SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Dipartimento di Fisica G. Occhialini

Dottorato di Ricerca in Fisica e Astronomia Ciclo XXX
Curriculum Fisica Subnucleare e Tecnologie Fisiche

Measurement of CP Violation using $B^0 \to D^{(*)\pm} D^{\mp}_{(s)}$ decays at LHCb

Cognome: Belloli **Nome: Nicoletta**

Matricola : 798762

Tutore: Prof. Marta Calvi

Coordinatore: Prof. Marta Calvi

Anno Accademico 2016/2017

Abstract

This thesis presents two measurements, performed using data collected by the LHCb experiment, operating at the Large Hadron Collider accelerator at CERN.

The first is the measurement of the CP violation observables S and C in the decays of B^0 and $\bar{B}^0$ mesons to the D^+D^- final state exploiting a data sample corresponding to an integrated luminosity of 3 fb^{-1} collected in proton-proton collision at center of mass energies of 7-8 TeV, and resulting in

$$S_{D^+D^-} = -0.54^{+0.17}_{-0.16}(stat) \pm 0.05(syst)$$

$$C_{D^+D^-} = 0.26^{+0.18}_{-0.17}(stat) \pm 0.02(syst)$$

This result combined with the constraint coming from $B^0 \to J/\psi K^0_s$ measurements, allows to constrain the phase shift $\Delta\phi$ to world's most precise value of $\Delta\phi$=-0.16$^{+0.19}_{-0.21}$ rad, implying only small contribution from higher-order Standard Model corrections.

The second aims to measure the CP violation observables S_{D^*D}, C_{D^*D}, ΔS_{D^*D} and ΔC_{D^*D} in the decays of B^0 and $\bar{B}^0$ mesons to the $D^{*+}D^-$ final state using a data sample corresponding to an integrated luminosity of 5 fb^{-1} collected at the LHCb experiment in proton-proton collision at center of mass energies of 7-8 TeV and 13-14 TeV, using a flavour-tagged, decay-time dependent analysis. From the preliminary results obtained it is possible to estimate the expected sensitivity for the CP observables to be

$$\sigma_{S_{D^*D}} = 0.09 \qquad \sigma_{\Delta S_{D^*D}} = 0.09$$

Sintesi

In questa tesi viene presentato il lavoro di due analisi, entrambe svolte all'interno della collaborazione LHCb, uno degli esperimenti presenti al Large Hadron Collider del CERN di Ginevra.

La prima analisi riguarda la misura delle osservabili di violazione di CP S e C nei decadimenti di mesoni B^0 e $\overline{B}^0$ nello stato finale D^+D^-. Per l'analisi é stato utilizzato un campione dati corrispondente ad una luminositá integrata di 3 fb^{-1}, raccolto in collisioni protone-protone per energie nel centro di massa di 7-8 TeV rispettivamente. Il risultato ottenuto per le osservabili é il seguente:

$$S_{D^+D^-} = -0.54^{+0.17}_{-0.16}(stat) \pm 0.05(syst)$$

$$C_{D^+D^-} = 0.26^{+0.18}_{-0.17}(stat) \pm 0.02(syst)$$

Questo risultato, combinato con il vincolo proveniente dalle misure nel canale $B^0 \to J/\psi K^0_s$, permette di determinare il valore della differenza di fase $\Delta\phi$ ottenendo il valore piú preciso al mondo, corrispondente a $\Delta\phi$=-0.16$^{+0.19}_{-0.21}$rad. Questo risultato indica piccoli contributi per le correzioni di ordine superiore del Modello Standard.

La seconda analisi presentata ha lo scopo di misurare le osservabili di violazione di CP S_{D^*D}, C_{D^*D}, ΔS_{D^*D} e ΔC_{D^*D} nei decadimenti di mesoni B^0 e $\overline{B}^0$ nello stato finale $D^{*+}D^-$. Per l'analisi viene utilizzato un campione dati corrispondente ad una luminositá integrata di 5 fb^{-1}, raccolto dall'esperimento LHCb in collisioni protone-protone per energie nel centro di massa di 7-8 Tev e 13-14 TeV rispettivamente. Utilizzando i risultati preliminari ottenuti dall'analisi si stima che la sensitivitá attesa per le osservabili di CP é

$$\sigma_{S_{D^*D}} = 0.09 \qquad \sigma_{\Delta S_{D^*D}} = 0.09$$

Preface

The work presented in this thesis was performed inside of the LHCb experiment, a collaboration of more than 800 scientists, and it would not have been possible without the joined effort of many. The Author worked specifically on the analysis of the data after they were collected by the detector and processed by a common software.

The CP violation measurement in the decay channel $B^0 \to D^+D^-$ is the result of the work of a group composed of: M. Calvi, B. Khanji, F. Meier, M. Schellenberg, P. Seyfert, J. Wishahi. The Author specifically worked on the selection and calibration of the flavour tagging using the control channel $B^0 \to D^+D_s^-$.

The measurement of CP violation in $B^0 \to D^{*+}D^-$ is mainly work of the Author, with the collaboration of a group composed of: M. Calvi, P. Ibis, B. Khanji, F. Meier, M. Schellenberg, P. Seyfert, G. Tellarini, S. Vecchi, J. Wishahi.

Ringraziamenti

Ringrazio prima di tutto la mia tutor Prof. Marta Calvi, per avermi dato la possibilitá di lavorare a questo progetto, e per la pazienza e disponibiltá con cui mi ha seguito ed incoraggiato in questo triennio.

Un ringraziamento va anche ai miei colleghi, Dott. Bassem Khanji e Dott. Paul Seyfert e ai colleghi del mio gruppo di lavoro delle Universitá di Dortmund e Ferrara, per tutti i consigli e gli insegnamenti che con pazienza mi hanno dato durante questo periodo.

Un particolare ringraziamento va alla mia famiglia, alle persone a me piú vicine ed al mio collega di universitá Davide, per avermi sopportato e supportato, in questo periodo cosí pieno ed intenso. Li ringrazio per la pazienza, l'aiuto materiale e l'affetto che mi hanno dimostrato.

Contents

Chapter 1

Introduction

The Standard Model (SM), describes elementary particles and their interactions. Even though this model is able to describe, with very high precision, elementary particles interactions, some experimental observations cannot be explained in this framework. The measurement of rotation velocity of spiral galaxies has lead to the need of an extension of the SM in order to account for the presence of dark matter in our universe, since it contributes to around of 86 % of matter in the universe. Another example comes from neutrino oscillations [1], this phenomen imply the neutrinos to have masses, nevertheless in the SM these are described as massless particles, for this reason also in this case an extension of the model is necessary. Finally the matter-antimatter asymmetry observed in the universe is another example for which the extension of the SM is necessary. In 1967 Sakharov established three conditions which should be satisfied in baryon production in order to guarantee that matter and antimatter are produced in equal amount [2]. One of these condition is related to CP violation which can be measured in B meson decays. At the Large Hadron Collider (LHC) in Geneva [3], the LHCb experiment is the one dedicated to the study of these decays and to the precision measurements related to other particles containing b and c quarks. In this work the decays of neutral B meson in $D^{+}D^{-}$ and $D^{*+}D^{-}$ meson final states are studied. Performing a measurement of the decay-time dependent decay rate the amount of CP violation can be inferred. These decays undergo the transition $b \rightarrow ccd$ and the CP violation is produced in the interference between $B^0 - \bar{B}^0$ mixing with or without decay. Eventually from CP observables the angle β of the Unitary Triangle can be obtained. The data exploited in the analysis presented here are collected at LHC where proton beams are accelerated and collided at center-of-mass energies up to 14 TeV. In 2011 and 2012 (Run 1 period) a data sample corresponding to 3 fb^{-1} were collected using LHCb detector, in 2015 and 2016, referred as Run 2 phase, the data sample collected was around 2 fb^{-1}. These samples represent the world's largest sample of B^0 mesons ever collected. This is the first measurement in these decays performed at LHCb, to update and improve the measurements coming from previous experiments BaBar [4] and Belle [5] (See Fig. 1.1 and Fig. 1.2). Moreover in $B \rightarrow D^{+}D^{-}$ and $B \rightarrow D^{*+}D^{-}$ decays higher order contributions are possible, making this measurement sensitive to effects of physics beyond the standard model. These contributions however need to be constrained

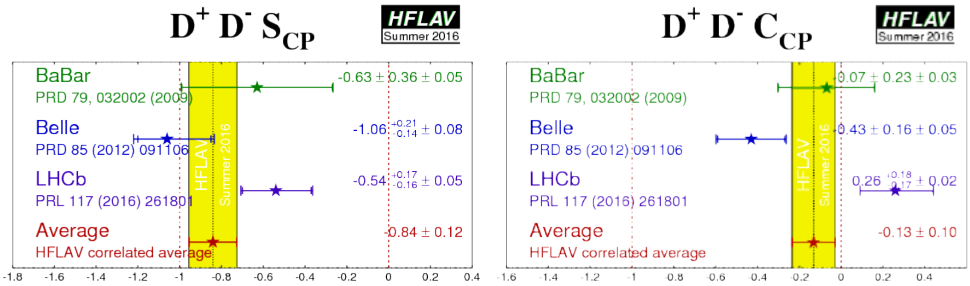


Figure 1.1: HFLAV average of the CP violation parameters in $B^0 \to D^+D^-$ [6].

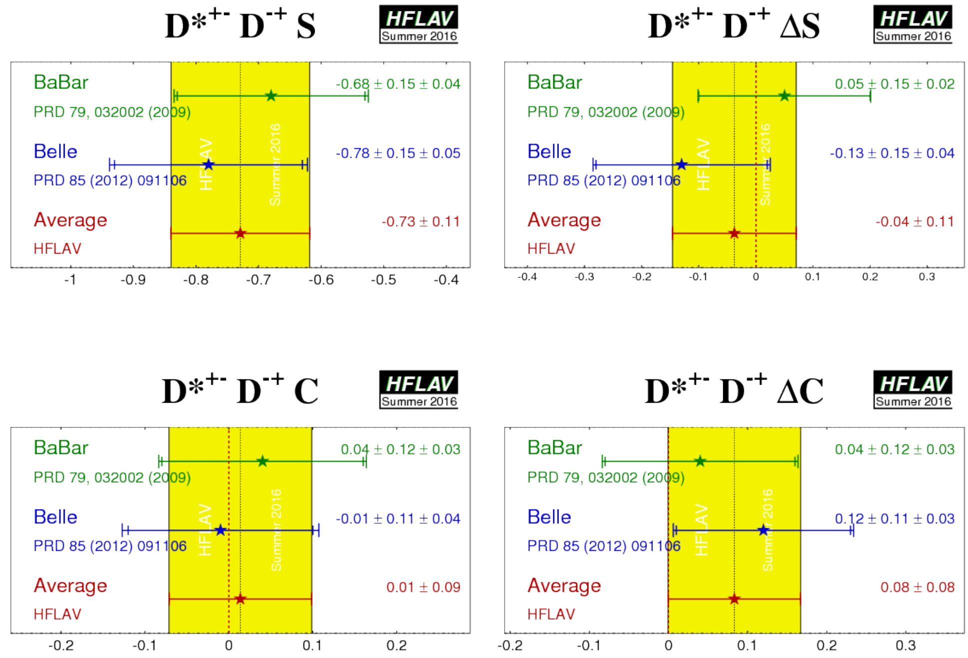


Figure 1.2: HFLAV average of the CP violation parameters in $B^0 \to D^{*+}D^-$ [6].

in order to distinguish them from new physics effects. To have handle on these higher order corrections the measurement of $\sin(2\beta)$ coming from [7] will be used as input. In this thesis after a brief introduction, an overview of the Standard Model is given in Ch. 2. In Ch. 3 a detailed explanation of CP Violation is given along with ways to measure it. In Ch. 4 the LHCb detector is described along with the data flow collection needed to perform this analysis. In Ch. 5 Flavour Tagging studies performed in the flavour specific decays $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*+} D_s^-$ and needed as input for the analysis are described. The selection procedure used to extract the signal in $B^0 \to D^+D^-$ is described in Ch. 6.

In Ch. 7 the result of the analysis using $B^0 \to D^{*+}D^-$ is described. In conclusion in Ch. 8 the final results are presented and compared with previous ones coming from former experiments and a discussion is developed.

Chapter 2

Standard Model of Particle Physics

The standard model of particle physics (SM) is a renormalizable gauge invariant quantum field theory, which describes the fundamental constituents of matter (Sec. 2.1) and three of the four fundamental interactions between them (Sec. 2.2) . This theory has been very successfully in describing experimental data and in predicting their results. However a number of experimental observations are unexplained by this theory, pointing to the need of an extension of SM to wider theories (Sec. 2.4).

2.1 Particles

The SM includes 12 elementary particles called fermions, with $\frac{1}{2}$ spin, and the same number of anti-fermions, which have the opposite charge [8, 9]. According to their interaction, fermions can be divided into six quark and six leptons. The quarks are further subdivided into three generation, each of them contain an up type and a down type quark. The common matter, protons and neutrons, is made of the quarks coming from the first generation, *i.e.* the quark up and the quark down. Their heavier partners are respectively the charm, the top, the strange and the bottom quarks. Quarks cannot be observed alone (confinement [10]) but they are always part of bound states , called hadrons, which can be distinguished into mesons, if composed of a pair of quark-antiquark, baryons if they are made of three quarks or the recently discovered tetraquarks/pentaquarks if they are composed of four or five quarks respectively [11–14]. Quarks are characterized by a colour charge, which can be of three different kinds and sum up such that hadrons are colourless, and an electric charge, which is of $+\frac{2}{3}$ of the elementary charge for up type quarks and $-\frac{1}{3}$ for down type quarks. Also leptons can be classified into three families, each consisting of a negatively charged particle ordered according to growing masses, and a corresponding neutral particle *i.e.* neutrino, which is massless in the SM framework. A graphic representation of SM particles is given in Fig. 2.1.

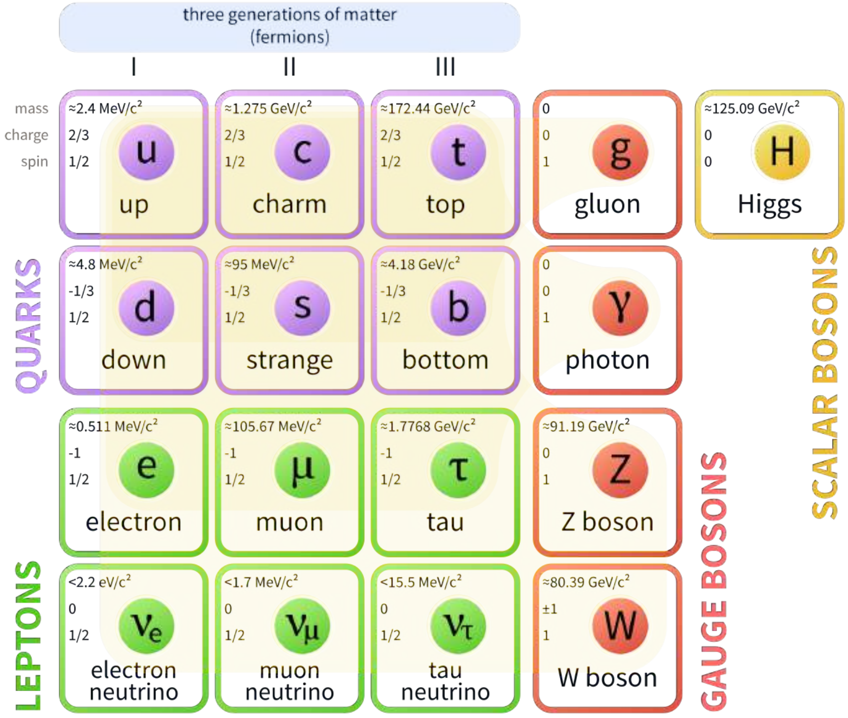


Figure 2.1: Schematic representation of SM particles. The corresponding antiparticles of the fermions have the same mass but charges and spins of opposite signs.

2.2 Interactions

The SM accounts for three of the four fundamental interactions, namely the electromagnetic, the weak and the strong forces. These are mediated by force carrier particles called gauge bosons, all characterized by an integer spin.

Since the SM is a gauge theory, the interactions between particles can be deduced from the internal local symmetries of the theory

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y,$$

where $SU(3)_C$ gives rise to quantum chromodynamics (QCD) which describes the strong force, while $SU(2)_L \otimes U(1)_Y$ describes the electro-weak force. The quantum chromodynamics (QCD) sector defines the interactions between quarks and gluons under SU(3) symmetry.

The Lagrangian describing the coupling of the quarks to the gluon fields is given by

$$L_{QCD} = \bar{\Psi}\gamma^{\mu}\Big(i\partial_{\mu} + gG^{a}_{\mu}T^{a} - m\Big)\Psi - \frac{1}{4}G^{a}_{\mu\nu}G^{\mu\nu}_{a},$$

where Ψ represent the quark field, γ^{μ} are the Dirac matrices, g is the strong coupling constant, G^{a}_{μ} is the SU(3) gauge field, T^{a} are the generators of SU(3) color group and $G^{a}_{\mu\nu}$ are the field strength tensors for the gluons.

The Lagrangian describing the electroweak sector can be written as

$$L_{EW} = \bar{\Psi}\gamma^{\mu}\Big(i\partial_{\mu} - \tilde{g}\frac{1}{2}Y_W B_{\mu} - g\frac{1}{2}\tau W^{\mu}\Big)\Psi - \frac{1}{4}W_a^{\mu\nu}W_{\mu\nu}^a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu},$$

where $\tilde{g}$ and g are the U(1) and SU(2) coupling constants, Y_W and τ are the generators of U(1) and SU(2) group respectively, B_μ and W_μ are the U(1) and SU(2) gauge fields and finally $W_a^{\mu\nu}$ and $B^{\mu\nu}$ are the field strength tensors.

The resulting gauge bosons are 8 gluons (g) for QCD which couple to colour charge, $W_{1,2,3}$ which are the boson of $SU(2)_L$ which couple to the weak isospin J and the gauge boson B of $U(1)_Y$ that couples to the hypercharge Y. Through the Higgs mechanism, a complex scalar doublet $SU(2)_L$ with a vacuum expectation value different from zero, causes a spontaneous symmetry breaking

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \rightarrow SU(3)_C \otimes U(1)_Q,$$

consequences of this are the existence of an electrically neutral massive scalar field (Higgs boson) and that the observable electro-weak force mediators are a linear combination, through the weak mixing angle θ_W, of the gauge bosons $W_{1,2,3}$ and B

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} cos\theta_W & sin\theta_W \\ -sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix},$$

$$W_\mu^{\pm} = \frac{1}{\sqrt{2}}(W_\mu^1 \pm W_\mu^2).$$

The three weak force bosons ($W^{\pm}$and Z^0) are massive, while the residual $U(1)_Q$ gives rise to the electromagnetic force that is carried by the photon (γ) and couples to the electric charge. The Gell-Mann-Nishijima relation $Q = J + \frac{Y}{2}$ expresses the connection between the electric charge and the quantum number of $SU(2)_L$ and $U(1)_Y$.

The Higgs field can be represented as a doublet of complex scalar fields

$$\Phi(x) = \begin{pmatrix} \Phi^+(x) \\ \Phi^0(x) \end{pmatrix},$$

the minimum of the potential is chosen as $\Phi(x) = 1/\sqrt{2}\Big(0, \sqrt{-\frac{\mu^2}{\lambda}} + h(x)\Big)$ with an expectation value on vacuum state of $|< 0 \mid \Phi^0(x) \mid 0 >| \equiv \frac{v}{\sqrt{2}}$, where $v = -\frac{\mu}{\sqrt{\lambda}}$. [15]

The masses of quarks and leptons are generated through the Yukawa interaction between Higgs boson and the fermion fields. For each fermion generation the Yukawa Lagrangian can be written as

$$L_Y = -\frac{1}{\sqrt{2}}(v + H)(c_d\bar{d}d + c_u\bar{u}u + c_l\bar{l}l), \tag{2.1}$$

where u and d represent respectively up-type and down-type quark, while l represents leptons. Fermion masses can then be calculated as

$$M_i = c_i\frac{v}{\sqrt{2}}, \qquad i = u, d, l \tag{2.2}$$

Neutrino remain massless because with the current accepted particle content there are no right-handed ν_R singlet states and one cannot write down couplings like in Eq. (2.1). With the evidence for massive neutrinos [1], one is forced to generate the masses in another way such as using Higgs triplets or adding right-handed neutrino singlets [16].

2.3 Symmetries and Conservation Laws

The SM is constructed to be invariant under gauge transformations, this means that the physics is independent of the choice of the gauge. This symmetry is only broken by the vacuum expectation value of the Higgs field, which is the origin of the masses of W and Z bosons. The invariance under space-time translation corresponds to the conservation of energy. However, according to Heisenberg's uncertainty principle the violation of the energy conservation for a very short period of time is allowed. This allows the existence of virtual, heavy particles, in decay processes. The invariance of a system under translation in space and rotation leads to the conservation of momentum and angular momentum, respectively. The electric charge and the colour charge are additive quantum numbers and the conservation of each of them corresponds to the invariance of the Lagrangian of that interaction under the transformations of the gauge group [17]. The quark flavours, the baryon number, the lepton flavours and the lepton numbers, do not correspond to a gauge symmetry and are not necessarily conserved (actually, quark and lepton flavours are not). For example, flavour transitions are only possible in the weak interaction. Thus, flavour symmetry as well as the U-spin symmetry are approximate symmetries. This reflects the fact that under the assumption that the masses of up, down and strange quarks are the same, processes are invariant under exchange of the two down-type quarks. This allows to transfer some findings from one decay mode to another, for example from decay modes of B^0 mesons to B_s^0 mesons. The parity P, the particle-antiparticle conjugation C and the time reversal T are discrete multiplicative transformations. The parity operation P, is an approximate symmetry and performs a spatial inversion of all coordinates

$$P\Psi(r) = \Psi(-r), \tag{2.3}$$

which means that it transforms left-handed into right-handed fermions. The weak interaction maximally violates parity, i.e. there are only left-handed neutrinos and only right-handed antineutrinos. Charge conjugation C changes the sign of all charges and the magnetic moment, transforming then particles into antiparticles

$$C \mid p >=\mid \bar{p} > . \tag{2.4}$$

Combining charge conjugation with parity CP makes a left-handed neutrino to become a right-handed antineutrino, also this symmetry is violated at the 10^{-4} level by the weak interaction. Combining CP with time reversal T, the CPT symmetry is obtained. The CPT theorem, is a fundamental concept of SM and states that the mass and lifetime of a particle and its antiparticle must be identical.

2.4 SM Extension

Even though the SM proved to be a very successfully theory, there are several open points which cannot be addressed inside this framework. This could point to the idea that the SM is only part of a more fundamental theory, usually called Grand Unified Theory (GUT). In these models the three interactions of the Standard Model which define the electromagnetic, weak, and strong interactions, are merged into one single force. This unified interaction is characterized by one larger gauge symmetry and thus several force carriers are predicted, but only one unified coupling constant. Particles predicted by GUT models cannot be observed directly with the actual experimental available energies, instead the effects of grand unification might be detected through indirect observations. Historically the concept of an unified theory was firstly proposed by Georgi and Glashow in 1974 [18]. In this theory an extension that starts with the generalization of the electroweak and the strong force, to include gravitation is provided. However at the moment gravitation is not included in the SM theory, mainly because this force cannot be probed at the energy scale that characterize the high energy physics experiments, and because it does not include a quantum mechanical derivation. Georgi-Glashow theory is based on the symmetry group SU(5) and the unification of forces is expected to happen at an energy scale of 10^{16} GeV. However, quantum corrections from those mass scales would heavily influence the Higgs mass, which is measured to be around 125 GeV/c^2 [19]. This issue known as hierarchy problem, is solved in the SM by a fine tuning of tree level and loop contributions, which exactly cancel out each other. An alternative approach is the possibility to extend the SM including new symmetries, like is done in extended supersymmetry models [20–23]. From a theoretical point of view, a frail aspect of the SM is the large number of free parameters needed to describe the theory, like the masses of the constituents or the number of generations. In addition, the SM only applies to the processes of ordinary matter, which represents only about 5% of the total energy density of our universe and doesn't contain an explanation for dark matter and dark energy. Furthermore, the amount of CP violation in the weak sector doesn't account for the baryon asymmetry in the universe, i.e. the prevalence of matter over antimatter. Finally the observation of neutrino oscillations [1, 24, 25] imply these particles to be massive, while they are described as massless in the SM.

