B}^0 \rightarrow D_s^+ K^-$* , Phys. Rev. **D82** (2010) 051103.
- [50] BaBar collaboration, B. Aubert *et al.*, *Measurement of time-dependent CP-violating asymmetries and constraints on $\sin(2\beta + \gamma)$ with partial reconstruction of $B \rightarrow D^{*\mp} \pi^\pm$ decays*, Phys. Rev. **D71** (2005) 112003.
- [51] BaBar collaboration, B. Aubert *et al.*, *Measurement of time-dependent CP asymmetries in $B^0 \rightarrow D^{*\pm} \pi^\mp$ and $B^0 \rightarrow D^\pm \rho^\mp$ decays*, Phys. Rev. **D76** (2006) 111101.
- [52] Belle collaboration, F. J. Ronga *et al.*, *Measurement of CP violation in $B^0 \rightarrow D^{*-} \pi^+$ and $B^0 \rightarrow D^- \pi^+$ decays*, Phys. Rev. **D73** (2006) 092003.
- [53] Belle collaboration, S. Bahinipati *et al.*, *Measurements of time-dependent CP asymmetries in $B^0 \rightarrow D^{*\mp} \pi^\pm$ decays using a partial reconstruction technique*, Phys. Rev. **D84** (2011) 021101.
- [54] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \rightarrow D_s^\mp K^\pm$ decays*, JHEP **11** (2014) 060.
- [55] L. Breiman, *Bagging Predictors*, Machine Learning **24** (1996) 123.
- [56] LHCb collaboration, *A single track HLT1 trigger*, LHCb-PUB-2011-003.
- [57] LHCb collaboration, *The LHCb inclusive B triggers*, LHCb-INT-2011-030.
- [58] W. D. Hulsbergen *et al.*, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. (2005), no. A552 566.
- [59] P. Koppenburg, *Statistical biases in measurements with multiple candidates*, [arXiv:1703.01128](#).
- [60] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.
- [61] 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](#).

- 1459 [62] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007)
1460 040, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- 1461 [63] R. E. Schapire and Y. Freund, *A decision-theoretic generalization of on-line learning
1462 and an application to boosting*, Jour. Comp. and Syst. Sc. **55** (1997) 119.
- 1463 [64] M. Pivk and F. R. L. Diberder, *sPlot: a statistical tool to unfold data distributions*,
1464 Nucl. Instrum. Meth. **A555** (2005) 356.
- 1465 [65] D. M. Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*,
1466 Nucl. Instrum. Meth. (2014), no. A764 150.
- 1467 [66] N. L. Johnson, *Systems of Frequency Curves Generated by Methods of Translation*,
1468 Biometrika (1949), no. 36 149.
- 1469 [67] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime
1470 and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986,
1471 DESY-F31-86-02.
- 1472 [68] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s in
1473 $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$ decays*, Phys. Rev. Lett. **113** (2014) 211801, [arXiv:1409.4619](https://arxiv.org/abs/1409.4619).
- 1474 [69] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of CP violation in
1475 $B_s^0 \rightarrow J/\psi K^+ K^-$ decays*, Phys. Rev. Lett. **114** (2015) 041801, [arXiv:1411.3104](https://arxiv.org/abs/1411.3104).
- 1476 [70] M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing
1477 and their efficient evaluation*, Nucl. Instrum. Meth. [arXiv:1407.0748](https://arxiv.org/abs/1407.0748).
- 1478 [71] Y. Xie, *sFit: a method for background subtraction in maximum likelihood fit*,
1479 [arXiv:0905.0724](https://arxiv.org/abs/0905.0724).
- 1480 [72] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38**
1481 (2014) 090001.
- 1482 [73] LHCb collaboration, R. Aaij *et al.*, *A precise measurement of the B^0 meson oscillation
1483 frequency*, Eur. Phys. J. **C76** (2016) 412, [arXiv:1604.03475](https://arxiv.org/abs/1604.03475).
- 1484 [74] LHCb collaboration, R. Aaij *et al.*, *Measurement of the \bar{B}^0 - B^0 and \bar{B}_s^0 - B_s^0 pro-
1485 duction asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739** (2014) 218,
1486 [arXiv:1408.0275](https://arxiv.org/abs/1408.0275).
- 1487 [75] P. Good, *Permutation, Parametric and Bootstrap Tests of Hypotheses*, Springer
1488 Series in Statistics, 2004.
- 1489 [76] BaBar, B. Aubert *et al.*, *Measurement of the Branching Fractions of the Rare Decays
1490 $B^0 \rightarrow D_s (*) + \pi^-$, $B^0 \rightarrow D_s (*) + \rho^-$, and $B^0 \rightarrow D_s (*) - K^{(*)} +$* , Phys. Rev.
1491 **D78** (2008) 032005, [arXiv:0803.4296](https://arxiv.org/abs/0803.4296).

Bibliography

- 1492 [77] Belle, A. Das *et al.*, *Measurements of Branching Fractions for $B^0 \rightarrow D_s^+ \pi^-$ and*
- 1493 $\bar{B}^0 \rightarrow D_s^+ K^-$, Phys. Rev. **D82** (2010) 051103, [arXiv:1007.4619](https://arxiv.org/abs/1007.4619).
- 1494 [78] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron,*
- 1495 *and τ -lepton properties as of summer 2014*, [arXiv:1412.7515](https://arxiv.org/abs/1412.7515), updated results and
- 1496 plots available at <http://www.slac.stanford.edu/xorg/hfag/>.
- 1497 [79] A. Rogozhnikov, *hep_ml*, https://arogozhnikov.github.io/hep_ml/.