Chapter 3

CP Violation

3.1 The CKM matrix

Quarks acquire their mass through the coupling to the Higgs field with the vacuum expectation value v and Yukawa interaction between the left-handed and right-handed quark content (Sec. 2.2) . This mechanism is called Spontaneous Symmetry Breaking of the electroweak symmetry group $SU(2) \otimes U(1)$. The Yukawa interaction between quarks and the Higgs doublet is given by the lagrangian in Eq. (3.1), where the Yukawa matrices Y_d and Y_u for down-type and up-type quarks involved are not necessarily diagonal:

$$\mathcal{L}_Y = -\frac{v}{\sqrt{2}}(\bar{d}_L Y_d d_R + \bar{u}_L Y_u u_R) + h.c. \tag{3.1}$$

The resulting mass eigenstates q' are not the same as the eigenstates of the weak interaction but can be obtained from them by using the unitary transformation in Eq. (3.2)

$$q'_i = V_{i,q} q_i, \qquad for\ \ q = u, d\ \ and\ \ i = L, R \tag{3.2}$$

with $V_{i,q}V^\dagger_{i,q} = 1$. Using this transformation the Lagrangian describing the charge-current interaction can be written as in Eq. (3.3)

$$\mathcal{L}_{CC} = -\frac{g_2}{\sqrt{2}}(\bar{u}_L \gamma^\mu W^+_\mu d_L + \bar{d}_L \gamma^\mu W^-_\mu u_L) = -\frac{g_2}{\sqrt{2}}(\bar{u}'_L \gamma^\mu W^+_\mu V_{L,u} V^\dagger_{L,d} d'_L + \bar{d}'_L \gamma^\mu W^-_\mu V_{L,d} V^\dagger_{L,u} u'_L), \tag{3.3}$$

where $V_{CKM} = V_{L,u}V^\dagger_{L,d}$ is the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The CKM matrix is not the unit matrix, so that the Yukawa matrices are not diagonalized by the same unitary transformation, allowing for flavour changes through the weak interaction. In this frame the CKM matrix represents the connection between the mass eigenstates and the weak eigenstates (Eq. (3.4))

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{3.4}$$

The CKM is a unitary, complex 3x3 matrix that can be described using 18 parameters. The unitary condition removes 9 degrees of freedom, a global rephasing allows to constrain five phases, thus only four free parameters describing the matrix remain. Three of these parameters are angles, the remaining one is a complex phase. This phase is the source of CP violation in the SM.

Introducing the notation $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ where the indices i and j run over the number of generations, a standard parametrization for the CKM matrix can be written as in Eq. (3.5)

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{3.5}$$

where δ is the irreducible phase responsible for CP violation. By examining the unitary condition of the CKM matrix, 12 equations can be written, of these six are orthogonality relations, which can be interpreted as triangles in the complex plane. The area of all triangles is the same and given by half of the Jarlskog invariant (Eq. (3.6))

$$J_{CP} = \pm\mathcal{I}m(V_{ik}V_{jl}V^*_{il}V^*_{jk}), \tag{3.6}$$

which quantifies the amount of CP violation in the SM [26–28]. One of the possible equations for the unitary triangles can be written as in Eq. (3.7), this triangle is particularly important for CPV studies, because its sides are of comparable length, it is called the B^0_d triangle, because its angles and sides can be measured through B^0_d decays.

$$V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0. \tag{3.7}$$

In order to show this triangle in the complex plane it is convenient to scale it by dividing all sides by $V_{cd}V^*_{cb}$, so that the base can match the real axis as shown in Fig. 3.1. The three angles of the unitary triangle are defined by

$$\alpha \equiv arg\Big(-\frac{V_{td}V^*_{tb}}{V_{ud}V^*_{ub}}\Big) \quad , \quad \beta \equiv arg\Big(-\frac{V_{cd}V^*_{cb}}{V_{td}V^*_{tb}}\Big) \quad , \quad \gamma \equiv arg\Big(-\frac{V_{ud}V^*_{ub}}{V_{cd}V^*_{cb}}\Big). \tag{3.8}$$

The measurements of CP-violating observables can be used to constrain these angles

- α
  is the phase between $V^*_{tb}V_{td}$ and $V^*_{ub}V_{ud}$, it can be measured only in time-dependent CP asymmetries where the decay mode $b \to u\bar{u}d$ dominates ($B \to \rho\rho$, $B \to \pi\pi$, $B \to \rho\pi$).

- β
  can be measured studying CP violation in B meson decays,where time-dependent CP asymmetry associated to neutral B decays can be written as

$$\mathcal{A}_f(t) = \frac{\Gamma(\bar{B}^0 \to f) - \Gamma(B^0 \to f)}{\Gamma(\bar{B}^0 \to f) + \Gamma(B^0 \to f)} = S_f \sin(\Delta m_d t) - C_f \cos(\Delta m_d t), \tag{3.9}$$

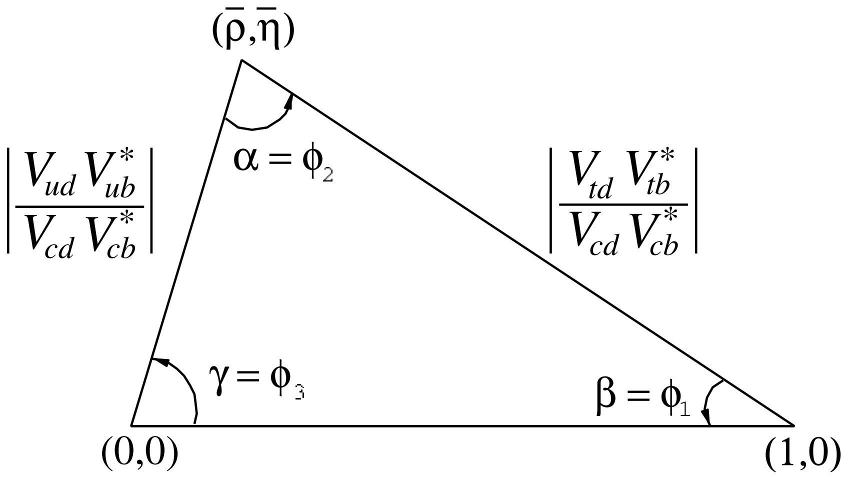


Figure 3.1: Schematic representation of the CKM unitary triangle.

where

$$S_f = \frac{2\mathcal{I}m\lambda_f}{1+\mid\lambda_f\mid^2}, \qquad C_f = \frac{1-\mid\lambda_f\mid^2}{1+\mid\lambda_f\mid^2}, \qquad \lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}, \tag{3.10}$$

while A_f and $\bar{A}_f$ are the decay amplitudes of B^0 and $\overline{B}^0$ into the final state f respectively

$$A_f = \langle f \mid H \mid B^0\rangle, \qquad \bar{A}_f = \langle f \mid H \mid \overline{B}^0\rangle, \tag{3.11}$$

and H is the Hamiltonian governing weak interactions. The ratio q/p describes $B^0 - \bar{B}^0$ mixing, and to a good approximation in the SM

$$\frac{q}{p} = \frac{V_{tb}^*V_{td}}{V_{tb}V_{td}} \simeq e^{-2i\beta}. \tag{3.12}$$

- γ does not depend on CKM elements involving the top quark, as can be seen from Eq. (3.8), so that it can be measured in tree-level B decays.

Another parametrization (Eq. (3.13)) for the CKM matrix can be obtained, expanding as a power series of the parameter $\lambda =\mid V_{us}\mid$, this is called Wolfenstein parametrization [28,29].

$$V = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{3.13}$$

The four parameters entering this matrix are defined as in Eq. (3.14)

$$\lambda = \frac{\mid V_{us}\mid}{\sqrt{\mid V_{ud}\mid^2 + \mid V_{us}\mid^2}},$$

$$A\lambda^2 = \lambda \mid \frac{V_{cb}}{V_{us}} \mid, \tag{3.14}$$

$$A\lambda^3(\rho + i\eta) = V_{ub}^*.$$

The CKM matrix elements and the angles of the Unitary Triangle can be most precisely determined using a global fit to all available measurements and imposing the SM constraints [28]. The fit must also use theoretical predictions for hadronic matrix elements, which sometimes have significant uncertainties. There are several approaches for combining the experimental data. CKMfitter [30, 31] use frequentist statistics, while UTfit [32, 33] uses a Bayesian approach. These approaches provide similar results. Fig. 3.2 illustrates the constraints on the $\bar{\rho}, \bar{\eta}$ plane from various measurements and the global fit result. The shaded 95% CL regions all overlap consistently around the global fit region. The determination of the CKM matrix elements, which are used as input in the UTfit, is possible using experimental processes, a summary of these along with their values are reported in Table 3.1. The result for the angles of the Unitary Triangle obtained from the UTfit are listed in Table 3.2. The constraints on the parameters used in the Wolfenstein parametrization are reported in Table 3.3.

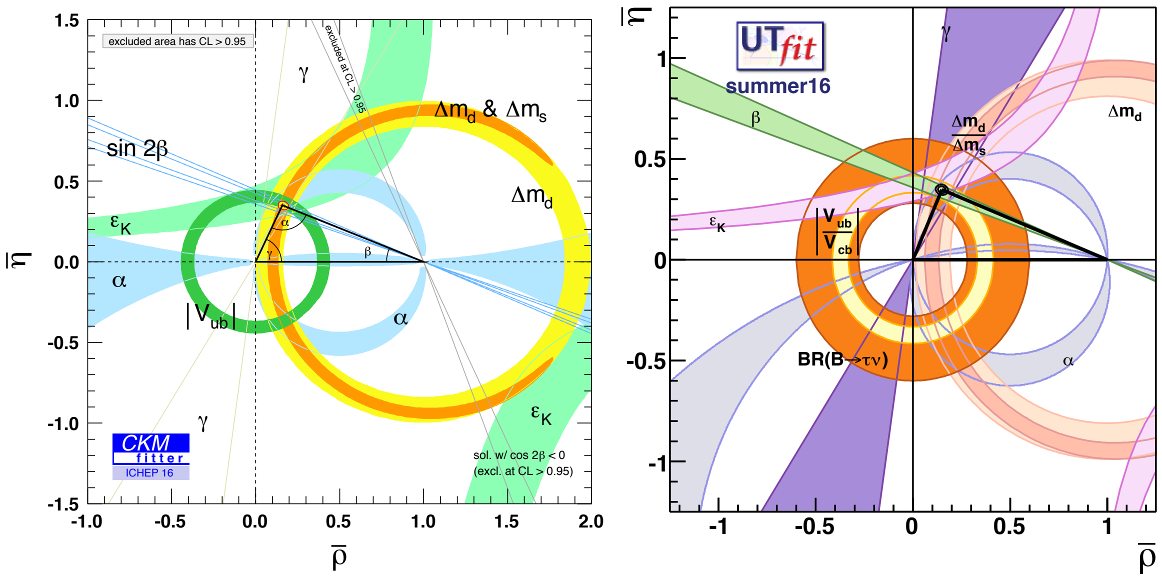


Figure 3.2: Global fit results for the Unitary Triangle, performed by the CKMfit [30] and UTfit [32] groups respectively.

3.2 Mixing in B^0-$\bar{B}^0$ meson system

Consider the two neutral states P^0 and $\bar{P}^0$ [28], where P^0 can be a K^0, B^0 or a D^0, and a state that is initially a superposition of P^0 and $\bar{P}^0$ (Eq. (3.15))

$$\mid \psi(t=0)\rangle = a(0) \mid P^0\rangle + b(0) \mid \bar{P}^0\rangle. \tag{3.15}$$

If we are interested only in the values of a(t) and b(t), and if the times t in which we are interested are much larger than the typical strong interaction scale, then we can use a

Matrix Element	Value	Channel
$\mid V_{ud} \mid$	0.97417±0.00021	Nuclear beta decays
$\mid V_{us} \mid$	0.2248±0.0006	Semileptonic kaon decays
$\mid V_{cd} \mid$	0.220±0.005	ν scattering from valence d quarks
$\mid V_{cs} \mid$	0.995±0.016	Semileptonic D meson decays
$\mid V_{cb} \mid$	$(40.5\pm1.5)\cdot10^{-3}$	Semileptonic B meson decays
$\mid V_{ub} \mid$	$(4.09\pm0.39)\cdot10^{-3}$	Semileptonic B meson decays
$\mid V_{td} \mid$	$(8.2\pm0.6)\cdot10^{-3}$	B^0 mixing assuming $\mid V_{tb} \mid = 1$
$\mid V_{ts} \mid$	$(40.0\pm2.7)\cdot10^{-3}$	B^0_s mixing assuming $\mid V_{tb} \mid = 1$
$\mid V_{tb} \mid$	1.009±0.031	Single-top-quark production

Table 3.1: Values of V_{CKM} matrix elements [28]

Parameter	Value
α [°]	92.0± 2.0
β [°]	22.11± 0.76
γ [°]	65.8± 1.9

Table 3.2: Values of the Unitary Triangle angles estimated from the UTfit [32].

Parameter	Value
A	0.821±0.012
λ	0.22534±0.00065
ρ	0.136± 0.024
η	0.361± 0.014

Table 3.3: Values of the Wolfenstein parameters, estimated by the UTfit group through a global fit [32].

simplified formalism, where the time evolution is described by the Schrödinger equation in Eq. (3.16)

$$i\hbar\frac{\partial}{\partial t}\psi(t) = H\psi(t). \tag{3.16}$$

The hamiltonian operator can be written as in Eq. (3.17)

$$H = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}, \tag{3.17}$$

where the mass matrix **M** and the decay matrix Γ are 2x2 hermitian matrices. Diagonal elements of M and Γ are associated with the flavor-conserving transitions $P^0 \to P^0$ and $\bar{P}^0 \to \bar{P}^0$, while off-diagonal elements are associated with flavor-changing transitions $P^0 \to \bar{P}^0$.

Introducing three complex parameters p, q and z when CP is violated, the mass eigenstates can be written as a linear combination of flavour eigenstates (Eq. (3.18))

$$\begin{aligned} \mid P_L\rangle &\propto p\sqrt{1-z} \mid P^0\rangle + q\sqrt{1+z} \mid \bar{P}^0\rangle, \\ \mid P_H\rangle &\propto p\sqrt{1+z} \mid P^0\rangle - q\sqrt{1-z} \mid \bar{P}^0\rangle, \end{aligned} \tag{3.18}$$

with $\mid p \mid^2 + \mid q \mid^2 = 1$ when z=0. The two eigenvalues $\omega_{L,H}$ are related to the mass and decay widths through Eq. (3.19)

$$\begin{aligned} \Delta m &\equiv m_H - m_L = \mathcal{R}e(\omega_H - \omega_L), \\ \Delta\Gamma &\equiv \Gamma_H - \Gamma_L = -2\mathcal{I}m(\omega_H - \omega_L). \end{aligned} \tag{3.19}$$

The solution to the eigenvalue problem leads to the results in Eq. (3.20) and Eq. (3.21)

$$\left(\frac{q}{p}\right)^2 = \frac{\mathbf{M}^*_{12} - \frac{i}{2}\mathbf{\Gamma}^*_{12}}{\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}}, \tag{3.20}$$

and

$$z = \frac{\delta m - (i/2)\delta\Gamma}{\Delta m - (i/2)\Delta\Gamma}, \tag{3.21}$$

where

$$\delta m = \mathbf{M}_{11} - \mathbf{M}_{22}, \qquad \delta\Gamma = \mathbf{\Gamma}_{11} - \mathbf{\Gamma}_{22}, \tag{3.22}$$

are the differences in effective mass and decay-rate expectation values for the flavour eigenstates P^0 and $\bar{P}^0$. If CP or CPT symmetry are conserved, then δm and $\delta\Gamma$ are zero so that z=0

$$CP\ \ or\ \ CPT\ \ invariance \longrightarrow M_{11} = M_{22},\ \ \Gamma_{11} = \Gamma_{22}$$

$$CP\ \ or\ \ T\ \ invariance \longrightarrow \mathcal{I}mM_{12} = 0 = \mathcal{I}m\Gamma_{12}$$

The time evolution of flavour eigenstates can be expressed as in Eq. (3.23)

$$\begin{aligned} \mid P^0(t)\rangle &= \Big(g_+(t) + zg_-(t) \mid P^0\rangle - \sqrt{1-z^2}\frac{q}{p}g_-(t) \mid \bar{P}^0\rangle, \\ \mid \bar{P}^0(t)\rangle &= \Big(g_+(t) - zg_-(t) \mid \bar{P}^0\rangle - \sqrt{1-z^2}\frac{q}{p}g_-(t) \mid P^0\rangle, \end{aligned} \tag{3.23}$$

where

$$g_\pm(t) = \frac{1}{2}\Big(e^{-im_Ht-\frac{1}{2}\Gamma_Ht} \pm e^{-im_Lt-\frac{1}{2}\Gamma_Lt}\Big). \tag{3.24}$$

Using the definitions given in Eq. (3.10) and Eq. (3.11) and introducing $x \equiv \Delta m/\Gamma$, $y \equiv \Delta\Gamma/2\Gamma$, the expressions in Eq. (3.25) for the transition amplitudes can be obtained

$$\begin{aligned} &\frac{d\Gamma(P^0 \longrightarrow f)/dt}{e^{-\Gamma t}N_f} = \\ &\qquad \Big(\mid A_f \mid^2 + \mid (q/p)\bar{A}_f \mid^2\Big)\cosh(y\Gamma t) + \Big(\mid A_f \mid^2 - \mid (q/p)\bar{A}_f \mid^2\Big)\cos(x\Gamma t) \\ &\qquad + 2\mathcal{R}e\Big(\frac{q}{p}A^*_f\bar{A}_f\Big)\sinh(y\Gamma t) - 2\mathcal{I}m\Big(\frac{q}{p}A^*_f\bar{A}_f\Big)\sin(x\Gamma t), \end{aligned} \tag{3.25}$$

$$\frac{d\Gamma(\bar{P}^0 \longrightarrow f)/dt}{e^{-\Gamma t}N_f} =$$

$$\Big(\mid (p/q)A_f \mid^2 + \mid \bar{A}_f \mid^2\Big)\cosh(y\Gamma t) - \Big(\mid (p/q)A_f \mid^2 - \mid \bar{A}_f \mid^2\Big)\cos(x\Gamma t)$$
$$+ 2\mathcal{R}e\Big(\frac{p}{q}A_f\bar{A}_f^*\Big)\sinh(y\Gamma t) - 2\mathcal{I}m\Big(\frac{p}{q}A_f\bar{A}_f^*\Big)\sin(x\Gamma t), \tag{3.26}$$

where N_f is a common normalization factor. Terms proportional to $\mid A_f \mid^2$ and $\mid \bar{A}_f \mid^2$ are associated with decays that occur without oscillation while terms proportional to $\mid (\frac{q}{p})\bar{A}_f \mid^2$ and $\mid (\frac{p}{q})A_f \mid^2$ are associated with decays following an oscillation. In Eq. (3.26), the terms $\sinh(y\Gamma t)$ and $\sin(x\Gamma t)$ are associated with the interference between the two cases.

Let's now consider in more detail the mixing phenomenology for the specific case of B^0 mesons [28]. B^0_d mesons are particles made by a b or $\bar{b}$ and a $\bar{d}$, d, characterized by a mass of $m_B \simeq 5$ GeV and a lifetime of $\tau_B \simeq 1.5$ ps. The flavour eigenstates B^0_d and $\bar{B}^0_d$ can mix, i.e. they can oscillate between the two flavour states. The $B^0 - \bar{B}^0$ oscillation is in lowest order Standard Model described by quantum loops involving charged currents, as shown in the Feynman diagrams in Fig. 3.3.

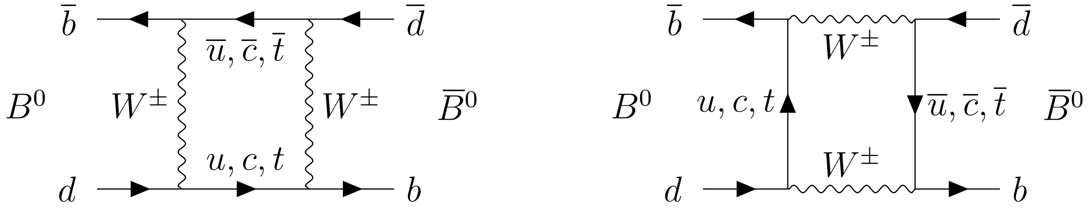


Figure 3.3: Box diagrams for $B^0 - \bar{B}^0$ oscillations.

For this system the mass eigenstates can be defined as in Eq. (3.27)

$$\mid B_{L,H}\rangle = p \mid B^0\rangle \pm q \mid \bar{B}^0\rangle, \tag{3.27}$$

and the time evolution is governed by Eq. (3.28)

$$\mid B_{L,H}\rangle = e^{-(iM_{L,H}+\Gamma_{L,H}/2)t} \mid B_{L,H}\rangle, \tag{3.28}$$

the mass and total decay width difference are described by Eq. (3.29)

$$\Delta m \equiv M_H - M_L = -2\mathcal{R}e\Big[\frac{q}{p}\Big(M_{12} - \frac{i}{2}\Gamma_{12}\Big)\Big],$$

$$\Delta\Gamma \equiv M_H - M_L = -2\mathcal{I}m\Big[\frac{q}{p}\Big(M_{12} - \frac{i}{2}\Gamma_{12}\Big)\Big]. \tag{3.29}$$

Combining Eq. (3.27) and Eq. (3.28) the time evolution of the flavour eigenstates can be obtained (Eq. (3.30))

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

$$\mid \bar{B}^0(t)\rangle = \frac{p}{q} g_-(t) \mid B^0\rangle + g_+(t) \mid \bar{B}^0\rangle, \tag{3.30}$$

where $g_\pm$ are defined according to Eq. (3.24).

It is now useful introducing the decay amplitudes for a $B^0/\overline{B}^0$ into a final state f (Eq. (3.31))

$$\begin{aligned} A_f &= \langle f \mid H \mid B\rangle, & \bar{A}_{\bar{f}} &= \langle \bar{f} \mid H \mid \bar{B}\rangle, \\ A_{\bar{f}} &= \langle \bar{f} \mid H \mid B\rangle, & \bar{A}_f &= \langle f \mid H \mid \bar{B}\rangle. \end{aligned} \tag{3.31}$$

Finally the decay rates for a B or a $\bar{B}$ into a final state f or into its CP conjugate $\bar{f}$ are given by Eq. (3.32)

$$\begin{aligned} \Gamma(B(t) \to f) &\propto \frac{\mid A_f \mid^2}{2} e^{-\Gamma_q t}[I_+(t) + I_-(t)], \\ \Gamma(\bar{B}(t) \to f) &\propto \frac{\mid A_f \mid^2}{2} \mid \frac{p}{q} \mid^2 e^{-\Gamma_q t}[I_+(t) - I_-(t)], \\ \Gamma(\bar{B}(t) \to \bar{f}) &\propto \frac{\mid \bar{A}_{\bar{f}} \mid^2}{2} e^{-\Gamma_q t}[\bar{I}_+(t) + \bar{I}_-(t)], \\ \Gamma(B(t) \to \bar{f}) &\propto \frac{\mid \bar{A}_{\bar{f}} \mid^2}{2} \mid \frac{p}{q} \mid^2 e^{-\Gamma_q t}[\bar{I}_+(t) - \bar{I}_-(t)], \end{aligned} \tag{3.32}$$

where

$$\begin{aligned} I_+(t) &= (1+ \mid \lambda_f \mid^2)\cosh(\Delta\Gamma_q t/2) - 2\mathcal{R}e(\lambda_f)\sinh(\Delta\Gamma_q t/2), \\ I_-(t) &= (1- \mid \lambda_f \mid^2)\cos(\Delta m_q t) - 2\mathcal{I}m(\lambda_f)\sin(\Delta m_q t), \\ \bar{I}_+(t) &= (1+ \mid \bar{\lambda}_f \mid^2)\cosh(\Delta\Gamma_q t/2) - 2\mathcal{R}e(\bar{\lambda}_f)\sinh(\Delta\Gamma_q t/2), \\ \bar{I}_-(t) &= (1- \mid \bar{\lambda}_f \mid^2)\cos(\Delta m_q t) - 2\mathcal{I}m(\bar{\lambda}_f)\sin(\Delta m_q t). \end{aligned} \tag{3.33}$$

These expressions represent the theoretical decay rates, they will be extended in Sec. 6.3 and Sec. 7.3 in order to take into account experimental effects such as production and detection asymmetries and experimental effects associated to the wrong determination of the B^0 production flavour.

3.3 Types of CP violation

We distinguish three types of CP-violating effects that can occur in the neutral meson system [34]:

1. CP violation in decay, which occurs when the decay amplitudes differ between CP conjugated processes
2. CP violation in mixing, when the mass eigenstates are no CP eigenstates
3. CP violation in the interference when there is interference between direct decays and decays to the same final state after mixing

3.3.1 CP violation in decay

For any final state f it is useful to write each contribution to its decay amplitude A_f as sum of three parts: its magnitude A_i, its weak phase $e^{i\phi_i}$, and its strong phase term $e^{i\delta_i}$. The first kind of phase appearing in the amplitude is associated to complex parameters in the Lagrangian implying that in A_f and $\bar{A}_{\bar{f}}$ they have opposite signs. In the Standard Model these phases occur only in the CKM matrix which is part of the electroweak sector of the theory, for this reason they are called weak phases. The weak phase of any single term is convention dependent. However the difference between the weak phases in two different terms in A_f is convention independent; the initial and final states are the same for every term and thus any phase rotation of the fields that appear in these states will affect all terms in the same way. The second type of phase can appear in scattering or decay amplitudes even when the Lagrangian is real. Such phases do not violate CP, since they appear in A_f and $\bar{A}_{\bar{f}}$ with the same sign. Their origin is the possible contribution from intermediate on-shell states in the decay process, that is an absorptive part of an amplitude that has contributions from coupled channels. Usually the dominant rescattering is due to strong interactions, hence the name strong phases for the phase shifts that is induced. Again only the relative strong phases of different terms in a scattering amplitude have physical content, an overall phase rotation of the entire amplitude has no physical consequences. Considering several amplitudes contributing to $B^0 \to f$, the amplitude A_f and the CP conjugate amplitude $\bar{A}_{\bar{f}}$ can be written as in Eq. (3.34)

$$A_f = \sum_i A_i e^{i(\delta_i + \phi_i)}, \qquad \bar{A}_{\bar{f}} = e^{2i(\xi_f - \xi_B)} \sum_i A_i e^{i(\delta_i - \phi_i)}, \tag{3.34}$$

here ξ_f and ξ_B are arbitrary phases coming from CP transformation on the B^0 meson and its final state. If f is a CP eigenstate the resulting eigenvalue is expressed by $e^{2i\xi_f} = \pm 1$. Direct CP violation is realized when (Eq. (3.35))

$$\left|\frac{\bar{A}\bar{f}}{A_f}\right| = \left|\frac{\sum_i A_i e^{i(\delta_i - \phi_i)}}{\sum_i A_i e^{i(\delta_i + \phi_i)}}\right| \neq 1. \tag{3.35}$$

3.3.2 CP violation in mixing

This type of CP violation called also CP violation in mixing, results from the mass eigenstates being different from the CP eigenstates. In this case the relative phase between M_{12} and Γ_{12} doesn't vanish. Therefore from Eq. (3.20) follows that

$$\mid q/p \mid \neq 1. \tag{3.36}$$

CP violation in mixing was observed unambiguously in the neutral kaon system. For the neutral B system, this effect could be observed through the asymmetries in semileptonic decays:

$$a_{sl} = \frac{\Gamma(\bar{B}^0_{phys}(t) \to l^+ \nu X) - \Gamma(B^0_{phys}(t) \to l \nu X)}{\Gamma(\bar{B}^0_{phys}(t) \to l^+ \nu X) + \Gamma(B^0_{phys}(t) \to l \nu X)}, \tag{3.37}$$

that can be expressed in terms of q/p,

$$a_{sl} = \frac{1- \mid q/p \mid^4}{1+ \mid q/p \mid^4}. \tag{3.38}$$

Effects of CP violation in mixing of neutral B_d^0 decays, such as the asymmetries in semileptonic decays, are expected to be of the order of 10^{-4}. However to calculate the deviation of q/p from a pure phase, one needs to calculate Γ_{12} and M_{12}. This involves large hadronic uncertainties. Thus even if such asymmetries are observed, it will be difficult to relate their rates to fundamental CKM parameters.

3.3.3 CP violation in interference

Let's consider a neutral B decay into final CP eigenstates f_{CP}, such states are accessible in both $B^0/\overline{B}^0$ decays. The quantity of interest here which is independent of phase conventions and physically meaningful is λ_f, defined according to Eq. (3.10). When CP is conserved, $q/p = 1$ and $\bar{A}_{\bar{f}CP}/A_{fCP} = 1$ and furthermore, the relative phase between (q/p) and $\bar{A}_{\bar{f}CP}/A_{fCP}$ vanishes. Therefore

$$\lambda_{fCP} \neq \pm 1, \tag{3.39}$$

implies CP violation. Note that both CP violation in decay and CP violation in mixing lead to Eq. (3.39) through $\lambda_{fCP} \neq 1$. However, it is possible that, to a good approximation, $q/p = 1$ and $\bar{A}_{fCP}/A_{fCP} = 1$, yet there is CP violation:

$$\lambda_{fCP} = 1, \qquad Im\lambda_{fCP} \neq 0. \tag{3.40}$$

This type of CP violation is called CP violation in the interference between decays with or without mixing, or it is abbreviated as interference between mixing and decay.

For the neutral B system, CP violation in the interference between decays with and without mixing can be observed by comparing decays into final CP eigenstates of a time-evolving neutral B state that begins at time zero as B^0 to those of the state that begins as a $\overline{B}^0$:

$$a_{fCP} = \frac{\Gamma(B^0_{phys}(t) \to f_{CP}) - \Gamma(\bar{B}^0_{phys}(t) \to f_{CP})}{\Gamma(B^0_{phys}(t) \to f_{CP}) + \Gamma(\bar{B}^0_{phys}(t) \to f_{CP})}. \tag{3.41}$$

This time-dependent asymmetry is given by:

$$a_{fCP} = \frac{2Im\lambda_{fCP}\sin(\Delta m_d t) - (1- \mid \lambda_{fCP} \mid^2)\cos(\Delta m_d t)}{(1+ \mid \lambda_{fCP} \mid^2)\cosh(\frac{\Delta\Gamma_d t}{2}) + 2Re\lambda_{fCP}\sinh(\frac{\Delta\Gamma_d t}{2})}. \tag{3.42}$$

This asymmetry only vanishes if $\mid \lambda_{fCP} \mid^2 \neq 1$ (direct or indirect CP violation) or if λ_{fCP} has an imaginary part different from zero.

Using the definitions given in Eq. (3.10) and assuming $\Delta\Gamma$=0 the time-dependent asymmetry can be written as in Eq. (3.43)

$$a_{fCP} = S_f \sin(\Delta m_d t) - C_f \cos(\Delta m_d t). \tag{3.43}$$

3.4 CP Violation in $B^0 \to D^+D^-$

In $B^0 \to D^+D^-$ decay, *CP* violation can arise both due to the interference of the direct decay and the decay after B^0-$\overline{B}^0$ mixing and in the direct decay. The $B^0 \to D^+D^-$ mode is caused by $\bar{b} \to \bar{c}c\bar{d}$ quark-level transitions, and in the SM receives contributions from different decay topologies, as illustrated in Fig. 3.4 [35]. The decay amplitude can be written as follows

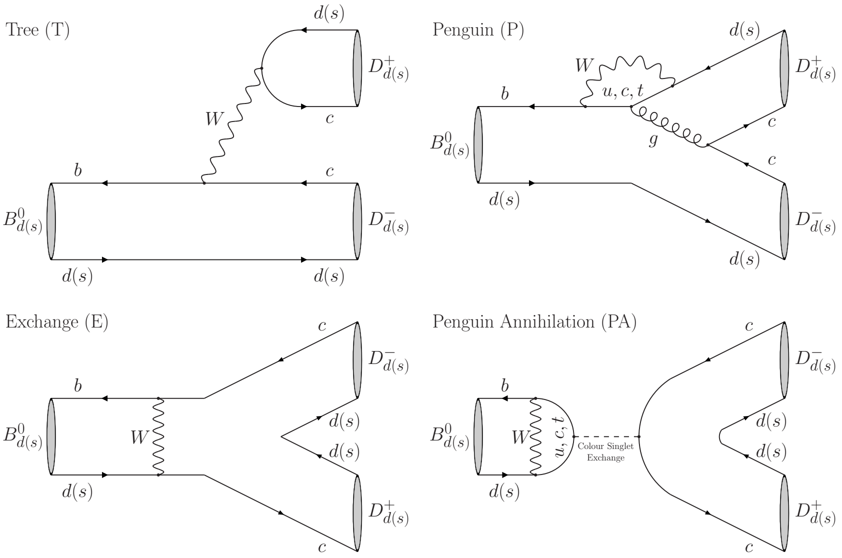


Figure 3.4: Illustration of topologies contributing to $B^0 \to D^+D^-$ decays [35].

$$A(B^0 \to D^+D^-) = V_{cb}^* V_{cd} \mathcal{A}[1 - ae^{i\theta}e^{-i\gamma}], \tag{3.44}$$

where γ serves as a *CP*-violating weak phase and is the angle of the unitary triangle of the CKM matrix, while

$$\mathcal{A} \equiv [T + E + \{P^{(c)} + PA^{(c)}\} - \{P^{(t)} + PA^{(t)}\}], \tag{3.45}$$

and

$$ae^{i\theta} \equiv R_b \Big[\frac{P^{(u)} + PA^{(u)} - P^{(t)} + PA^{(t)}}{T + E + P^{(c)} + PA^{(c)} - P^{(t)} + PA^{(t)}}\Big], \tag{3.46}$$

are related to *CP*-conserving hadronic parameters. Here T and $P^{(q)}$ denote the strong amplitudes of the (colour-allowed) tree and penguin topologies (with internal q-quark exchanges) respectively, which can be expressed in terms of hadronic matrix elements of the corresponding low-energy effective Hamiltonian. The amplitudes describing exchange E and penguin annihilation $PA^{(q)}$ topologies, must also be included. The parameter

$$R_b = \Big(1 - \frac{\lambda^2}{2}\Big)\frac{1}{2}\Big|\frac{V_{ub}}{V_{cb}}\Big|, \tag{3.47}$$

measures the side of the UT originating from the origin of the complex plane with the angle γ between the real axis, while λ is defined according to Wolfenstein parametrization. The corresponding $\overline{B}^0$ decay amplitude can then be written as

$$A(\overline{B}^0 \to D^+D^-) = V_{cb}V_{cd}^*\mathcal{A}[1 - ae^{i\theta}e^{i\gamma}]. \tag{3.48}$$

Using the ratio of the mixing coefficients (Eq. (3.20)) along with the definition of the unitary triangle angle β (Eq. (3.8)), the parameter describing *CP* violation can be written following Eq. (3.10)

$$\begin{aligned}\lambda_{D^+D^-} &= \frac{V_{tb}^*V_{td}}{V_{tb}^*V_{td}}\frac{V_{cb}V_{cd}^*}{V_{cb}^*V_{cd}}\frac{1-ae^{i\theta}e^{i\gamma}}{1-ae^{i\theta}e^{-i\gamma}}\\ &= e^{-i2\beta}\frac{1-ae^{i\theta}e^{i\gamma}}{1-ae^{i\theta}e^{-i\gamma}}.\end{aligned} \tag{3.49}$$

The hadronic parameters must be determined through a measurement of the *CP* observables, that can be written as (see Eq. (3.10))

$$\begin{aligned}C_{D^+D^-} &= A_{CP}^{dir}(B^0 \to D^+D^-) = \frac{1-\mid\lambda\mid^2}{1+\mid\lambda\mid^2} = \frac{2b_d\sin\rho_d\sin\gamma}{1-2b_d\cos\rho_d\cos\gamma+b_d^2},\\ S_{D^+D^-} &= A_{CP}^{mix}(B^0 \to D^+D^-) = \frac{2Im\lambda}{1+\mid\lambda\mid^2} = \eta_d\Big[\frac{\sin\phi_d - 2b_d\cos\rho_d\sin(\phi_d+\gamma)+b_d^2\sin(\phi_d+2\gamma)}{1-2b_d\cos\rho_d\cos\gamma+b_d^2}\Big],\end{aligned} \tag{3.50}$$

where $b_de^{i\rho_d} = ae^{i\theta}$. The parameter η_q represents the *CP* eigenvalue of the final state and is given by +1. The direct *CP* asymmetry $A_{CP}^{dir}(B^0 \to D^+D^-)$ is caused by the interference between tree and penguin contributions, while the mixing-induced *CP* asymmetry $A_{CP}^{mix}(B^0 \to D^+D^-)$ originates from the interference between B^0-$\overline{B}^0$ mixing and decay processes, and depend on the mixing phase ϕ_d. This phase takes the general form

$$\phi_d = 2\beta + \phi_d^{NP}, \tag{3.51}$$

where β is the angle of the UT. In the SM the *CP*-violating phase ϕ_d^{NP} vanishes, allowing to take into account NP contributions to B^0-$\overline{B}^0$ mixing. It is useful to introduce the so-called effective mixing phase

$$\phi_d^{eff} \equiv \phi_d + \Delta\phi_d, \tag{3.52}$$

using the following expression

$$\frac{A_{CP}^{mix}(B^0 \to D^+D^-)}{\sqrt{1-(A_{CP}^{dir}(B^0 \to D^+D^-))^2}} = -\frac{S_{D^+D^-}}{\sqrt{1-C_{D^+D^-}^2}} = -\sin(\phi_d^{eff}), \tag{3.53}$$

where the hadronic penguin phase shift $\Delta\phi_d$ is characterized by

$$\tan\Delta\phi_d = \frac{-2a\cos\theta\sin\gamma + a^2\sin 2\gamma}{1-2a\cos\theta\cos\gamma+a^2\cos 2\gamma}, \tag{3.54}$$

Following the HFLAV convention the two different decay-time dependent partial decay rates of initial $B^0/\overline{B}^0$ mesons decaying into their common final state D^+D^-, can be written as Eq. (3.55), which give access to the previously introduced *CP* observables.

$$\begin{aligned}\frac{d\Gamma_{B^0}(t)}{dt} &= \frac{A_{B^0}e^{-t/\tau}}{2}\Big(\cosh\Big(\frac{\Delta\Gamma t}{2}\Big) + D\sinh\Big(\frac{\Delta\Gamma t}{2}\Big) - S\sin(\Delta m t) + C\cos(\Delta m t)\Big),\\ \frac{d\Gamma_{\overline{B}^0}(t)}{dt} &= \frac{A_{\overline{B}^0}e^{-t/\tau}}{2}\Big(\cosh\Big(\frac{\Delta\Gamma t}{2}\Big) + D\sinh\Big(\frac{\Delta\Gamma t}{2}\Big) + S\sin(\Delta m t) - C\cos(\Delta m t)\Big).\end{aligned} \tag{3.55}$$

3.5 CP Violation in $B^0 \to D^{*\pm}D^{\mp}$

The $B^0 \to D^{*\pm}D^{\mp}$ decays allow to probe independently from $B^0 \to D^{\pm}D^{\mp}$ decays for penguin effects in double-charm modes. As in $B^0 \to D^+D^-$ decay, an interference between direct decays and decays following B^0-$\overline{B}^0$ oscillation emerges, and mixing-induced and direct *CP* violation similar to that of $B^0 \to D^+D^-$ decay are expected. However unlike D^+D^-, the $D^{*+}D^-$ and $D^{*-}D^+$ configurations are not *CP* eigenstates. Both these states are accessible from B^0 and $\overline{B}^0$ decays with amplitudes of comparable magnitudes. Therefore four different flavor-charge configurations have to be considered, and the time-dependent decay rates can be written as [6]

$$\begin{aligned}\frac{d\Gamma_{\overline{B}^0,f}(t)}{dt} &= e^{-t/\tau}(1+A_{f\bar{f}})\Big[1 + S_f\sin(\Delta m t) - C_f\cos(\Delta m t)\Big],\\ \frac{d\Gamma_{B^0,f}(t)}{dt} &= e^{-t/\tau}(1+A_{f\bar{f}})\Big[1 - S_f\sin(\Delta m t) + C_f\cos(\Delta m t)\Big],\\ \frac{d\Gamma_{\overline{B}^0,\bar{f}}(t)}{dt} &= e^{-t/\tau}(1-A_{f\bar{f}})\Big[1 + S_{\bar{f}}\sin(\Delta m t) - C_{\bar{f}}\cos(\Delta m t)\Big],\\ \frac{d\Gamma_{B^0,\bar{f}}(t)}{dt} &= e^{-t/\tau}(1-A_{f\bar{f}})\Big[1 - S_{\bar{f}}\sin(\Delta m t) + C_{\bar{f}}\cos(\Delta m t)\Big],\end{aligned} \tag{3.56}$$

here $f = D^{*+}D^-$ and $\bar{f} = D^{*-}D^+$, while the parameter $A_{f\bar{f}}$ represents an overall asymmetry in the production of the f and $\bar{f}$ final states, and can be written as

$$A_{f\bar{f}} = \frac{\Big(\mid A_f\mid^2 + \mid \bar{A}_f\mid^2\Big) - \Big(\mid A_{\bar{f}}\mid^2 + \mid \bar{A}_{\bar{f}}\mid^2\Big)}{\Big(\mid A_f\mid^2 + \mid \bar{A}_f\mid^2\Big) + \Big(\mid A_{\bar{f}}\mid^2 + \mid \bar{A}_{\bar{f}}\mid^2\Big)}. \tag{3.57}$$

The parameters S_f, $S_{\bar{f}}$, C_f and $C_{\bar{f}}$ are defined according to Eq. (3.10) including the decay amplitudes and assuming that both *CP* violation in mixing as in decay hold

$$\begin{aligned}
S_f &= \frac{-2 \mid A_f \mid\mid \bar{A}_f \mid \sin(\phi_{mix} + \phi_{dec} - \delta_f)}{\mid A_f \mid^2 + \mid \bar{A}_f \mid^2}, \\
S_{\bar{f}} &= \frac{-2 \mid A_{\bar{f}} \mid\mid \bar{A}_{\bar{f}} \mid \sin(\phi_{mix} + \phi_{dec} + \delta_f)}{\mid A_{\bar{f}} \mid^2 + \mid \bar{A}_{\bar{f}} \mid^2}, \\
C_f &= \frac{\mid A_f \mid^2 - \mid \bar{A}_f \mid^2}{\mid A_f \mid^2 + \mid \bar{A}_f \mid^2}, \\
C_{\bar{f}} &= \frac{\mid A_{\bar{f}} \mid^2 - \mid \bar{A}_{\bar{f}} \mid^2}{\mid A_{\bar{f}} \mid^2 + \mid \bar{A}_{\bar{f}} \mid^2},
\end{aligned} \tag{3.58}$$

where $\phi_{mix} = arg(q/p)$ and $\phi_{dec} = arg(\bar{A}_f/A_f)$. Moreover $S_{D^{*\pm}D^{\mp}}$ and $C_{D^{*\pm}D^{\mp}}$ for the $D^{*+}D^-$ and $D^{*-}D^+$ configurations are not independent, but related by the following equation

$$\frac{S_{D^{*\pm}D^{\mp}}}{\sqrt{1 - C^2_{D^{*\pm}D^{\mp}}}} = -\sin(2\tilde{\phi}_d^{eff} \pm \delta). \tag{3.59}$$

The expression includes dependencies on the mixing phase $\tilde{\phi}_d$, which can be modified by penguin contributions to the effective phase $\tilde{\phi}_d^{eff}$, and on the relative strong phase δ between $B^0 \to D^{*+}D^-$ and $B^0 \to D^{*-}D^+$ decay amplitudes.

When introducing the following parametrization for the *CP* observables

$$\begin{aligned}
S_{D^*D} &= \frac{1}{2}(S_{D^{*+}D^-} + S_{D^{*-}D^+}), \\
\Delta S_{D^*D} &= \frac{1}{2}(S_{D^{*+}D^-} - S_{D^{*-}D^+}), \\
C_{D^*D} &= \frac{1}{2}(C_{D^{*+}D^-} + C_{D^{*-}D^+}), \\
\Delta C_{D^*D} &= \frac{1}{2}(C_{D^{*+}D^-} - C_{D^{*-}D^+}),
\end{aligned} \tag{3.60}$$

the decay rates of Eq. (3.56) can be written as

$$\begin{aligned}
\frac{d\Gamma_{\bar{B}^0,f}(t)}{dt} &= e^{-t/\tau}(1 + A_{f\bar{f}})\Big[1 + (S_{D^*D} + \Delta S_{D^*D})\sin(\Delta m t) - (C_{D^*D} + \Delta C_{D^*D})\cos(\Delta m t)\Big], \\
\frac{d\Gamma_{B^0,f}(t)}{dt} &= e^{-t/\tau}(1 + A_{f\bar{f}})\Big[1 - (S_{D^*D} + \Delta S_{D^*D})\sin(\Delta m t) + (C_{D^*D} + \Delta C_{D^*D})\cos(\Delta m t)\Big], \\
\frac{d\Gamma_{\bar{B}^0,\bar{f}}(t)}{dt} &= e^{-t/\tau}(1 - A_{f\bar{f}})\Big[1 + (\Delta S_{D^*D} - S_{D^*D})\sin(\Delta m t) - (\Delta C_{D^*D} - C_{D^*D})\cos(\Delta m t)\Big], \\
\frac{d\Gamma_{B^0,\bar{f}}(t)}{dt} &= e^{-t/\tau}(1 - A_{f\bar{f}})\Big[1 - (\Delta S_{D^*D} - S_{D^*D})\sin(\Delta m t) + (\Delta C_{D^*D} - C_{D^*D})\cos(\Delta m t)\Big],
\end{aligned} \tag{3.61}$$

here S_{D^*D} parametrizes mixing-induced and C_{D^*D} flavor-dependent direct CP violation. The parameters ΔS_{D^*D} and ΔC_{D^*D} are not sensitive to CP violation. The parameter

ΔS_{D^*D} is related to the relative strong phase δ between the decay amplitudes and ΔC_{D^*D} describes the asymmetry between the rates $\Gamma(B^0 \to D^{*+}D^-) + \Gamma(\overline{B}^0 \to D^{*-}D^+)$ and $\Gamma(B^0 \to D^{*-}D^+) + \Gamma(\overline{B}^0 \to D^{*+}D^-)$.

If the contributions of penguin amplitudes to $B^0 \to D^{*\pm}D^{\mp}$ decays are negligible, the relative strong phase between the $B^0 \to D^{*+}D^-$ and $B^0 \to D^{*-}D^+$ decay amplitudes is zero and their magnitudes are the same, then

$$\begin{aligned} A_{D^*D} &= 0, \\ S_{D^*D} &= -\sin(2\beta), \\ C_{D^*D} &= \Delta S_{D^*D} = \Delta C_{D^*D} = 0. \end{aligned} \tag{3.62}$$

In this case, the time-dependent CP violating asymmetry in $B^0 \to D^{*\pm}D^{\mp}$ decays measures directly $-\sin(2\beta)$.

The analysis of $B^0 \to D^+D^-$ allows to determine the phase shift $\Delta\phi$ and hence the hadronic parameters a and θ. These can then be transferred to the measurement of *CP* violation in $B^0_s \to D^+_s D^-_s$ decay. In fact $B^0 \to D^+_s D^-_s$ is related to $B^0 \to D^+D^-$ via U-spin symmetry , however in this channel the relation between the phase shift and the hadronic parameters is reduced by an additional term ϵ, as shown in Eq. (3.63)

$$\tan\Delta\phi_s = \frac{2\epsilon a' \cos\theta' \sin\gamma + \epsilon^2 a^{\text{‘}2} \sin 2\gamma}{1 + 2\epsilon a' \cos\theta' \cos\gamma + \epsilon^2 a'^2 \cos 2\gamma}, \tag{3.63}$$

where $\epsilon \equiv 0.0536 \pm 0.0003$ [35]. The study of $B^0 \to D^{*\pm}D^{\mp}$ decays also enable the determination of $\tilde{\phi}_d^{eff}$, and eventually the associated hadronic parameters, however in this decay other factors make the framework more complicated. First the $D^{*\pm}$ $D^{\mp}$ is not a *CP* eigenstate , implying an higher number of *CP* observables to be taken into account. Moreover experimentally the final state is not symmetrical in terms of the charges of pions and kaons, thus a detection asymmetry has to be included to account for the different phase space (kinematic and geometrical distributions) of the final state π coming from the $D^{*\pm}$ with respect to the one coming from the $D^{\pm}$, this will be discussed in Ch. 7.

Chapter 4

The LHCb experiment at LHC

4.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest accelerator, able to achieve the highest energies ever available at a collider up to date [3, 36, 37]. Proposed and realized by the European Organization for Nuclear Research (CERN), it was designed to collide proton beams as well as heavy ions, at an unprecedent rate, in order to solve the until now unaddressed fundamental questions in particle physics.

4.1.1 LHC design and performances

The LHC lies in the 27 Km long LEP tunnel [38], situated at a depth of about 100 m underground near the border between Switzerland and France. The tunnel contains two parallel beam pipes, where proton or ions travel in opposite directions, and intersect each other in four different points, where the main experimental halls are built and detectors are located (See Fig. 4.1). The beams are kept on their circular path by the usage of 1232 dipole magnets, while 392 additional quadrupole magnets are used to keep the beam focalized, in order to maximize the probability for an interaction to occur at one of the four cross points. In total about 1600 superconducting magnets are used, and approximately 96 tons of liquid helium are needed to maintain the magnet at the operating temperature of 1.9 K. The field in the dipole magnets ranges from 0.53 T up to 8.3 T, and the proton beams are accelerated from 450 GeV up to 7 TeV. Before being injected into the main accelerator, protons undergo different accelerating stages (Fig. 4.2). The first system is the linear particle accelerator (LINAC 2), that generates protons at an energy of 50 MeV, those are then driven into the Proton Synchrotron Booster (PSB). At this stage protons are accelerated up to an energy of 1.4 GeV and then injected into the Proton Synchrotron which raise their energy up to 26 GeV. Finally the Super Proton Synchrotron is used to further increase their energy up to 450 GeV, from here they are injected into the LHC, where protons are accumulated, accelerated (over a period of about 20 minutes) and finally circulated.

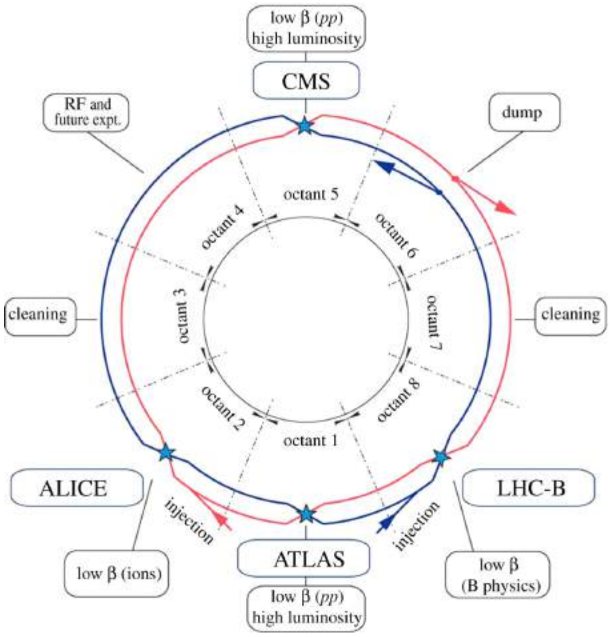


Figure 4.1: A schematic picture of LHC injection scheme [39].

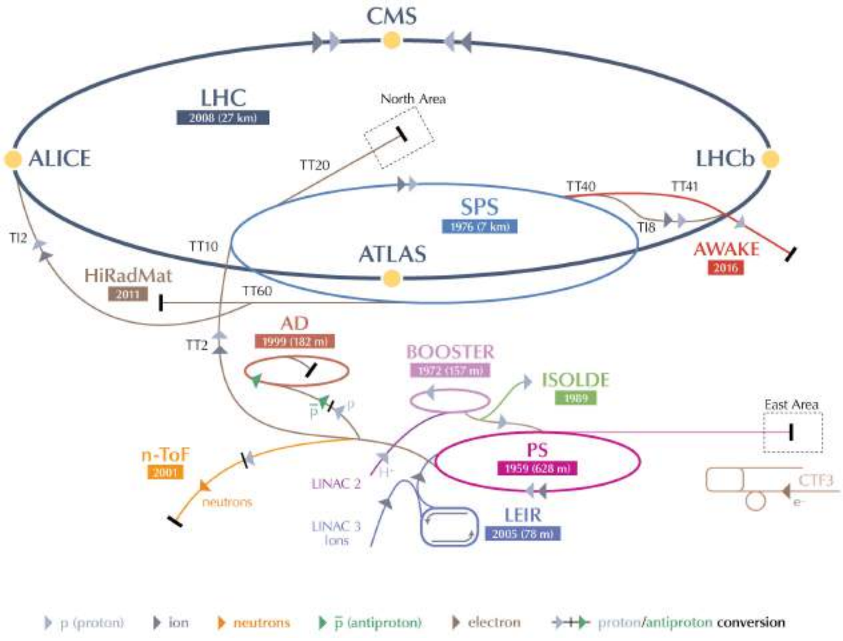


Figure 4.2: The full complex of accelerators that dumps protons into the LHC [40].

4.1.2 LHC experiments

The Large Hadron Collider hosts seven different experiments, each experiment has a different composition, geometry and a different physics program (See Fig. 4.3).

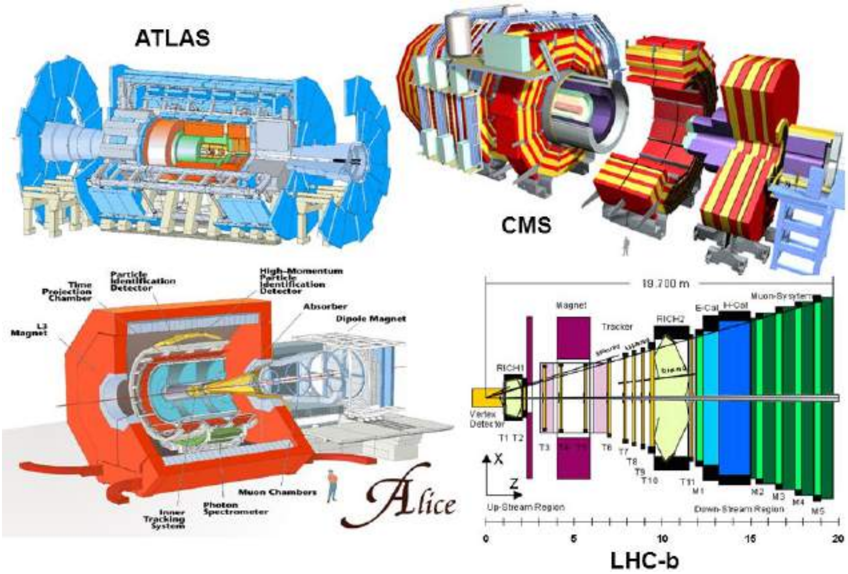


Figure 4.3: LHC main experiments [41].

ALICE (A Large Ion Collider Experiment) [42] is a detector designed for Pb-Pb collisions, in order to study the properties of matter like at very early stage of our universe, in a phase of matter called Quark Gluon Plasma, a state characterized by an high temperature and high density.

ATLAS (A Toroidal Lhc Apparatus) is a general purpose experiment [43], it is characterized by a cylindrical geometry around the beamline. It was designed with the goal of discovery new particles at TeV scale, for this reason its detectors are optimized for the reconstruction of high energy objects with great accuracy. ATLAS uses a toroidal field produced by three sets of air-core toroids complemented by a small solenoid in the inner region.

CMS (Compact Muon Solenoid) is a general purpose detector with a cylindrical geometry around the beamline. It was designed and built with the main goal of discovery new particles up to TeV scale, it is characterized by a specific configuration for the magnetic field, indeed CMS uses the world's largest superconducting solenoid [44].

LHCb is the LHC experiment specialized in studying the physics of heavy quarks with a particular attention to the b quark and its mesons, it will be described in more detail in Sec. 4.2.

TOTEM and LHCf TOTEM (TOTal Elastic and diffractive cross section Measurement) [45, 46] and LHCf (Large Hadron Collider forward) are forward detector located closer to CMS and ATLAS respectively. They are placed 100 m away with respect to the interaction points of the main experiments with the aim of studying diffractive physics happening in the very forward region of the collision. These detectors were put far from the interaction point so that the products of such very forward (with a very small angle with respect to the beam-line) inelastic or elastic collisions may exit the beam-pipe.

MoEDAL (Monopole and Exotics Detectors at the Large Hadron Collider) [47] is a passive detector dedicated to the search of magnetic monopoles or other highly ionizing stable and pseudo-stable massive particles. It is composed of plastic nuclear track detectors attached to the wall and ceiling around LHCb vertex locator. The passage of a massive highly ionizing particle would leave, in plastic nuclear track detectors, damage that can be revealed by controlled etching in hot sodium hydroxide solution.

4.1.3 LHC Run 1 and Run 2

LHC operations are organized in periods of data taking followed by long shutdowns during which maintenance and upgrade can be performed to the accelerator and the detectors. The first period of data taking, known as Run 1, is the period that goes from 2010 until 2013. The nominal center of mass energy during this period was $\sqrt{s} = 7$ TeV during 2011 and $\sqrt{s} = 8$ TeV during 2012. Between 2013 and 2015 the first long shutdown took place, this stopping period of the acceleration complex was used to make interventions on detector components and to improve some of the components in order to be ready for the LHC collisions at 14 TeV. The second phase of data taking period, called Run 2, started in 2015 and will be concluded in 2018. In 2015 and 2016 the nominal center of mass energy was $\sqrt{s} = 13$ TeV. In this work samples coming from both the data taking period are used to perform the analysis. Fig. 4.4 shows the luminosity collected by LHCb experiment from 2010. The luminosity collected is much smaller with respect to the delivered one, nevertheless this was 1.75 times more than the design luminosity for 2011 and twice more for 2012, for this reason in 2011 a luminosity levelling procedure was introduced at the LHCb interaction point [48]. By adjusting the transverse overlap of the beams at LHCb, the instantaneous luminosity could be kept stable to within about 5% during a fill, as illustrated in Fig. 4.5, this allow LHCb to have a lower number of primary vertices per bunch crossing (usually 1 or 2 in LHCb while around 20 in ATLAS and CMS).

4.2 The LHCb detector

LHCb is a dedicated heavy flavour physics experiment at the LHC [48] . Its main goal is to search for indirect evidence of new physics in CP violation and rare decays of

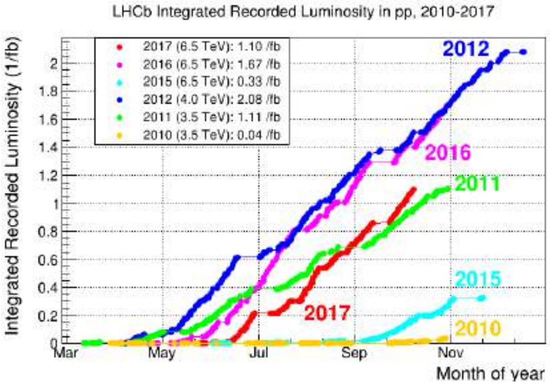


Figure 4.4: Integrated luminosity in LHCb during LHC Run 1 and Run 2. The figure shows the curves for the recorded integrated luminosities [49].

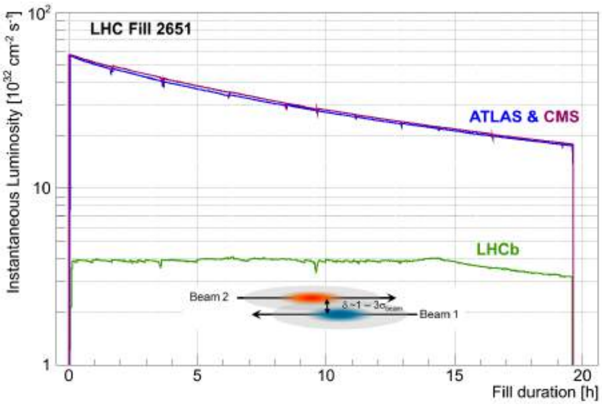


Figure 4.5: Development of the instantaneous luminosity for ATLAS, CMS and LHCb during LHC fill 2651. After ramping to the desired value of $4 \cdot 10^{32} cm^{-2} s^{-1}$ for LHCb, the luminosity is kept stable in a range of 5% for about 15 hours by adjusting the transversal beam overlap. The difference in luminosity towards the end of the fill between ATLAS, CMS and LHCb is due to the difference in the final focusing at the collision points, commonly referred to as the beta function, β^* [48].

beauty and charm hadrons, by looking for the effects of new particles in processes that are precisely predicted in the Standard Model (SM) and by using the distinctive flavour structure of the SM with no tree-level flavour-changing neutral currents. To examine such possibilities, CP violation and rare decays of hadrons containing b and c quarks must be studied with large data samples, using many different decay modes. LHCb is a single-arm spectrometer with a forward angular coverage from approximately 15 mrad to 300 (250) mrad in the bending (non-bending) plane. In proton-proton collisions $b\bar{b}$ and $c\bar{c}$ pairs are produced mainly through gluon fusion and quark-antiquark annihilation with the two incident partons having very different momenta in the laboratory frame. The choice of the detector geometry is driven by the fact that at high energies production of the b and

$\bar{b}$ hadrons is highly correlated, such that they are predominantly produced in the same forward or backward cone. Of course, half of the $b\bar{b}$ quark pairs are directed backward, and therefore outside of detector coverage, but around 25% of the pairs produced are inside the instrumented region. The layout of LHCb detector is shown in Fig. 4.6. Most detector subsystems are assembled in two halves, which can be moved out horizontally for assembly and maintenance purposes, as well as to provide access to the beam-pipe. They are referred to as the detector A and C sides. In order to fulfill its physics program LHCb detector must satisfy some experimental requirements. To separate the decays of interest from the background, both displaced vertex and high transverse momentum signatures are exploited. Excellent vertex resolution is required to measure impact parameters and to achieve a good decay time resolution, and to reject various sources of background. Good momentum and invariant mass resolution are important to minimize combinatorial background and resolve heavy-flavour decays with kinematically similar topologies. Charged particle identification is essential in any flavour physics program, for instance to isolate suppressed decays and for b-quark flavour tagging. Finally, to benefit from the high event rate at the LHC, a high-bandwidth data acquisition system and a robust and selective trigger system are required.

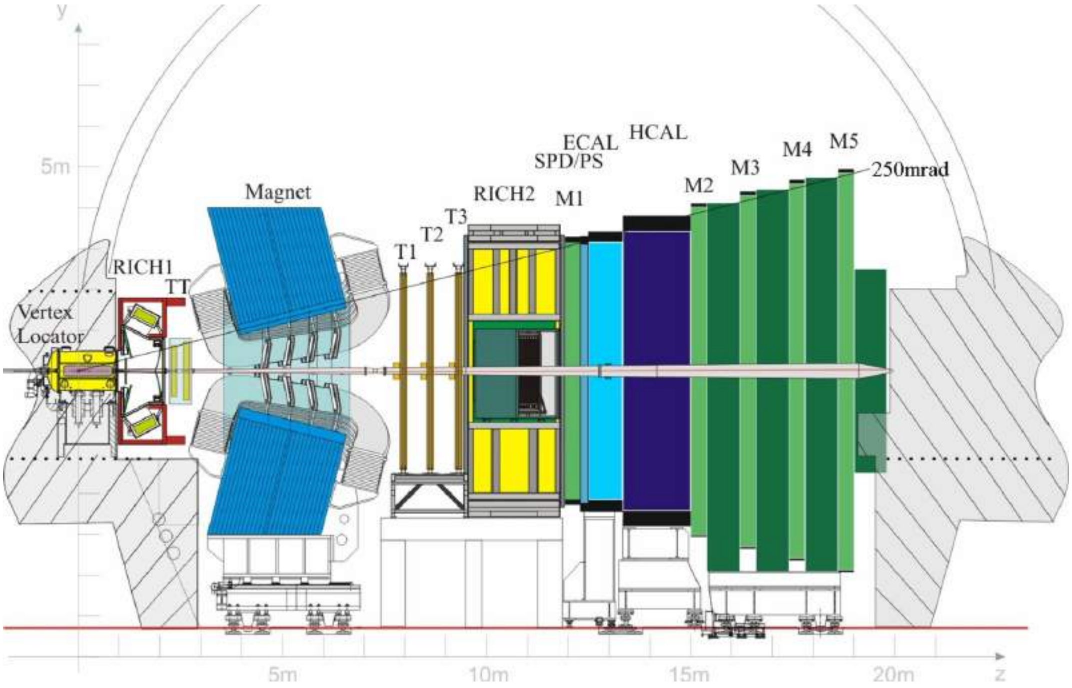


Figure 4.6: View of the LHCb detector [48].

4.2.1 Vertex Locator

The first element of the tracking system consists of the VErtex LOcator (VELO), situated around the interaction region inside a vacuum tank and designed to reconstruct tracks close to the interaction point in order to separate primary vertices, due to proton-proton interactions, from secondary ones, due to the decay of short- lived particles [50]. A schematic view of the VELO structure can be seen in Fig. 4.7. The VELO contains 42 silicon modules divided in two halves (A and C side) and arranged along the beam, each providing a measurement of the r (R sensors) and ϕ (ϕ sensors) coordinates. The layout has been optimized to allow each track in the detector acceptance to traverse at least four modules. The R-sensors consist of circular silicon strips divided in four sectors each 45° wide (See Fig. 4.8). The pitch within a module varies from 38 μm at the inner radius of 8.2 mm, increasing linearly to 102 μm at the outer radius of 42 mm, to keep the strip occupancy approximately constant. The ϕ-sensors contain straight silicon strips and are divided radially into two sections to reduce occupancy and to prevent too large a strip pitch at the outer edge. The strips are not perfectly radial but are inclined by a so called stereo angle of 10°(20°) for the inner (outer) part to improve the pattern recognition capability. The best hit precision measured is around 4 μm for an optimal projected angle of 8°and the smallest pitch of approximately 40 μm. The minimum distance of a track to a primary vertex, the impact parameter, is measured with a resolution of $(15 + 28/p_T)$ μm, where p_T is the transverse momentum, expressed in GeV. For detector safety, the VELO modules are retracted by 29 mm in the horizontal direction during injection of the

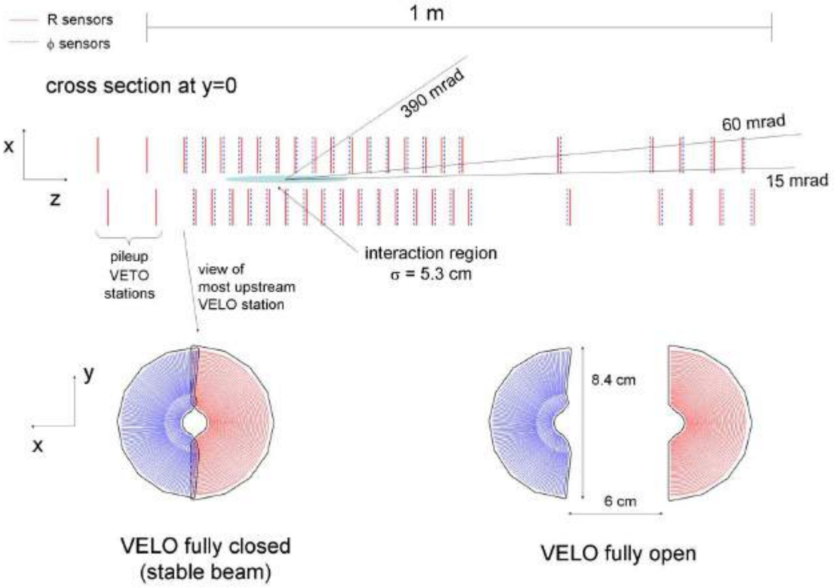


Figure 4.7: Cross section in the (x,z) plane of the VELO silicon sensors, at y = 0, with the detector in the fully closed position. The front face of the first modules is also illustrated in both the closed and open positions [50].

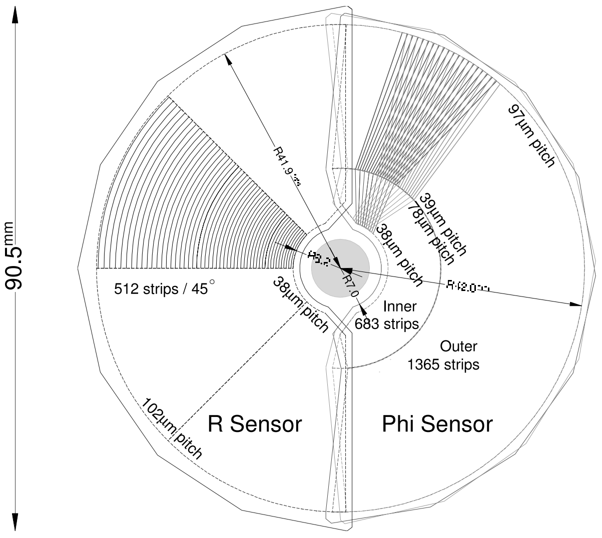


Figure 4.8: Sketch illustrating the rϕ geometry of the VELO sensors. For clarity, only a portion of the strips are illustrated. In the ϕ-sensor, the strips on two adjacent modules are indicated, to highlight the stereo angle. The different arrangement of the bonding pads leads to the slightly larger radius of the R-sensor; the sensitive area is identical [50].

LHC beams and are subsequently moved back close up to 8 mm from the nominal beam position, using a fully automated procedure once stable conditions have been declared.

4.2.2 Tracking System

The second part of the tracking system is composed of four planar tracking stations: the tracking stations T1-T3 positioned downstream and the Tracker Turicensis (TT) upstream of the magnet [50]. A schematic view of the tracking system is visible in Fig. 4.9. The TT is composed of four layers of silicon strips. For the three tracking stations located downstream of the magnet, two different technologies are used, this is due to the fact that the particle flux is much higher in the central region than in the outer part. For the Inner Tracker (IT) silicon microstrips are used, each of the tracking stations has four detection layers in an x-u-v-x arrangement with vertical strips in each of the two x layers, and strips rotated by a stereo angle of $-5°$ and $5°$ in the u and v layers, respectively. The Outer Tracker is a drift-tube gas detector consisting of approximately 200 modules. As a counting gas, a mixture of Argon (70%), CO_2 (28.5%) and O_2 (1.5%) is chosen to guarantee a drift time below 50 ns and a spatial resolution of 200 μm. The magnet spectrometer, required for the momentum measurement of charged particles, is a warm dipole magnet providing an integrated field of about 4 Tm, which deflects charged particles in the horizontal plane. Since the LHCb magnet deflects positive and negative particles in opposite directions in the x-z plane, a difference in performance of the left and right sides of the detector leads to charge detection asymmetries. To reach its design sensitivity in CP violation

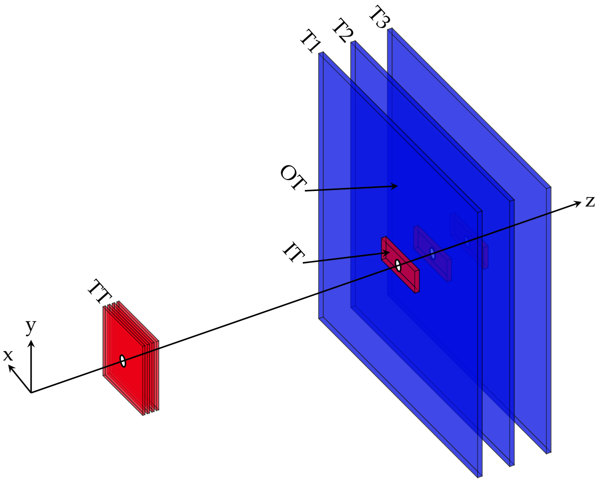


Figure 4.9: Layout of the LHCb tracking system, highlighting the IT and OT portions of T1-T3. The TT and IT collectively form the ST [51].

measurements, LHCb aims to control such detection asymmetries to a precision of 10^{-3} or better. This is achieved by changing the direction of the magnetic field regularly and then combining data sets with different polarity to cancel left-right asymmetries.

4.2.3 Cherenkov Detectors

In a Cherenkov detector a particle travelling faster than the local speed of light in the detector emits a cone of radiation at a certain angle related to the speed of the particle and the refractive index of the medium. Combining this information with the momentum, it is possible to determine the mass of a particle and then to discriminate between kaons, pions and protons. In LHCb two different RICH detectors are used [50, 52], which are installed between VELO and TT and downstream of the tracking stations respectively. RICH1 covers the low momentum charged particle range from about 2 to 60 GeV/c and is filled with C_4F_{10} as radiator (during Run 1 Aerogel was added to the gas mixture). The RICH2 detector covers the high momentum range from about 15 GeV/c to 100 GeV/c, and is filled with CF_4 as radiator. A scheme of these two detectors is shown in Fig. 4.10, while in Fig. 4.11 Cherenkov bands corresponding to the identification of muons, pions, kaons and protons are displayed. Both RICH detectors use two sets of mirrors: spherical primary mirrors reflect the Cherenkov photons onto the plane while secondary mirrors deflect the photons outside the LHCb acceptance where they are detected by Hybrid Photon Detectors (HPD).

4.2.4 Calorimeters

The calorimeter system provides the identification of electrons, photons and hadrons as well as the measurement of their energies and positions, and selects candidates with

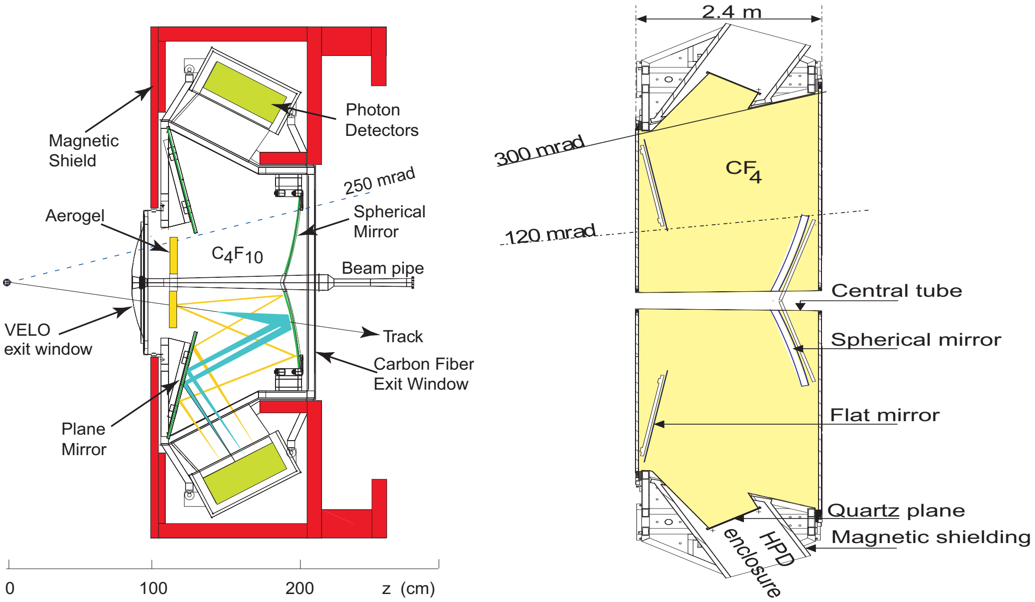


Figure 4.10: Schematic layout of the RICH 1 detector (*left*). (Top view schematic of the RICH 2 detector *right*) [50].

high transverse energy for the first trigger level (L0). It is composed of a Scintillating

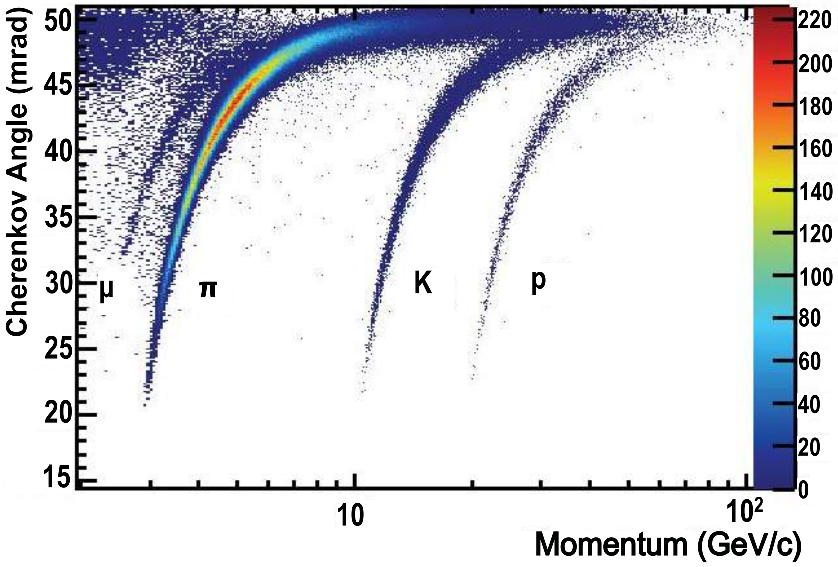


Figure 4.11: Reconstructed Cherenkov angle for isolated tracks, as a function of the track momentum in the C_4F_{10} radiator. The cherenkov bands for muons, pions, kaons and protons are clearly distinguish [48].

Pad Detector (SPD), a Preshower (PS), an electromagnetic calorimeter (ECAL) and an hadronic calorimeter (HCAL) [50], the layout of the detectors can be seen in Fig. 4.12. A 15mm lead converter with a thickness of 2.5 radiation lengths (X_0) is placed between the planes of rectangular scintillating pads of the SPD and the PS. The SPD interacts only with charged particles allowing to discriminate between electrons and photons before showering occurs in ECAL. The lead converter makes photons convert into electron-positron pairs that interact with the PS, allowing to discriminate between photons and neutral pions that do not interact with the PS. The ECAL is made of a sampling scintillator/lead structure with a total thickness of 25 X_0, enough to contain the full electromagnetic shower. It is characterized by an energy resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$. The hadron calorimeter (HCAL) is a sampling device made of iron and scintillating tiles, as absorber and active material, respectively. The main purpose of this calorimeter is to provide information to the hardware trigger, for this reason it requires fast response even without very high energy resolution: $\sigma_E/E \sim 70\%/\sqrt{E} \oplus 10\%$. Given the dimensions of the hadronic showers, the HCAL is segmented into two zones with different lateral dimensions. The thickness of the HCAL is limited to 5.6 nuclear interaction lengths (λ_i) due to space constraints.

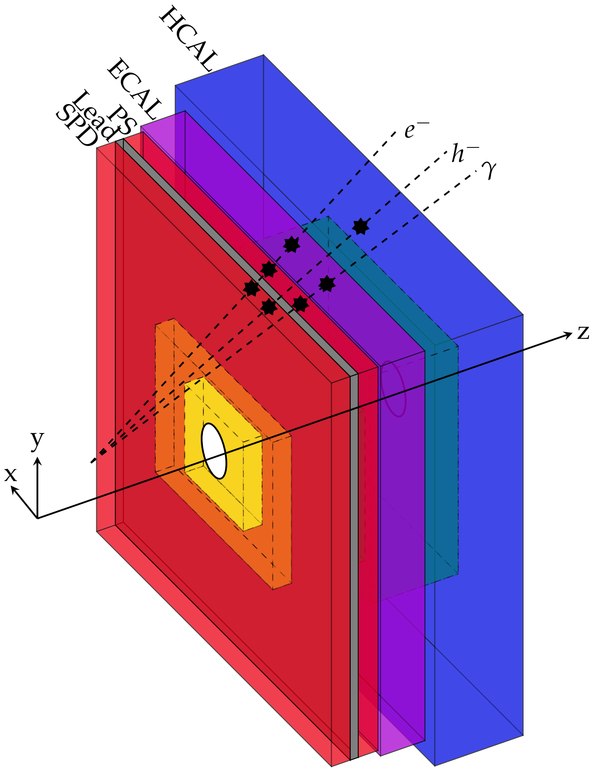


Figure 4.12: Layout of the SPD/PS, ECAL and HCAL showing the segmentation and the interactions of different particle species. The relative dimensions of the ECAL and HCAL are correct, but the the z-scale of the SPD/PS is amplified [51].

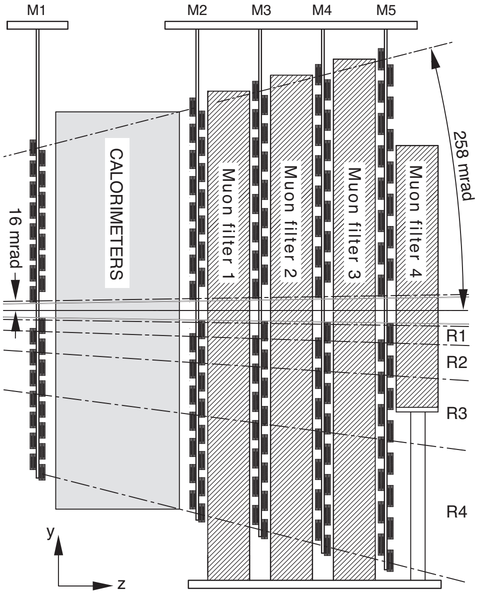


Figure 4.13: Side view of the muon system [50].

4.2.5 Muon System

The muon detection system provides muon identification and contributes to the L0 trigger of the experiment. It is composed of five stations (M1-M5) of rectangular shape provided by Multi Wire Proportional Chambers (MWPC), except in the highest rate region of M1, where Gas Electron Multiplier (GEM) detectors are used [50]. The full system comprises 1380 chambers and covers a total area of 435 m^2. Station M1 is placed in front of the calorimeters and is used to improve the p_T measurement in the trigger. Stations from M2 to M5 are placed downstream of the calorimeters and are interspersed with 80 cm thick iron absorbers to select penetrating muons. The structure of muon stations is shown in Fig. 4.13. The geometry is chosen such that the channel occupancies are comparable in each of the four regions of a given station. The minimum momentum that a muon must have to traverse the five stations is approximately 6 GeV/c.

4.3 The LHCb software

4.3.1 Trigger

The bunch crossing rate at the LHC is designed to be 40MHz, however only a smaller fraction of the collisions produce a $b\bar{b}$ or $c\bar{c}$ pair with all the decay products inside the detector acceptance. Moreover only a tiny fraction of these contain a decay of interest. Also only a part of the events can be stored on disk for later analysis, namely 5kHz in Run 1 and 12.5 kHz in Run 2. For this reason it is necessary to have a trigger which identifies the signal of interest among the large amount of background [50]. The LHCb trigger is composed of three levels:

L0 is the hardware trigger which combines the information coming from the pile-up system, the calorimeters and the muon system. It selects events with high p_T and $E_T = \sqrt{m^2 + p_T^2}$ calorimeter clusters. It reduces the rate to 1 MHz.

HLT is implemented in C++ and runs on a dedicated processor farm constituted of several thousand CPU nodes, it has access to the full event information read out from the LHCb detector. It is divided into two stages:

- **HLT1** is composed of a set of inclusive selection requirements, for example, the presence of a track with high impact parameter or the presence of a muon with high p_T. It reduces the rate down to 40 - 80 kHz.
- **HLT2** can use the full event information to perform a complete reconstruction of the event, the reconstructed tracks are combined to select composite particles. At this stage there are many lines written for specific physics analysis. In the offline selection, trigger signals are associated with reconstructed particles. Selection requirements can therefore be made on the trigger selection itself and on whether the decision was due to the signal candidate (TOS), other particles produced in the pp collision (TIS), or a combination of both. More details can be found in [53].

4.3.2 Reconstruction

The reconstruction procedure consists mainly of two parts: finding of tracks and assignment of particle hypothesis to the found tracks. The tracks, which are the trajectories of charged particles are reconstructed in LHCb combining the hits that is the measured positions of the intersections between tracks and detectors coming from different tracking detectors (VELO, TT, IT and OT). Not all charged particles will leave hits in all subdetectors: some are produced in the backward direction and leave hits only in the VELO before flying outside the LHCb acceptance, other particles with low momentum are deflected outside the acceptance by the magnet, while particles coming from the decay of neutral long-lived particles like K_s^0 or Λ are often produced after the VELO and hence leave hits only in the tracker. The track reconstruction has to take into account these different possibilities, get a high efficiency in finding the tracks and, at the same time, have a low probability of reconstructing spurious tracks, called ghosts. The different kind of tracks categorized in LHCb are shown in Fig. 4.14. The first step of the track reconstruction is the pattern recognition in which, independently in the VELO and in the T stations, sequences of hits are collected together and identified as coming from the same track. These VELO tracks and T tracks are then used as input to find long, upstream and downstream tracks. The long tracks are found either by extrapolating the VELO tracks into the T stations and looking for matching hits, or by matching directly the VELO tracks with T tracks, then

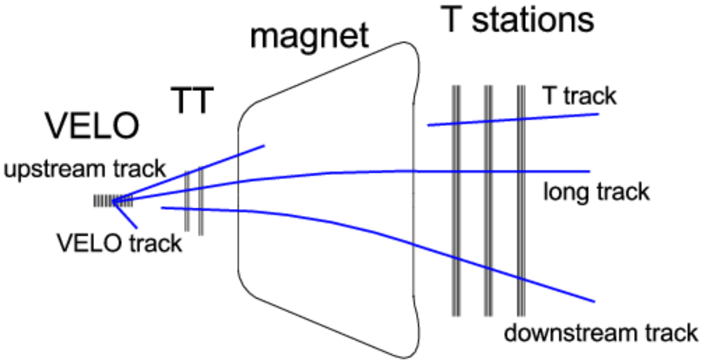


Figure 4.14: Track types reconstructed in LHCb by the track finding algorithms [54].

TT hits are added. Upstream tracks are found by extrapolating VELO tracks into the TT while the Downstream tracks combine T tracks with TT information. The second step of the track reconstruction is the track fitting that, in LHCb, is done using a Kalman filter [55], which takes into account effects from multiple scattering and energy loss due to ionization. Using the output of a neural network that takes as an input various quality variables like the track χ^2 and the number of hits in each subdetector, ghost tracks are then removed.The last step of the track reconstruction is the clone killing which consists of removing tracks that are also subtracks of other tracks, for example a VELO track used to build a long track. The tracking efficiency for charged tracks that pass through the full tracking system varies as a function of the kinematics of the track and the occupancy of the detector and is above 95 %. The particle identification (PID) of charged hadrons is performed using the allocation of rings in the RICH detectors to the tracks and calculating likelihoods for the different particle hypotheses. Maximizing this likelihood, particles identities are then assigned. The estimation for a particle to be more likely a kaon or a pion can be done by taking the difference between the logarithms of the likelihood under the two different mass hypotheses for that particle. In LHCb these variables are called PIDK, PIDp etc. where the first mass hypothesis tested is that of a kaon or a proton while the second mass hypothesis is always a pion. A second class of commonly used PID variables are called ProbNNx. They are the output of a neural network that takes as an input the various PIDx and also complementary information from all the subdetectors including the tracking detectors. The output is a variable between zero and one that can be interpreted as absolute probability for a particle to have a certain PID. On average the efficiency to correctly identify a kaon is $\sim$ 95 % with a misidentification probability to identify a pion as a kaon of $\sim$ 5 % [48].

4.3.3 Stripping

Once tracking and PID are performed the full decay chain can be fitted. However, the data size after reconstruction is huge. For this reason a centralized selection called Stripping is used to handle it. Stripping lines are a set of requirements characteristic for each decay mode, which are applied on data to select them, this is done inside the LHCb framework called DaVinci [56]. Many selection steps can be shared between various stripping lines, this allows to save resources and computing time. Only data which are selected by a stripping line can be analyzed offline by the final users.

4.3.4 Monte Carlo simulation

Monte Carlo simulations (MC) are a fundamental part of any data analysis. Indeed MC samples can be used in different points of the analysis, like for example for the calculation of efficiencies, to find appropriate parametrizations to model data distributions or for the development of selection strategies. The main advantage is that the simulated samples are very large, typically considerably larger than the real data samples. It is important that simulation samples are as similar as possible with respect to real data, constantly checks to evaluate this similarity are done and in case a deviation between data and MC is found, methods to compensate it are applied. A known case is the one related to particle identification system performance, which is overestimated and for this reason has to be corrected applying a data driven resampling procedure, which will be further discussed in Ch. 7. The production of MC samples is divided into two phases: the generation and the simulation. In the first phase the proton proton interaction is simulated using the event generator Pythia [57, 58], the decays are then simulated using EvtGen [59] and the final state radiation is generated using Photos [60]. In the second step, using Geant4 [61, 62], the interaction of the generated particles with the detector is simulated. After that, using Boole [63], the simulation of the detector response to produce digitalized data is performed. From this point on the processing is the same as the one of real collision data, starting with the trigger implemented in Moore [64] , followed by the reconstruction done using Brunel [65] and the stripping performed by the DaVinci package [56]. Along with the reconstructed properties of the particles, the true information is available in MC samples. This is useful in comparison between MC and data, to study resolution and acceptance effects, and flavour tagging as will be discussed in Ch. 6, Ch. 7 and Ch. 5.

Chapter 5

Flavour tagging calibration using $B^0 \to D_s^+ D^{-(*)}$ decays

Measurements of *CP* violation in $B^0 \to D^+D^-$ and $B^0 \to D^{*\pm}D^{\mp}$ decays, require the determination of the B^0 candidate flavour at production. The method to determine the initial flavour of a reconstructed candidate, that is whether it contains a b or a $\overline{b}$ quark at production, is called *flavour tagging*.

5.1 Flavour tagging algorithms

The LHCb collaboration has developed several flavour-tagging algorithms, which can be classified as *same side taggers* (SS taggers) and *opposite side taggers* (OS taggers). A schematic representation of the taggers that can be used for tagging B_d^0 mesons is shown in Fig. 5.1. The SS taggers infer the production flavour of the signal B meson by selecting charged particle candidates that have a high chance of being remnants of the hadronization process of the B candidate [66]. For B^0 mesons, the same side pion tagger, which exploits $\pi^{\pm}$ mesons produced in the hadronization of the B^0 meson, and the same side proton tagger, which looks for co-produced protons, have been developed. The response of the two taggers can be combined into a common SS response.

In contrast, the OS taggers exploit the predominant production process of B mesons via $b\overline{b}$ quark pair production [67]. They partially reconstruct the decay of the b-hadron produced along with each reconstructed signal B meson, and infer its initial flavour. Several OS taggers have been developed inside LHCb collaboration, even in this case these taggers can be combined, and the combination of the OS kaon, muon, electron, charm, and vertex charge tagging algorithms represents the current standard OS combination. An additional OS tagger, the OS Charm tagger, can be exploited, and can be combined with the OS standard combination.

Given a reconstructed candidate, each flavour tagging algorithm provides a flavour tag decision d and a probability η for the the tag to be wrong. This mistag probability η ranges from 0 to 0.5 and is based on the output of multivariate classifiers, which are

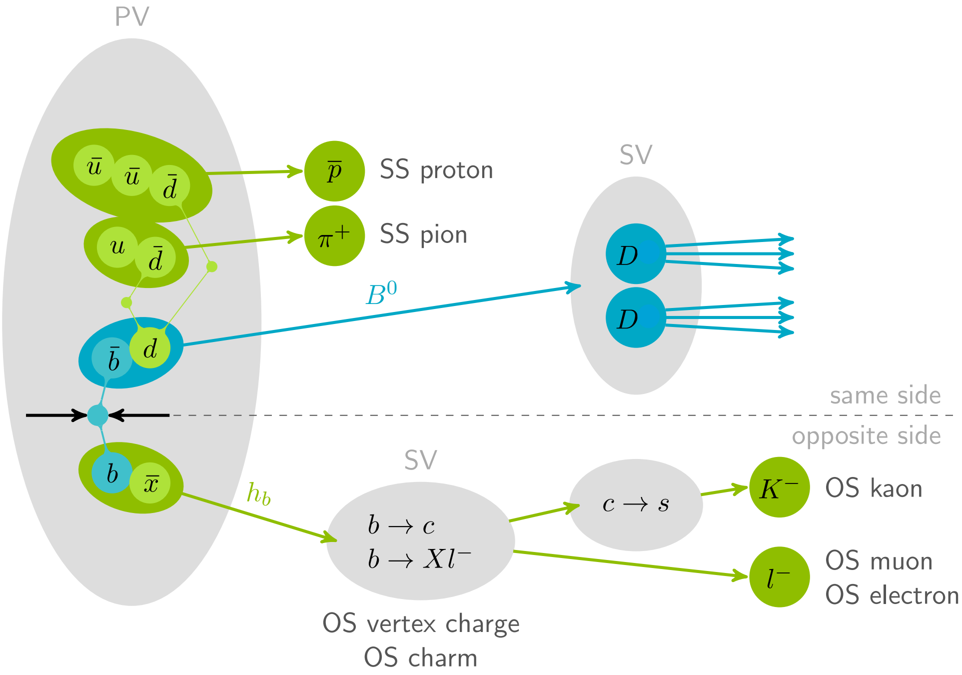


Figure 5.1: Tagging algorithms used in the analysis of $B^0 \to D^+D^-$ and $B^0 \to D^{*\pm}D^{\mp}$ decays.

trained on datasets of flavour specific decays, and combine several kinematic and geometric information on the tagging particle(s) and the event. The flavour tag decision can take the values $d = +1$ for an initial B^0, $d = -1$ for an initial $\overline{B}^0$, and $d = 0$ when no decision on initial flavour can be assigned. Tags are based on the charge of the tagging particle(s) and the output of the multivariate classifiers.

A more detailed description of flavour tagging at LHCb can be found in [68, 69].

5.1.1 Performance characteristics

The performances of flavour tagging algorithms are not perfect, for this reason a set of quantities that reflect their performance can be defined. Considering N reconstructed candidates of which N_U are untagged, N_W are incorrectly tagged, and N_R are correctly tagged, the tagging efficiency can be defined as

$$\varepsilon_{\text{tag}} = \frac{N_R + N_W}{N_R + N_W + N_U}, \tag{5.1}$$

while the mistag probability can be written as

$$\omega = \frac{N_W}{N_R + N_W}. \tag{5.2}$$

The mistag probability is related to the tagging dilution, $D = 1 - 2\omega$, which takes a value of 1 in case of perfect tagging, and 0 in case of random tagging. The quantity that is interpreted as the figure of merit as it states the effective loss in statistics compared to a perfect tagged sample, and that hence should be maximized, is the effective tagging efficiency, called also tagging power:

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}(1-2\omega)^2 = \varepsilon_{\text{tag}}D^2. \tag{5.3}$$

5.1.2 Calibration of the tagging output

The output of the flavour tagging algorithms is the result of training multivariate classifiers using datasets of flavour specific B decays, and transforming the classifier output into mistag estimates η through regression. However the training and validation samples are different from the signal sample used in the *CP* measurement (e.g. in terms of trigger and selection criteria that affect the distribution of the MVA's input variables), for this reason the output of the flavour tagging algorithms needs to be calibrated. This is done using control samples of flavour specific decays, on these channels calibration functions $\omega(\eta)$ are extracted in order to transform the mistag estimates η of the algorithm to the mistag probability ω measured in the control sample.

A common choice for the calibration function is a linear function (see Fig. 5.7),

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

The use of the arithmetic mean $\langle\eta\rangle$ of the η distribution aims for a decorrelation of p_0 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 is 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 $\overline{B}^0$. These tagging asymmetries can dilute or enhance the observed raw tag asymmetry and need to be corrected for. The asymmetries of the mistag probability, defined as the difference of the tagging calibration parameters p_0 and p_1 for initial B^0 and $\overline{B}^0$, can be parametrized by two independent calibration functions

$$\begin{aligned}\omega^{B^0}(\eta) &= p_0^{B^0} + p_1^{B^0}\left(\eta - \langle\eta\rangle\right)\ ,\\ \omega^{\overline{B}^0}(\eta) &= p_0^{\overline{B}^0} + p_1^{\overline{B}^0}\left(\eta - \langle\eta\rangle\right)\ .\end{aligned} \tag{5.5}$$

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

$$p_i^{B^0} = p_i + \frac{\Delta p_i}{2}\ ,\quad p_i^{\overline{B}^0} = p_i - \frac{\Delta p_i}{2}\ . \tag{5.6}$$

The difference of a B^0 meson mistag with respect to the anti-meson one can be expressed as

$$\begin{aligned}\Delta\omega(\eta) &= \omega^{B^0}(\eta) - \omega^{\overline{B}^0}(\eta)\\ &= \Delta p_0 + \Delta p_1\left(\eta - \langle\eta\rangle\right)\ .\end{aligned} \tag{5.7}$$

5.1.3 Strategy

In $B^0 \to D^+D^-$ analysis, the OS combination and the SS combination, as described in Sec. 5.1, are used. The calibration parameters are determined from a weighted, unbinned maximum likelihood fit to the distribution of the decay time t, using the control channel $B^0 \to D_s^+D^-$ coming from Run 1. Using the control channels $B^0 \to D_s^+D^-$ and $B^0 \to D^{*+}D_s^-$, coming from both Run 1 and Run 2 data samples, the calibration parameters are determined from a weighted, unbinned maximum likelihood fit to the distribution of the decay time t and from a weighted sample, making use of the Espresso Performance Monitor (EPM) tool [70], respectively. These calibration parameters are used in the $B^0 \to D^{*\pm}D^{\mp}$ analysis, where $D^{*\pm} \to D^0\pi^{\pm}$, $D^0 \to K^{\mp}\pi^{\mp}\pi^{\pm}\pi^{\pm}$ and $D^0 \to K^{\mp}\pi^{\pm}$, again the OS along with the SS tagger combinations are exploited. A detailed description is given in Sec. 5.2 and Sec. 5.3 respectively. Additionally studies in the signal channel $B^0 \to D^+D^-$ are presented in Sec. 5.6, while those relative to $B^0 \to D^{*\pm}D^{\mp}$ can be found in Sec. 5.7. Here the study of possible correlations between the mistag estimates of the OS and SS taggers, and an evaluation of the tagging performance after applying the calibrations are presented.

5.2 Studies in the control channel $B^0 \to D_s^+D^-$

The choice of control channels strongly depends on the signal channel and the employed tagging algorithms. In general flavour specific decays with large branching fractions w.r.t. the signal channel are chosen as control channels to keep the statistical uncertainties on the calibration parameters small w.r.t. the uncertainties on the *CP* asymmetries in the signal channel. Moreover, these control samples should resemble the signal sample in terms of kinematics, event properties, and selection criteria. Otherwise, additional assumptions on the portability of the calibration parameters are required, and systematic uncertainties on the portability need to be evaluated.

The decay channel $B^0 \to D_s^+D^-$ has been chosen as one of the flavour tagging control and calibration channels, as its similarity should lead to a good portability of the calibration results to the signal channels $B^0 \to D^+D^-$ and $B^0 \to D^{*\pm}D^{\mp}$.

5.2.1 Selection

The selection closely follows the selection described in Sec. 6.1. The differences are that the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state is selected, the mass of $KK\pi$ is chosen around the D_s^+ instead of the $D^{\pm}$ mass, and the vetoes are applied only to $D^{\pm}$ candidates and not to D_s^+ ones. This leads to the following selection:

Stripping In the procedure used to make tuples from raw data *i.e.* stripping, the daughter particles (kaons and pions) of the D mesons have to fulfill a set of criteria which are listed in Table 6.2. The further requirements on the D combination are given in Table 6.3, where the two D meson candidates have to form a common vertex. The stripping requirements on the B^0 combination are listed in Table 6.4.

Preselection At this stage of the selection the mass window around the $D_s^{\pm}$ and $D^{\pm}$ mass is tightened to be $\pm 40\,\mathrm{MeV/c^2}$ with respect to the nominal masses. The complete set of cuts which are implemented in this part of the selection are listed in Table 6.5.

Cuts against (partially) charmless backgrounds Three kinds of selection requirements are applied, to reduce the amount of (partially) charmless contributions, consisting of $B^0 \to hhhhhh$ and $B^0 \to Dhhh$, mainly due to $B^0 \to D^{\pm}K^{\mp}K^{\pm}\pi^{\pm}$. First the window around the invariant mass of both D mesons is tightened to $\pm 25\,\mathrm{MeV/c^2}$. Moreover the χ^2 of the flight distance of both D mesons w.r.t. the B decay vertex has to be be larger than 2. Finally, the $D^{\pm}$ decay time significance and the $D_s^{\pm}$ decay time significance w.r.t. the B^0 decay vertex ($t_{D^{\pm}}/\sigma_{t_{D^{\pm}}}$) are required to be greater than 3.

$D_s^{\pm}$ mass veto For $D^{\pm}$ mesons which are reconstructed in the $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ final state a $K \to \pi$ mis-ID can lead to background contributions from $D_s^{\pm} \to K^{\mp}K^{\pm}\pi^{\pm}$. To reduce these $D_s^{\pm}$ contributions the kaon mass hypothesis is assigned to the pion with the higher transverse momentum. If $\texttt{ProbNN}K > \texttt{ProbNN}\pi$ for the pion with the higher p_{T} and the invariant mass of the three hadrons is compatible with the $D_s^{\pm}$ mass of $M_{D_s^{\pm}} = 1968.30\,\mathrm{MeV/c^2}$ [28], within $\pm 25\,\mathrm{MeV/c^2}$ the candidate is rejected.

$\Lambda_c^{\pm}$ mass veto To reduce $p \to \pi$ mis-ID the proton mass hypothesis is assigned to the pion with the higher p_{T} of the $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ final state and the invariant mass of the $Kp\pi$ combination is calculated. If $\texttt{ProbNN}p > \texttt{ProbNN}\pi$ for the pion with the higher p_{T} and the invariant mass of the $Kp\pi$ combination is compatible with the $\Lambda_c^{\pm}$ mass of $M_{\Lambda_c^{\pm}} = 2286.46\,\mathrm{MeV/c^2}$ [28], within $\pm 25\,\mathrm{MeV/c^2}$ the candidate is rejected.

MVA classifier The BDT for the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state of the $B^0 \to D^+D^-$ selection is used without adaptation, and the same cut at -0.05 found in $B^0 \to D^+D^-$ BDT cut optimization is applied.

Final selection In the final selection step, the range of the invariant $D_s^+D^-$ mass is restricted to $5220\,\mathrm{MeV/c^2}$ to $5500\,\mathrm{MeV/c^2}$ and the range of the decay time to 0.25 ps to 10.25 ps. An analysis of all candidates surviving the formerly described selection reveals that for some events there is more than one candidate, and since a-priori they have the same possibility to be the true signal candidate, these multiple candidates are removed randomly [71]. After all selection steps, 18477 events remain in the data sample, while 16966 candidates remain in the region corresponding to 2σ of the width of the mass peak, of these 16736 are signal candidates (see Sec. 5.2.2), corresponding to a purity of 98%.

5.2.2 Fit Model

As the $B^0 \to D_s^+ D^-$ is a neutral mode, the final state flavour alone does not allow to recover the information on the production flavour, for this reason a mixing analysis is required. A fit to the invariant mass distribution of the $B^0 \to D_s^+ D^-$ candidates is used to extract ${}_s\mathcal{W}eights$ via the *sPlot* method [72]. These are then used in a simultaneous fit of the distributions of decay time, final state, tags, and mistag estimates.

Mass Fit The mass fit is performed considering the final state $KK\pi K\pi\pi$ in the mass range 5220 MeV/c² to 5500 MeV/c². This removes backgrounds with D^* mesons and leaves enough candidates in the upper mass sideband to determine the shape of the combinatorial background.

For the mass fit the extended maximum likelihood method is used, the extended likelihood function is defined as in Eq. (5.8)

$$\mathcal{L}(\vec{\lambda}; m) = \frac{e^{-N} N^n}{n!} \prod_s \prod_{i=1}^{N^s} \mathcal{P}^s(m_i; \vec{\lambda}_s) \ . \tag{5.8}$$

Here, m_i is the invariant mass of the B^0 candidate and $\vec{\lambda}_s$ are the parameters for which the optimal values that maximize this function have to be found. In practice, an equivalent approach is performed. The negative log-likelihood $-\ln \mathcal{L}$ is minimized numerically using MINUIT implemented by the RooFit package [73]. The index s sums over the different simultaneous categories which contain N^s events. These add up to the number of all observed events $n = \sum_s N^s$. The Poisson term in the likelihood reflects the probability that n events are observed, when N are expected. The expected number of events N is then estimated by the fit.

The PDF includes three components: the $B^0 \to D_s^+ D^-$ signal, the $B_s^0 \to D_s^- D^+$ background, and the combinatorial background. The B^0 signal component is parametrized with a sum of two crystal-ball PDFs [74], with the same mean μ_{B^0} but different width parameters σ_i and tails to opposite directions ($\alpha_1 > 0$, $\alpha_2 < 0$). The parameter n is fixed to 10. The parameters α_1 and α_2 of the power law functions are taken from simulation, the values are listed in Table 5.1. The fraction $f_{B^0 \to D_s^+ D^-}$ of both crystal-ball PDFs is fixed to 0.5.

Table 5.1: Mass parameters fixed to values from signal MC fit.

Parameter	Fixed Value
$\alpha_{1,B^0 \to D_s^+ D^-}$	1.28
$\alpha_{2,B^0 \to D_s^+ D^-}$	-1.32
$f_{B^0 \to D_s^+ D^-}$	0.5
$n_{B^0 \to D_s^+ D^-}$	10

A fit to simulated events for the signal is shown in Fig. 5.2, even if the tails are not perfectly described, this doesn't affect the data mass fit, since this contribution is very small, moreover the mass model is included in the systematics.

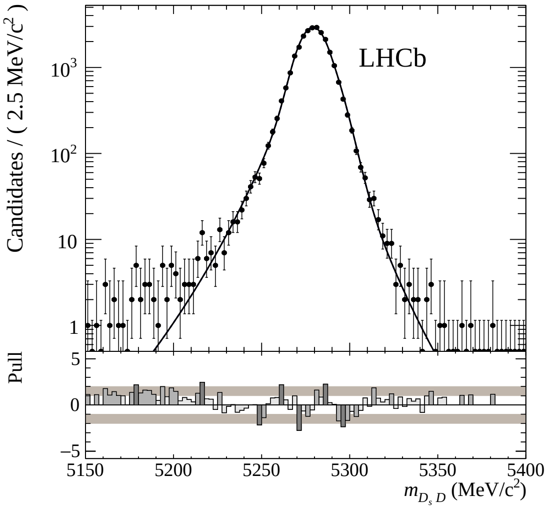


Figure 5.2: Fit of B^0 signal to Run1 simulated events.

Along with the $B^0 \to D_s^+ D^-$ signal the decay of the heavier B_s^0 to the same final state is also present. It is parametrized with the same PDF as the B^0 signal component sharing the width and tail parameters. The difference between the peak positions is fixed to the world average $\Delta m_{B_s^0 - B^0} = \mu_{B_s^0} - \mu_{B^0} = 87.35\,\mathrm{MeV/c^2}$ [28]. The reconstructed mass PDF of the combinatorial background component is modelled with an exponential function

$$\mathcal{P}_{\mathrm{Bkg}}(m;\beta) \propto \mathrm{e}^{\beta m} \ . \tag{5.9}$$

In Table 5.2 the results of the floating shape parameters of the mass fit are shown along with the fit results for the yields, while Fig. 5.3 shows the distribution of the data overlaid with the PDF projections.

Decay Time Fit After obtaining the $_s\mathcal{W}eights$ from the fit to the invariant mass distribution described in the previous section, the calibration parameters are determined from a weighted, unbinned maximum likelihood fit to the distribution of the decay time t, the final state r, the tag d, and the predicted per-event mistag probability η with the PDF

$$\mathcal{P}\,(t,d,r) = \epsilon(t) \cdot \left(\mathcal{R}(t-t') \otimes \mathcal{B}(t',d,r,\eta)\right) \ . \tag{5.10}$$

Here, $\epsilon(t)$ is the acceptance function for reconstructed decay times t, it is parametrized using cubic splines obtained from a data-driven approach. The time range considered is $[0.25;10.25]\,\mathrm{ps}$, and four spline coefficients are used in the parameterization,

Table 5.2: Results of the floating shape parameters in the mass fit to $B \to D_s D$ Run 1 and Run 2 data samples.

Sample	Parameter	Fitted Value
Run 1	$\mu_{B^0 \to D_s^+ D^-}$ (MeV/c^2)	$5278.9 \pm ^{+0.08}_{-0.08}$
	$\sigma_{1,B^0 \to D_s^+ D^-}$ (MeV/c^2)	$7.55 \pm ^{+0.21}_{-0.19}$
	$\sigma_{2,B^0 \to D_s^+ D^-}$ (MeV/c^2)	$9.73 \pm ^{+0.24}_{-0.25}$
	β (1/MeV/c^2)	-0.0068 ± 0.0004
	$N_{B^0 \to D_s^+ D^-}$	16736 ± 134
Run 2	$\mu_{B^0 \to D_s^+ D^-}$ (MeV/c^2)	$5278.5 \pm ^{0.09}_{0.09}$
	$\sigma_{1,B^0 \to D_s^+ D^-}$ (MeV/c^2)	$7.72 \pm ^{0.25}_{0.22}$
	$\sigma_{2,B^0 \to D_s^+ D^-}$ (MeV/c^2)	$11.05 \pm ^{0.32}_{0.35}$
	β (1/MeV/c^2)	-0.00057 ± 0.00032
	$N_{B^0 \to D_s^+ D^-}$	13757 ± 130

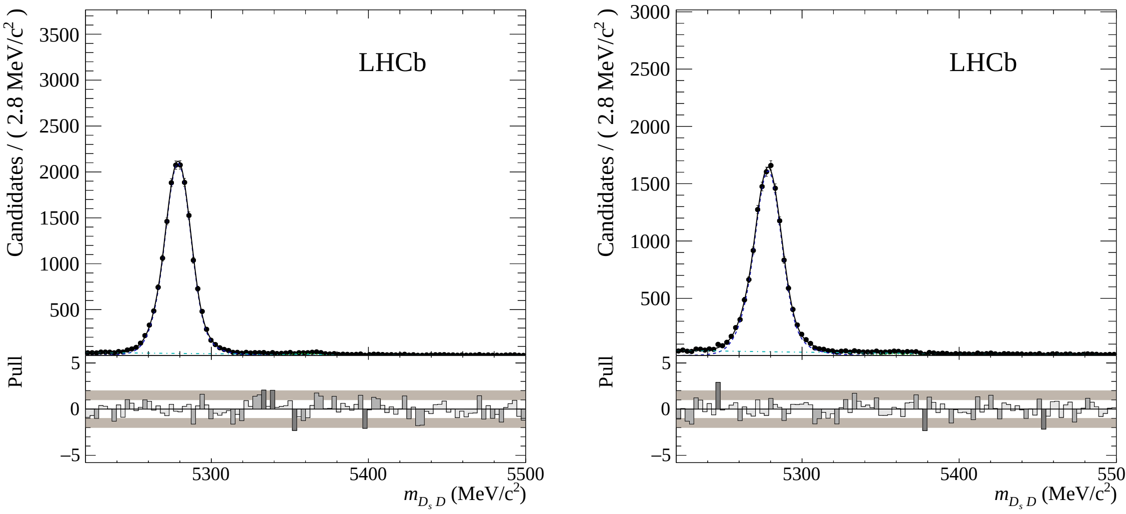


Figure 5.3: Fit to the B^0 mass, in order to extract ${}_s\mathcal{W}eights$. The solid line is the PDF projection, the blue dotted line represent the signal component, while the dotted cyan and green lines represents the main backgrounds. The fit is performed on Run 1 (left) and Run 2 (right) data samples respectively.

where the second to last is fixed to 1 for normalization. The splines knot positions $\{0.25, 0.8, 2.0, 10.25\,\text{ps}\}$ are chosen as in the nominal fit to the $B^0 \to D^+D^-$ mode, see Sec. 6.3.

The $\mathcal{R}(t - t')$ term represents the decay time resolution. As the B^0 mixing period is very large compared to the decay time resolution, an under- or overestimation of the decay time resolution at the level of 50% has negligible effects on the results. The resolution parameters are fixed from results obtained in a $B^0 \to D_s^+ D^-$ Monte Carlo sample. To account for different sources introducing the decay time resolution an effective model consisting of two Gaussians with per-event widths is used. Besides this common resolution effect the decay time resolution model is also supposed to describe the effect of events matched to the wrong PV. This can cause a large deviation between the correct and the reconstructed decay time. The wrong PV component is parametrized with a broad Gaussian distribution using the same mean μ_t as the other two Gaussians and one width parameter σ_{PV} [75]. The complete parametrization of the resolution model follows the one used for $B^0 \to D^+ D^-$ and is given by

$$\begin{aligned}\mathcal{R}(t - t_{\mathrm{true}}|\sigma_t) = &\sum_{i=1}^{2} g_i \cdot \frac{1}{\sqrt{2\pi}(c_i + b_i \cdot \sigma_t)} \exp\left(-\frac{(t - t_{\mathrm{true}} - \mu_t)^2}{2(c_i + b_i \cdot \sigma_t)^2}\right) \\ &+ f_{\mathrm{PV}} \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{PV}}} \exp\left(-\frac{(t - t_{\mathrm{true}} - \mu_t)^2}{2\sigma_{\mathrm{PV}}^2}\right) .\end{aligned} \tag{5.11}$$

The first two Gaussian components have different calibration parameters b_i and c_i and thus different widths. Together with the fraction f_{PV} of the wrong PV component the fractions of the two Gaussian components g_1 and g_2 sum up to unity. The wrong PV component, the fraction of wrong PV events, and the fractions for the two Gaussians are fixed to $B^0 \to D^+ D^-$ values, the results for the remaining parameters obtained from the fit on $B^0 \to D_s^+ D^-$ MC are reported in Table 5.3 and the corresponding plot is shown in Fig. 5.4.

Table 5.3: Fit parameters of decay time resolution function determined on $B^0 \to D_s^+ D^-$ Run 1 signal MC.

Parameter		Value
μ_t	(ps)	$-0.003\,75 \pm 0.00029$
b_1		0.866 ± 0.019
c_1	(ps)	$0.012\,38 \pm 0.00066$
b_2		1.299 ± 0.063
c_2	(ps)	0.0245 ± 0.0021
g_2		0.14 (fixed)
σ_{PV}	(ps)	0.16 (fixed)
f_{PV}		0.0029 (fixed)

Before being multiplied with the acceptance function, the resolution function is convolved with the B physics PDF, which describes the distribution of true decay time t', tag d, and final state r. Ignoring negligible CP asymmetries, it can be written as

$$\mathcal{B}(t', d, r, \eta) = e^{-t'/\tau} \cdot \Big(a(d, r, \eta) \cdot \cosh(\Delta\Gamma_d/2t') + b(d, r, \eta) \cdot \cos(\Delta m t')\Big) , \tag{5.12}$$

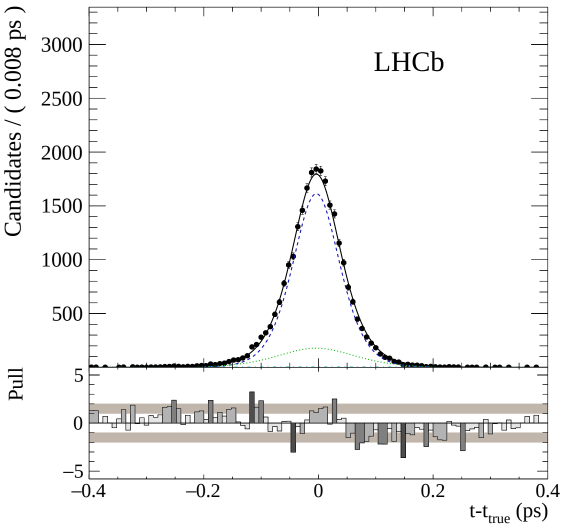


Figure 5.4: Fit of per event resolution model to the difference of true and reconstructed decay time in Run 1 signal MC. The black solid line is the projection of the full PDF. The blue dashed and the green dotted line represent the two per-event components and the turquoise dashed-dotted line shows the wrong PV component.

where τ is the B^0 lifetime and Δm ($\Delta\Gamma$) is the mass (decay width) difference of the heavy and light mass eigenstates. The coefficients a and b are parametrized as

$$\begin{aligned} a(d,r,\eta) =&\Big(1 - r\cdot A_{\text{det}}\Big)\cdot\Big(1-\frac{A_{\text{sl}}^{\text{d}}}{2}\cdot\frac{1+r}{2}\Big)\cdot\Bigg(\Big(1+A_{\text{prod}}+\Delta\varepsilon_{\text{tag}}\Big)\cdot\Big(\frac{1-d}{2}+d(\omega+\Delta\omega)\Big)\\ &+\Big(1-A_{\text{prod}}-\Delta\varepsilon_{\text{tag}}\Big)\cdot\Big(\frac{1+d}{2}-d(\omega-\Delta\omega)\Big)\cdot\Big(1+\frac{A_{\text{sl}}^{\text{d}}}{2}\Big)\Bigg)\ , \end{aligned} \tag{5.13}$$

$$\begin{aligned} b(d,r,\eta) =&-r\Big(1 - r\cdot A_{\text{det}}\Big)\cdot\Big(1-\frac{A_{\text{sl}}^{\text{d}}}{2}\cdot\frac{1+r}{2}\Big)\cdot\Bigg(\Big(1+A_{\text{prod}}+\Delta\varepsilon_{\text{tag}}\Big)\cdot\Big(\frac{1-d}{2}+d(\omega+\Delta\omega)\Big)\\ &-\Big(1-A_{\text{prod}}-\Delta\varepsilon_{\text{tag}}\Big)\cdot\Big(\frac{1+d}{2}-d(\omega-\Delta\omega)\Big)\cdot\Big(1+\frac{A_{\text{sl}}^{\text{d}}}{2}\Big)\Bigg)\ . \end{aligned} \tag{5.14}$$

Where the tagging efficiency asymmetry is defined as:

$$\Delta\varepsilon_{\text{tag}} = \frac{{\varepsilon_{\text{tag}}}^{B^0} - {\varepsilon_{\text{tag}}}^{\bar{B}^0}}{{\varepsilon_{\text{tag}}}^{B^0} + {\varepsilon_{\text{tag}}}^{\bar{B}^0}}\ , \tag{5.15}$$

while the mistag probability ω, and the asymmetry of the mistag probability $\Delta\omega$ are defined according to Eq. (5.4) and Eq. (5.7). Other asymmetries need to be considered, first the detection asymmetry defined as

$$A_{\text{det}} = \frac{\varepsilon_{\text{det}}(D_s^+ D^-) - \varepsilon_{\text{det}}(D_s^- D^+)}{\varepsilon_{\text{det}}(D_s^+ D^-) + \varepsilon_{\text{det}}(D_s^- D^+)} \ , \tag{5.16}$$

and the production asymmetry

$$A_{\text{P}} = \frac{R_{\overline{B}^0} - R_{B^0}}{R_{\overline{B}^0} + R_{B^0}}, \tag{5.17}$$

where $R_{\overline{B}^0}$ and R_{B^0} are the production rates of $\overline{B}^0$ and B^0, and the flavour-specific semileptonic asymmetry can be written as

$$A_{\text{sl}}^{\text{d}} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \ . \tag{5.18}$$

In the decay time fit, all parameters except for the calibration parameters and the asymmetry of the tagging efficiency are fixed to the values listed in Table 5.4. The B^0 oscillation frequency and the B^0 lifetime are fixed to the values in [28]. The value of the production asymmetry is obtained from the LHCb measurement [76], applying a weighting procedure based on transverse momentum and pseudorapidity of the data set, while the values of the flavour-specific semileptonic asymmetry and the detection asymmetry are fixed to the result reported in [77]. The statistical precision on these values is used as variation in the study of systematics.

Table 5.4: Fixed parameters in the decay time fit to $B^0 \to D_s^+ D^-$ data.

Parameter		Value	Variation for Systematics
τ	(ps)	1.520	(±0.004)
Δm	($\hbar\,\text{ps}^{-1}$)	0.510	(±0.003)
$\Delta\Gamma$	(ps^{-1})	0.	none
A_{det}		0.0	(±0.01)
A_{P}		-0.0062	(±0.01)
A_{sl}^{d}		-0.0002	(±0.0036)

5.2.3 Calibration in Run 1

Calibration

In the $B^0 \to D^+ D^-$ analysis a set of combinations for the available taggers is considered, composed of two inclusive samples (OS, SS). OS represents the combination of the Opposite Side standard combination with the Opposite Side charm tagger, while SS is the combination of the Same Side proton tagger with the optimized Same Side pion tagger. To calculate the combined mistag and the combined tagging decision the definitions given in [68] are used.

Results on OS: In the inclusive OS Combination sample, a tagging efficiency of $(41.05 \pm 0.50)\%$ and an effective tagging efficiency of $(4.62 \pm 0.42)\%$ are measured. The calibration parameters (defined in Sec. 5.1.2) and the asymmetry of the tagging efficiency are determined to be

$$\begin{aligned} p_1 &= 1.069 \pm 0.072 \ , \\ p_0 &= 0.3691 \pm 0.0080 \ , \\ \langle \eta \rangle &= 0.3627 \ , \\ \Delta p_1 &= 0.03 \pm 0.11 \ , \\ \Delta p_0 &= -0.009 \pm 0.012 \ , \\ \Delta \epsilon_{\text{tag}} &= 0.012 \pm 0.017 \ , \end{aligned} \tag{5.19}$$

with the correlation matrix

$$\rho_{\text{OS}} = \left(\begin{array}{c|ccccc} & \Delta\epsilon_{\text{tag}} & \Delta p_0 & \Delta p_1 & p_0 & p_1 \\ \hline \Delta\epsilon_{\text{tag}} & 1 & 0.167 & -0.131 & -0.004 & 0.0004 \\ \Delta p_0 & & 1 & 0.086 & -0.013 & 0.007 \\ \Delta p_1 & & & 1 & 0.008 & -0.024 \\ p_0 & & & & 1 & 0.15 \\ p_1 & & & & & 1 \end{array} \right) . \tag{5.20}$$

Results on SS: In the inclusive SS Combination sample, a tagging efficiency of $(78.89 \pm 0.69)\%$ and an effective tagging efficiency of $(2.49 \pm 0.33)\%$ is measured. The calibration parameters and the asymmetry of the tagging efficiency are determined as

$$\begin{aligned} p_1 &= 0.842 \pm 0.090 \ , \\ p_0 &= 0.4296 \pm 0.0060 \ , \\ \langle \eta \rangle &= 0.4282 \ , \\ \Delta p_1 &= 0.07 \pm 0.13 \ , \\ \Delta p_0 &= -0.0065 \pm 0.0087 \ , \\ \Delta \epsilon_{\text{tag}} &= -0.012 \pm 0.012 \ , \end{aligned} \tag{5.21}$$

with the correlation matrix

$$\rho_{\text{SS}} = \left(\begin{array}{c|ccccc} & \Delta\epsilon_{\text{tag}} & \Delta p_0 & \Delta p_1 & p_0 & p_1 \\ \hline \Delta\epsilon_{\text{tag}} & 1 & 0.093 & -0.07 & 0.01 & -0.006 \\ \Delta p_0 & & 1 & 0.02 & 0.025 & -0.006 \\ \Delta p_1 & & & 1 & -0.007 & 0.022 \\ p_0 & & & & 1 & 0.052 \\ p_1 & & & & & 1 \end{array} \right) . \tag{5.22}$$

Results on the full sample: Applying the calibration parameters from the OS and SS samples, the effective tagging efficiency of the full sample when using a combination of the OS and SS tagging response is found as $(7.29 \pm 0.48)\%$. When considering the sample where the events are tagged by both OS and SS combinations, an effective tagging efficiency of $(4.43 \pm 0.33)\%$ is found.

5.2.4 Calibration in Run 2

Calibration

Also in the $B^0 \to D^{*\pm}D^{\mp}$ analysis a set of combinations for the available taggers is considered. It is composed of two inclusive samples (OS, SS), where OS is the Opposite Side standard combination , while SS represents the combination of the Same Side proton with the BDT optimized Same Side pion.

Results on OS: A tagging efficiency of $(36.42 \pm 0.43)\%$ and an effective tagging efficiency of $(2.91 \pm 0.41)\%$ are measured, in the inclusive OS standard combination sample. The results obtained for the calibration parameters and for the asymmetry of the tagging efficiency are reported in Eq. (5.23)

$$\begin{aligned} p_1 &= 0.810 \pm 0.104 \ , \\ p_0 &= 0.3797 \pm 0.0104 \ , \\ \langle\eta\rangle &= 0.3703 \ , \\ \Delta p_1 &= 0.272 \pm 0.149 \ , \\ \Delta p_0 &= 0.0122 \pm 0.0150 \ , \\ \Delta\epsilon_{\text{tag}} &= -0.046 \pm 0.017 \ , \end{aligned} \tag{5.23}$$

the corresponding correlation matrix is

$$\rho_{\text{OS}} = \left(\begin{array}{c|ccccc} & \Delta\epsilon_{\text{tag}} & \Delta p_0 & \Delta p_1 & p_0 & p_1 \\ \hline \Delta\epsilon_{\text{tag}} & 1 & -0.129 & 0.078 & 0.011 & 0.005 \\ \Delta p_0 & & 1 & 0.038 & -0.118 & 0.014 \\ \Delta p_1 & & & 1 & 0.018 & -0.159 \\ p_0 & & & & 1 & 0.094 \\ p_1 & & & & & 1 \end{array}\right) . \tag{5.24}$$

Results on SS: A tagging efficiency of $(87.86 \pm 0.33)\%$ and an effective tagging efficiency of $(3.54 \pm 0.45)\%$ are measured, in the inclusive SS combination sample. The results obtained for the calibration parameters and for the asymmetry of the tagging efficiency are reported in Eq. (5.25)

$$\begin{aligned} p_1 &= 1.147 \pm 0.108 \ , \\ p_0 &= 0.4296 \pm 0.0069 \ , \\ \langle\eta\rangle &= 0.4303 \ , \\ \Delta p_1 &= 0.018 \pm 0.163 \ , \\ \Delta p_0 &= -0.0048 \pm 0.0098 \ , \\ \Delta\epsilon_{\text{tag}} &= 0.0069 \pm 0.0055 \ , \end{aligned} \tag{5.25}$$

the corresponding correlation matrix is

$$\rho_{\mathrm{SS}} = \left(\begin{array}{c|ccccc} & \Delta\epsilon_{\mathrm{tag}} & \Delta p_0 & \Delta p_1 & p_0 & p_1 \\ \hline \Delta\epsilon_{\mathrm{tag}} & 1 & -0.038 & 0.035 & -0.002 & 0.005 \\ \Delta p_0 & & 1 & 0.021 & -0.034 & -0.009 \\ \Delta p_1 & & & 1 & -0.010 & -0.005 \\ p_0 & & & & 1 & 0.055 \\ p_1 & & & & & 1 \end{array}\right). \tag{5.26}$$

Results on the full sample: The effective tagging efficiency of the full sample when applying the calibration parameters for the OS and SS samples considering an offline combination of the OS and SS tagging response results to be $(6.23 \pm 0.0.59)\%$.

5.2.5 Portability of the calibration from control to $B^0 \to D^+D^-$ signal channel

In order to check if the control channel $B^0 \to D_s^+D^-$ and the signal channel $B^0 \to D^+D^-$ have a similar tagging behavior, a study of kinematic observables that are relevant for tagging and of the output of the tagging algorithms are compared between the two decay modes using MC samples. All nominal selection requirements are applied, and the reconstructed candidates in simulated samples are required to be truth-matched.

MC studies A comparison between the transverse momentum in the control and the signal channel $B^0 \to D^+D^-$ is shown in Fig. 5.5, proving that after applying the selection requirements the distributions are similar. Furthermore, the mistag distributions for the

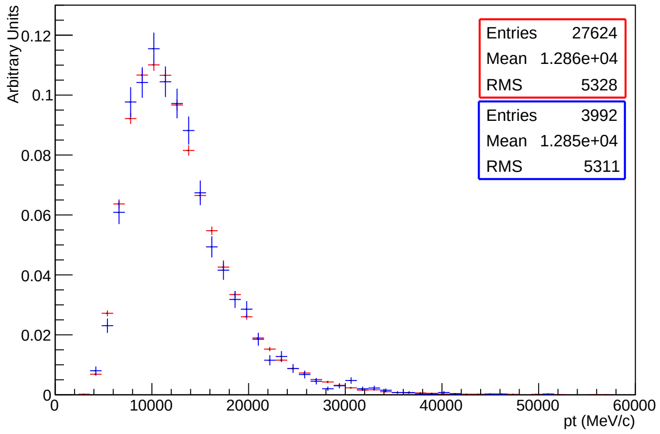


Figure 5.5: Comparison between B^0 transverse momentum distributions for $B^0 \to D^+D^-$ (blue) and $B^0 \to D_s^+D^-$ (red) in Run 1 MC sample.

correctly and wrongly tagged candidates of the two channels are compared. For this, the tag decision is compared with the MC-truth information on the production state of the B^0 candidate. As an example in Fig. 5.6 the OS mistag distributions are shown. The distributions show a good agreement between the channels. Following Eq. (5.2) the ratio between the distribution of wrong tagged events and the distribution of all tagged events is calculated and a comparison of the calibration parameters in the control and signal channel is done, the result presented in Fig. 5.7 shows that the values for these parameters in the two channels are compatible (see Table 5.6). For this reason no additional systematic uncertainty on portability is assigned.

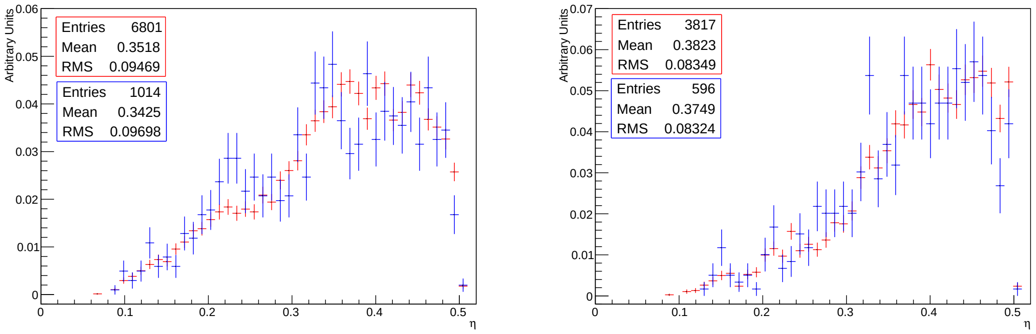


Figure 5.6: Comparison between mistag distributions for $B^0 \to D_s^+ D^-$ (red) and $B^0 \to D^+ D^-$ (blue) Run 1 signal MC for correctly (left) and wrongly (right) tagged events, using Opposite Side tagger.

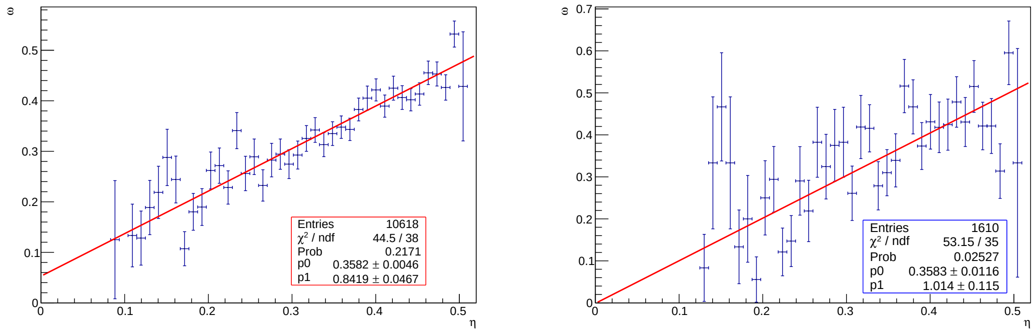


Figure 5.7: Mistag probability ω as a function of the predicted mistag probability η using OS tagger for $B^0 \to D_s^+ D^-$ (left) and $B^0 \to D^+ D^-$ (right) Run 1 MC samples.

Studies with data After applying the full selection, a study on the data samples is performed, in order to check the similarity between the control and signal channel

Table 5.5: Calibration parameters extracted from $B^0 \to D_s^+ D^-$ and $B^0 \to D^+ D^-$ Run 1 Monte Carlo samples for the OS tagger combination.

	$B^0 \to D_s^+ D^-$	$B^0 \to D^+ D^-$
p_0	0.3582 ± 0.0046	0.358 ± 0.012
p_1	0.842 ± 0.047	1.01 ± 0.12
$\langle \eta \rangle$	0.3627	0.3545
$\varepsilon_{\text{tag}}(\%)$	38.44 ± 0.44	40.33 ± 1.19
$\varepsilon_{\text{eff}}(\%)$	4.01 ± 0.22	4.69 ± 0.65

$B^0 \to D^+ D^-$. The distributions for the variables of interest are shown in Fig. 5.8 and Fig. 5.9 respectively. Small differences are visible, but generally the phase space of $B^0 \to D_s^+ D^-$ and $B^0 \to D^+ D^-$ is very similar. For the flavour tagging calibration it is important that the parameters used to calibrate the predicted mistag rates (which take possible differences between the channels already into account) are the same. The compatibility of the calibration parameters is discussed in the previous section and the tests performed on data confirm the validity of the simulation study after applying the selection on data.

5.2.6 Portability of the calibration from control to $B^0 \to D^{*\pm} D^{\mp}$ signal channel

Similarly to what described in Sec. 5.2.5 a compatibility study involving relevant kinematic variables, is performed using the control channel and the signal $B^0 \to D^{*\pm} D^{\mp}$. The full selection is applied, and in MC samples the reconstructed candidates are required to be truth-matched.

MC studies A comparison between the transverse momentum in the control and the signal channel $B^0 \to D^{*\pm} D^{\mp}$ is presented in Fig. 5.10. The study shows that after applying the selection requirements there is a deviation between the control and the two signal decays, however the control channel appears to be intermediate between them. This effect has to be taken into account, applying a reweighting in the control channel in order to match the kinematic features of the signal.

In Fig. 5.11 the OS combination and the SS combination mistag distributions are compared for the control channel and the signal channel $B^0 \to D^{*\pm} D^{\mp}$. The distributions show a good agreement between the two decay modes.

Following Eq. (5.4) a comparison of the calibration parameters in the control and signal channel is done (see Fig. 5.12), showing that the values for these parameters in the two channels are compatible, thus no additional systematic uncertainty on portability is assigned. The results are listed in Table 5.6.

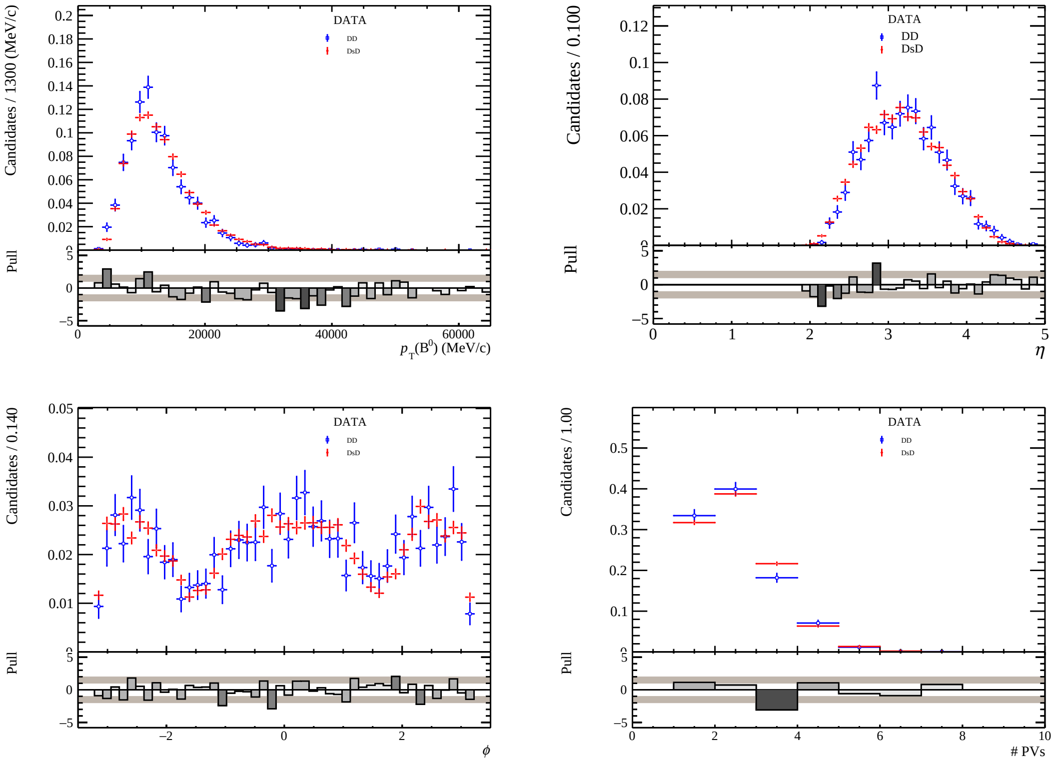
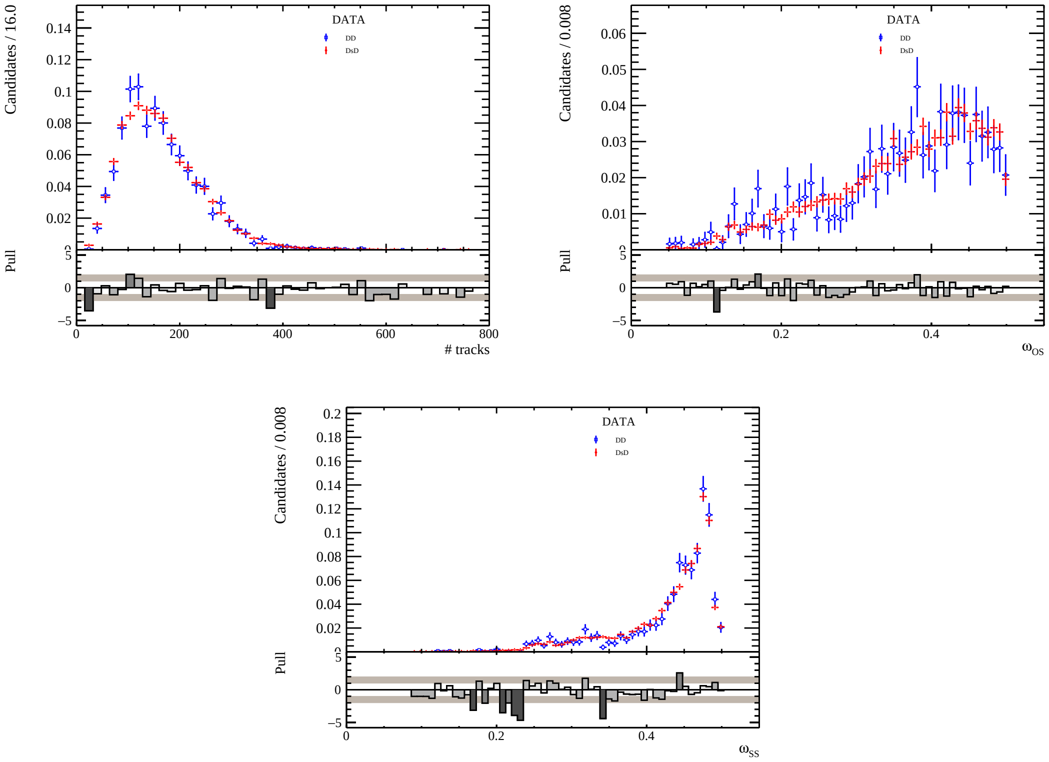


Figure 5.8: Comparison between $B^0 \to D_s^+ D^-$ Run1 data sample (red) and $B^0 \to D^+ D^-$ (blue) for B^0 transverse momentum (top left), pseudorapidity (top right), azimuthal angle (bottom left) and number of primary vertex (bottom right).

Table 5.6: Calibration parameters obtained from $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*\pm} D^{\mp}$ Run 1 Monte Carlo samples for the OS tagger combination.

	$B^0 \to D_s^+ D^-$	$B^0 \to D^{*\pm} D^{\mp}(k3\pi)$	$B^0 \to D^{*\pm} D^{\mp}(k\pi)$
$p_0 - \langle\eta\rangle$	-0.005 $\pm$ 0.005	-0.007 $\pm$ 0.005	0.004 $\pm$ 0.002
p_1	0.84 $\pm$ 0.05	0.88 $\pm$ 0.06	1.04 $\pm$ 0.03
$\langle\eta\rangle$	0.363	0.358	0.359
$\varepsilon_{\text{tag}}(\%)$	38.44 $\pm$ 0.44	41.81 $\pm$ 0.37	41.33 $\pm$ 0.16
$\varepsilon_{\text{eff}}(\%)$	4.01 $\pm$ 0.22	4.44 $\pm$ 0.33	4.28 $\pm$ 0.14

Studies with data In order to check the similarity between the control and the $B^0 \to D^{*\pm} D^{\mp}$ signal channel, a study is performed, after applying the full selection, on the respective ${}_s$*weighted* data samples. The distributions for the variables of interest are shown in Fig. 5.13 and Fig. 5.14. Small difference are visible, but generally the phase space of $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*\pm} D^{\mp}$ are similar. The tests on data show that the



3

Figure 5.9: Comparison between $B^0 \to D_s^+ D^-$ Run1 data sample (red) and $B^0 \to D^+ D^-$ (blue) for number of tracks (top left), mistag probability for the Opposite Side tagger combination (top right) and mistag probability for the Same Side tagger combination (bottom).

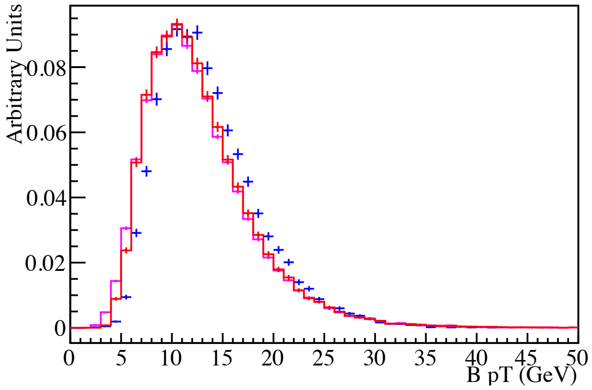


Figure 5.10: Comparison between B_d transverse momentum distributions for $B^0 \to D_s^+ D^-$ (red), $B^0 \to D^{*\pm} D^{\mp}(k3\pi)$ (blue) and $B^0 \to D^{*\pm} D^{\mp}(k\pi)$ (magenta) in Run 1 MC.

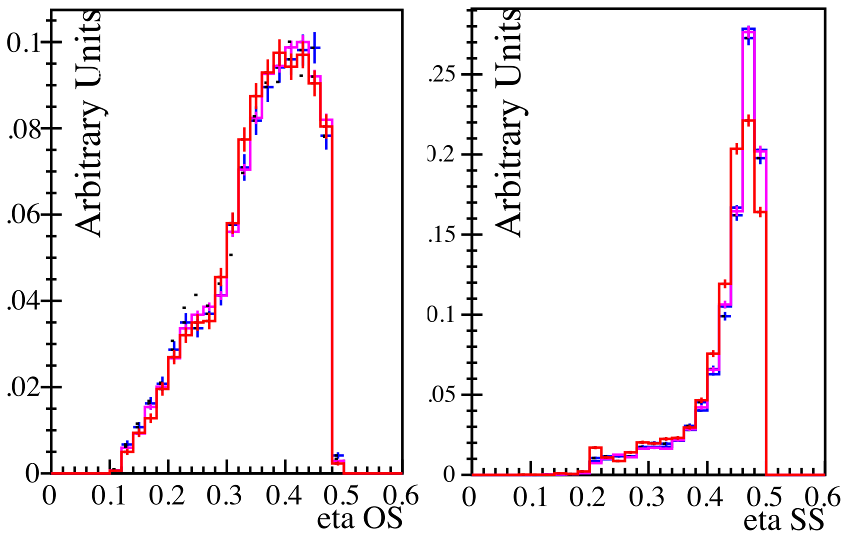


Figure 5.11: Comparison between mistag distributions for $B^0 \to D_s^+ D^-$ (red), $B^0 \to D^{*\pm} D^{\mp}(k3\pi)$ (blue) and $B^0 \to D^{*\pm} D^{\mp}(k\pi)$(magenta) Run 1 signal MC using Opposite Side (left plot) and Same Side (right plot) tagger combinations respectively.

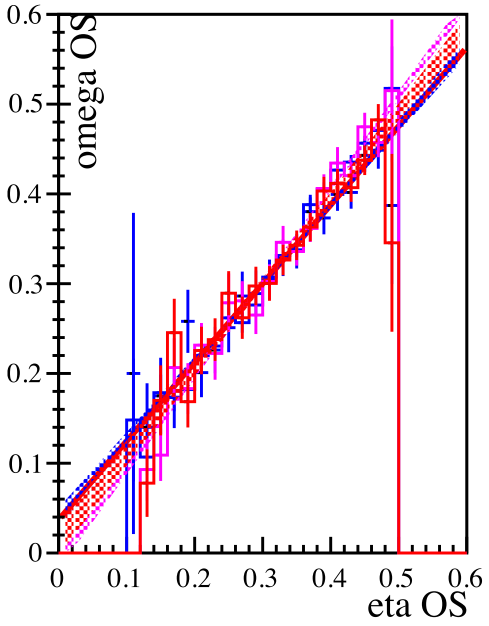


Figure 5.12: Mistag probability ω as a function of the predicted mistag probability η using OS tagger for the $B^0 \to D_s^+ D^-$ (red), $B^0 \to D^{*\pm} D^{\mp}(k3\pi)$ (blue) and $B^0 \to D^{*\pm} D^{\mp}(k\pi)$ (magenta) Run 1 MC samples.

validity of the simulation study is not spoiled by a different phase space coverage of the two channels after applying the selection on data.

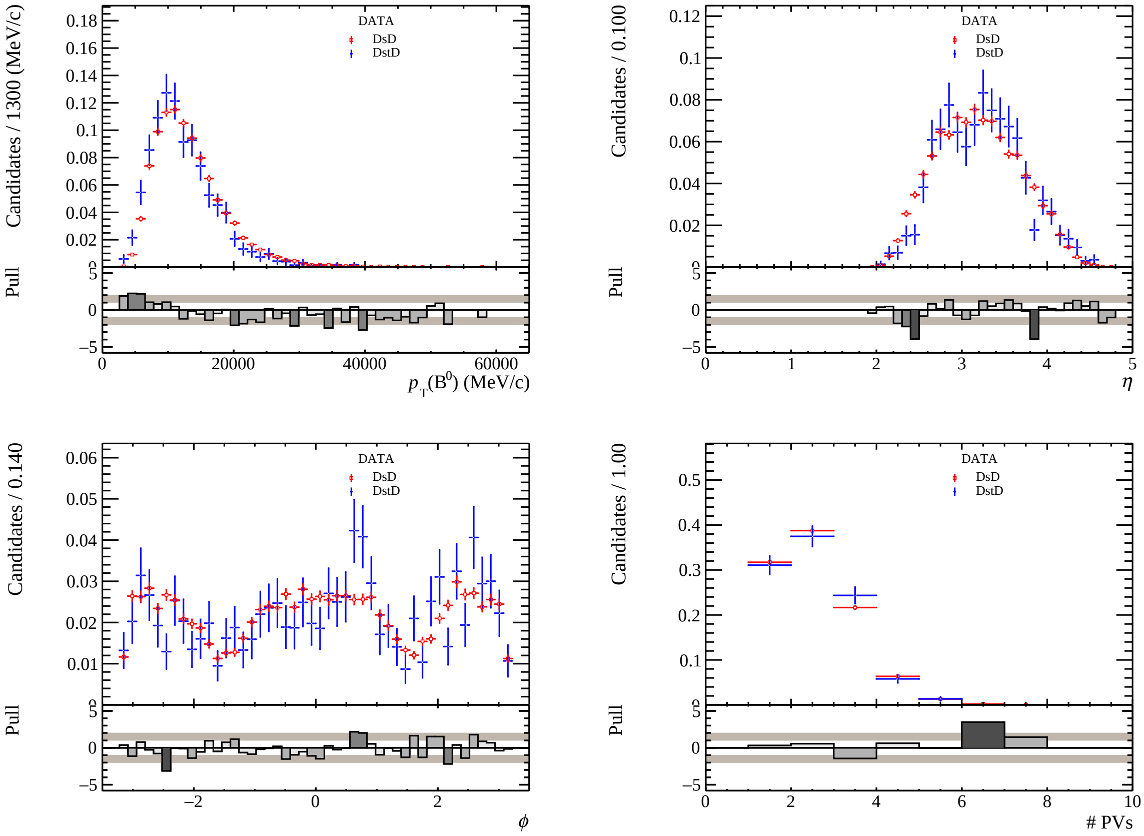


Figure 5.13: Comparison between $B^0 \to D_s^+ D^-$ Run 1 data sample (red) and $B^0 \to D^{*\pm} D^{\mp}$ (blue) for B^0 transverse momentum (top left), pseudorapidity (top right), azimuthal angle (bottom left) and number of primary vertex (bottom right).

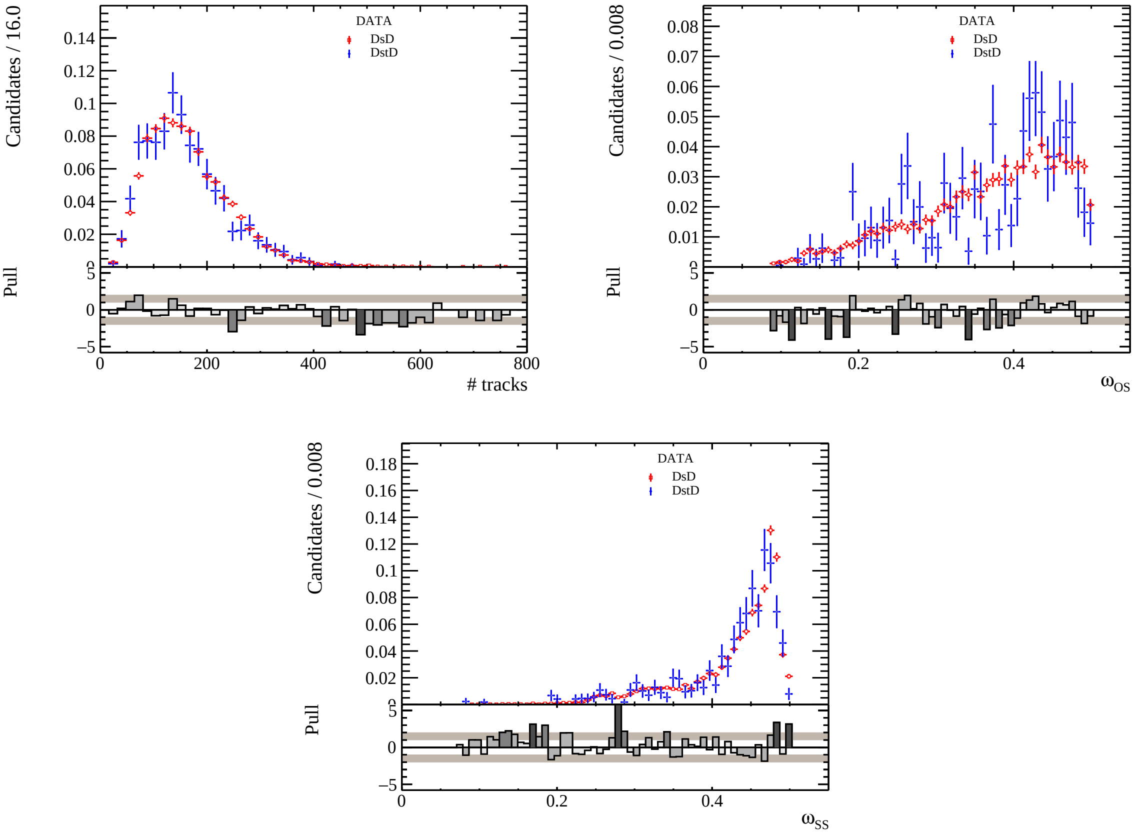


Figure 5.14: Comparison between $B^0 \to D_s^+ D^-$ Run 1 data sample (red) and $B^0 \to D^{*\pm} D^{\mp}$ (blue) for number of tracks (top left), mistag probability for the Opposite Side tagger combination (top right) and mistag probability for the Same Side tagger combination (bottom).

5.3 Studies in the control channel $B^0 \rightarrow D^{+*}D_s^-$

As already explained in Sec. 5.2, the choice of the control channel strongly depends on signal features. As described in Sec. 5.1.3 the $B^0 \rightarrow D^{*+}D_s^-$ decay channel is chosen as flavour tagging control and calibration channel, since its similarity is supposed to result in a good portability of the calibration parameters to the signal channel $B^0 \rightarrow D^{*\pm}D^{\mp}$.

5.3.1 Selection

The selection applied to $B^0 \rightarrow D^{*+}D_s^-$ closely follows the one applied to $B^0 \rightarrow D^{*\pm}D^{\mp}$ (see Sec. 7.1). For this decay mode, the D_s^+ meson is reconstructed as:

- $D_s^{\mp} \rightarrow K^{\pm}K^{\mp}\pi^{\mp}$

while the $D^{*\pm}$ mesons are reconstructed as $D^{*\pm} \rightarrow D^0\pi^{\pm}$, where the D^0 final states considered are:

- $D^0 \rightarrow K^-\pi^-\pi^+\pi^+$
- $D^0 \rightarrow K^-\pi^+$
- $D^0 \rightarrow K^+K^-$

Stripping In the procedure used to make tuples from raw data i.e. stripping, the daughter particles (kaons and pions) of the $D_s^{\pm}$ and D^0 mesons have to fulfil a set of criteria, following those listed in Table 7.2. The further requirements on the $D^{*\pm}D_s^{\mp}$ combination are given in Table 7.3, where the D^* and $D_s^{\pm}$ mesons candidates have to form a common vertex. The stripping requirements on the B^0 combination are listed in Table 7.4.

Preselection During preselection the mass window around the $D^{*\pm}D_s^{\mp}$ is restricted to the range 5220 MeV/c^2 − 5500 MeV/c^2. The complete list of cut applied are reported in Table 5.7.

Table 5.7: Offline cuts applied in the selection of $B^0 \rightarrow D^{*+}D_s^-$.

$PID(K, D_s^+)$	K_Dp_PIDK > 0.
D^*, D^0 masses	$\mid m_{D^*} - m_{D^0} \mid < 152.$ MeV/c^2
D^0 mass	$\mid m_{D^0} - 1864.83 \mid < 35.$ MeV/c^2
$D_s^{\pm}$ mass	$\mid m_{D_s^{\pm}} - 1968.27 \mid < 45.$ MeV/c^2
Candidates Number	nCandidates ==0

5.3.2 Fit Model

As described in Sec. 5.2.2, also for this control mode a mixing analysis is required in order to extract calibration parameters to be used in the $B^0 \to D^{*\pm} D^{\mp}$ decay time fit. A fit to the invariant mass distribution of the $B^0 \to D^{*+} D_s^-$ candidates is used to obtain ${}_s\mathcal{W}eights$.

Mass Fit The mass fit is performed in the range 5220 MeV/c^2 to 5500 MeV/c^2. For the mass fit the extended maximum likelihood method is used. The selection requirements strongly suppress backgrounds events so that the PDF includes only two components, namely the $B^0 \to D^{*+} D_s^-$ signal, and the combinatorial background. The B^0 signal component is parametrized with a sum of two Gaussian PDFs with the same mean μ_{B^0} but different width parameters σ_i, while the reconstructed mass PDF of the combinatorial background component is modelled with an exponential function.

In Table 5.8 the results of the floating shape parameters of the mass fit are shown along with the fit results for the yields, while Fig. 5.15 shows the distributions of the data overlaid with the PDF projections.

Table 5.8: Results of the floating shape parameters in the mass fit to $B^0 \to D^{*+} D_s^-$ for Run 1 (left) and Run 2 (right) data samples.

Sample	Parameter	Fitted Value
Run 1	$\mu_{B^0 \to D^{*+} D_s^-}$ (MeV/c^2)	5278.85 ± 0.14
	$\sigma_{1,B^0 \to D^{*+} D_s^-}$ (MeV/c^2)	11.54 ± 0.18
	$\sigma_{2,B^0 \to D^{*+} D_s^-}$ (MeV/c^2)	40.10 ± 9.10
	$f_{B^0 \to D^{*+} D_s^-}$	0.911 ± 0.029
	β (1/MeV/c^2)	-0.0027 ± 0.0003
	$N_{B^0 \to D^{*+} D_s^-}$	12331 ± 129
Run 2	$\mu_{B^0 \to D^{*+} D_s^-}$ (MeV/c^2)	5279.17 ± 0.12
	$\sigma_{1,B^0 \to D^{*+} D_s^-}$ (MeV/c^2)	5.57 ± 0.79
	$\sigma_{2,B^0 \to D^{*+} D_s^-}$ (MeV/c^2)	12.85 ± 0.27
	$f_{B^0 \to D^{*+} D_s^-}$	0.112 ± 0.0034
	β (1/MeV/c^2)	-0.0004 ± 0.0001
	$N_{B^0 \to D^{*+} D_s^-}$	15903 ± 143

Espresso Performance Monitor After obtaining the ${}_s\mathcal{W}eights$ from the fit to the invariant mass distribution described in the previous section, the calibration parameters are determined from the weighted sample, making use of the Espresso Performance Monitor (EPM) tool, developed by the LHCb collaboration. The EspressoPerformanceMonitor package consists of several parts. It is based on a C++ library designed to perform

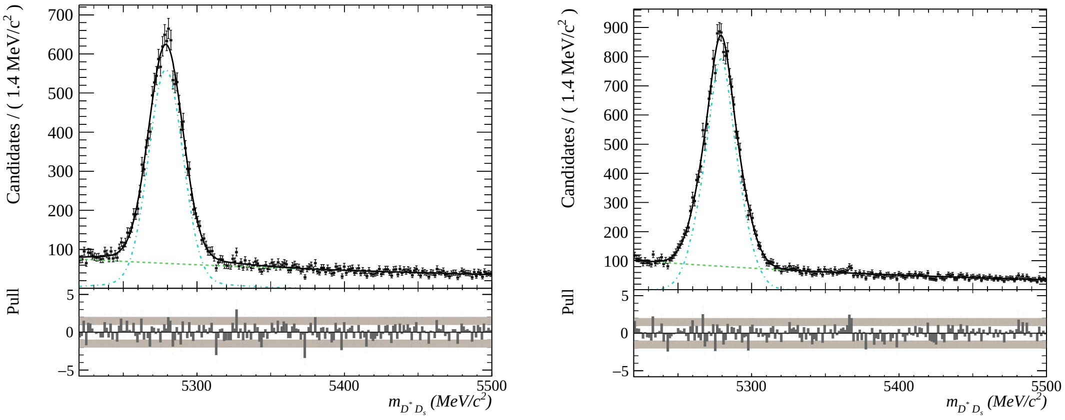


Figure 5.15: Fit to the B^0 mass, in order to extract ${}_s\mathcal{W}eights$ for Run 1 (left) and Run 2 (right) data samples. The solid line is the PDF projection, the cyan dotted line represent the signal component, while the green dotted line represents the combinatorial background.

numerical, unbinned calibrations of flavour tagging algorithms with binomial regression. On top of this base, it adds support for looping over ROOT files and extracting calibrations and performance numbers and also provides support for performing toy studies [70].

5.3.3 Calibration in Run 1

Calibration

For the $B^0 \to D^{*\pm}D^{\mp}$ analysis the combination of the calibration parameters coming from $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*+}D_s^-$ is used. Also for $B^0 \to D^{*+}D_s^-$ calibration channel, a combination of the available taggers is considered, namely the Opposite Side standard combination and the combination of Same Side proton and optimized Same Side pion taggers.

Results on OS: In the inclusive OS Combination sample, a tagging efficiency of $(36.86 \pm 0.52)\%$ and an effective tagging efficiency of $(4.11 \pm 0.63)\%$ are measured. The calibration parameters are determined as

$$\begin{aligned} p_1 &= 1.093 \pm 0.128 \ , \\ p_0 &= 0.3752 \pm 0.0134 \ , \\ \langle \eta \rangle &= 0.3675 \ , \\ \Delta p_1 &= 0.176 \pm 0.171 \ , \\ \Delta p_0 &= 0.0256 \pm 0.0172 \ , \end{aligned} \tag{5.27}$$

with the correlation matrix

$$\rho_{\mathrm{OS}} = \left(\begin{array}{c|ccccc} & \Delta\epsilon_{\mathrm{tag}} & \Delta p_0 & \Delta p_1 & p_0 & p_1 \\ \hline \Delta\epsilon_{\mathrm{tag}} & 1 & na & na & na & na \\ \Delta p_0 & & 1 & 0.074 & -0.011 & -0.019 \\ \Delta p_1 & & & 1 & -0.017 & 0.063 \\ p_0 & & & & 1 & 0.111 \\ p_1 & & & & & 1 \end{array}\right) . \tag{5.28}$$

Results on SS: In the inclusive SS Combination sample, a tagging efficiency of $(68.98 \pm 0.50)\%$ and an effective tagging efficiency of $(2.82 \pm 0.51)\%$ are measured. The calibration parameters result to be

$$\begin{aligned} p_1 &= 1.056 \pm 0.132 \ , \\ p_0 &= 0.4261 \pm 0.0100 \ , \\ \langle\eta\rangle &= 0.4274 \ , \\ \Delta p_1 &= 0.007 \pm 0.184 \ , \\ \Delta p_0 &= -0.0005 \pm 0.0129 \ , \end{aligned} \tag{5.29}$$

with the correlation matrix

$$\rho_{\mathrm{SS}} = \left(\begin{array}{c|ccccc} & \Delta\epsilon_{\mathrm{tag}} & \Delta p_0 & \Delta p_1 & p_0 & p_1 \\ \hline \Delta\epsilon_{\mathrm{tag}} & 1 & na & na & na & na \\ \Delta p_0 & & 1 & 0.044 & -0.006 & -0.006 \\ \Delta p_1 & & & 1 & -0.004 & 0.080 \\ p_0 & & & & 1 & 0.076 \\ p_1 & & & & & 1 \end{array}\right) . \tag{5.30}$$

Results on the full sample: Applying the calibration parameters for the OS and SS samples, the effective tagging efficiency of the full sample when using a combination of the OS and SS tagging response is found to be $(6.91 \pm 0.83)\%$.

5.3.4 Calibration in Run 2

Results on OS: In the inclusive OS Combination sample, a tagging efficiency of $(35.88 \pm 0.44)\%$ and an effective tagging efficiency of $(4.15 \pm 0.56)\%$ are measured. The calibration parameters are

$$\begin{aligned} p_1 &= 1.058 \pm 0.125 \ , \\ p_0 &= 0.3635 \pm 0.0123 \ , \\ \langle\eta\rangle &= 0.3684 \ , \\ \Delta p_1 &= 0.074 \pm 0.159 \ , \\ \Delta p_0 &= 0.0145 \pm 0.0154 \ , \end{aligned} \tag{5.31}$$

with the correlation matrix

$$\rho_{\mathrm{OS}} = \left(\begin{array}{c|ccccc} & \Delta\epsilon_{\mathrm{tag}} & \Delta p_0 & \Delta p_1 & p_0 & p_1 \\ \hline \Delta\epsilon_{\mathrm{tag}} & 1 & na & na & na & na \\ \Delta p_0 & & 1 & 0.059 & 0.0007 & -0.031 \\ \Delta p_1 & & & 1 & -0.030 & 0.042 \\ p_0 & & & & 1 & 0.094 \\ p_1 & & & & & 1 \end{array}\right) . \tag{5.32}$$

Results on SS: In the inclusive SS Combination sample, a tagging efficiency of $(84.17 \pm 0.33)\%$ and an effective tagging efficiency of $(3.07 \pm 0.53)\%$ are measured. The calibration parameters are determined as

$$\begin{aligned} p_1 &= 1.055 \pm 0.132 \ , \\ p_0 &= 0.4327 \pm 0.0079 \ , \\ \langle \eta \rangle &= 0.4269 \ , \\ \Delta p_1 &= -0.210 \pm 0.164 \ , \\ \Delta p_0 &= 0.0221 \pm 0.0098 \ , \end{aligned} \tag{5.33}$$

with the correlation matrix

$$\rho_{\mathrm{OS}} = \left(\begin{array}{c|ccccc} & \Delta\epsilon_{\mathrm{tag}} & \Delta p_0 & \Delta p_1 & p_0 & p_1 \\ \hline \Delta\epsilon_{\mathrm{tag}} & 1 & na & na & na & na \\ \Delta p_0 & & 1 & -0.012 & -0.004 & -0.010 \\ \Delta p_1 & & & 1 & -0.010 & -0.013 \\ p_0 & & & & 1 & 0.0030 \\ p_1 & & & & & 1 \end{array}\right) . \tag{5.34}$$

Results on the full sample: Applying the calibration parameters for the OS and SS samples, the effective tagging efficiency of the full sample when combining the OS and SS tagging responses is found as $(6.78 \pm 0.72)\%$.

5.3.5 Portability of the calibration from control to $B^0 \to D^{*\pm} D^{\mp}$ signal channel

Similarly to what described in Sec. 5.2.6, a compatibility study is performed using the control channel and the signal $B^0 \to D^{*\pm} D^{\mp}$ samples. The full selection is applied, and for the MC samples the reconstructed candidates are required to be truth-matched.

MC studies A comparison between the transverse momentum in the control and the signal channel $B^0 \to D^{*\pm} D^{\mp}$ is shown in Fig. 5.16. The study shows that after applying the selection requirements there is a difference between the two decay modes. This effect has to be taken into account, applying a reweighting in the control channel in order to match the kinematic features of the signal.

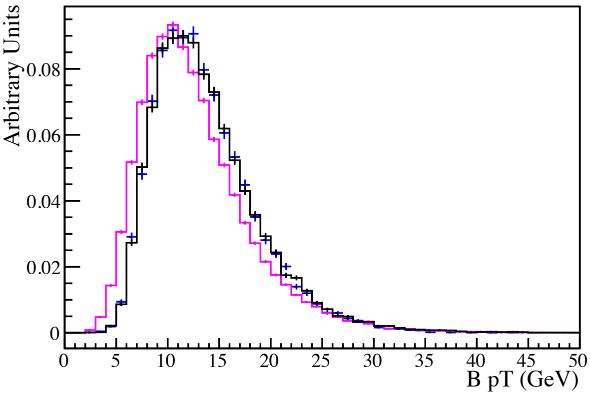


Figure 5.16: Comparison between B_d transverse momentum distributions for $B^0 \to D^{*+}D_s^-(k3\pi/k\pi/kk)$ (black), $B^0 \to D^{*\pm}D^{\mp}(k3\pi)$ (blue) and $B^0 \to D^{*\pm}D^{\mp}(k\pi)$ (magenta) in Run 1 MC.

In Fig. 5.17 the OS combination and the SS combination mistag distributions are compared for the control channel and the signal channel $B^0 \to D^{*\pm}D^{\mp}$. The distributions show a good agreement between the two decay modes. Following Eq. (5.4) a comparison

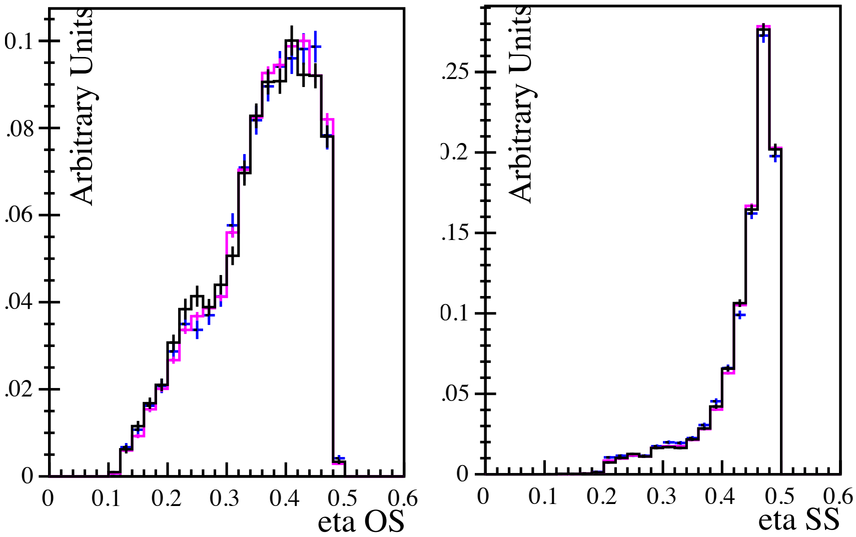


Figure 5.17: Comparison between mistag distributions for $B^0 \to D^{*+}D_s^-(k3\pi/k\pi/kk)$ (black), $B^0 \to D^{*\pm}D^{\mp}(k3\pi)$ (blue) and $B^0 \to D^{*\pm}D^{\mp}(k\pi)$ Run 1 signal MC (magenta) using Opposite Side and Same Side tagger combinations respectively.

of the calibration parameters in the control and signal channel is done (see Fig. 5.18), the results listed in Table 5.9 show that the values for these parameters in the two channels are compatible. Thus, no additional systematic uncertainty on portability is assigned.

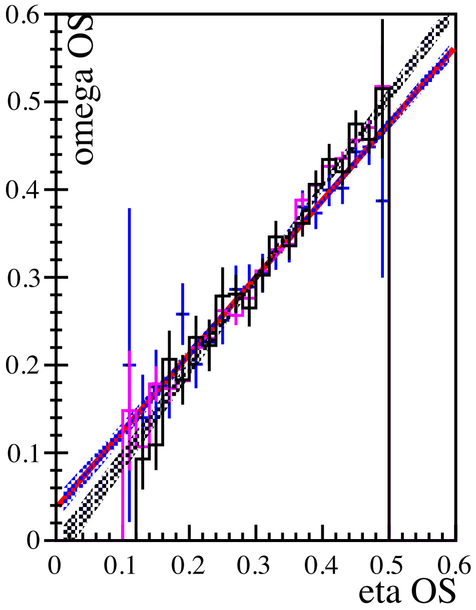


Figure 5.18: Mistag probability ω as a function of the predicted mistag probability η using OS tagger for the $B^0 \to D^{*+}D_s^-(k3\pi/k\pi/kk)$ Run1 Monte Carlo sample (black), $B^0 \to D^{*\pm}D^{\mp}(k3\pi)$ (blue) and $B^0 \to D^{*\pm}D^{\mp}(k\pi)$ MC sample (magenta).

Table 5.9: Calibration parameters obtained from $B^0 \to D^{*+}D_s^-$ and $B^0 \to D^{*\pm}D^{\mp}$ Run1 Monte Carlo samples for the OS tagger combination.

	$B^0 \to D^{*+}D_s^-$	$B^0 \to D^{*\pm}D^{\mp}(k3\pi)$	$B^0 \to D^{*\pm}D^{\mp}(k\pi)$
$p_0 - \langle\eta\rangle$	-0.001 ± 0.005	-0.007 ± 0.005	0.004 ± 0.002
p_1	1.04 ± 0.05	0.88 ± 0.06	1.04 ± 0.03
$\langle\eta\rangle$	0.355	0.358	0.359
$\varepsilon_{\text{tag}}(\%)$	39.44 ± 0.46	41.81 ± 0.37	41.33 ± 0.16
$\varepsilon_{\text{eff}}(\%)$	4.61 ± 0.22	4.44 ± 0.33	4.28 ± 0.14

Studies with data In order to check the similarity between the control and the $B^0 \to D^{*\pm}D^{\mp}$ signal channel, a study is performed, after applying the full selection, on the respective data samples. The distributions of the variables of interest are shown in Fig. 5.19 and Fig. 5.20. Some difference are visible, but generally the phase space of $B^0 \to D^{*+}D_s^-$ and $B^0 \to D^{*+}D^-$ are similar. The tests on data show the validity of the simulation studies for the two channels after applying the selection on $_s$*weighted* data.

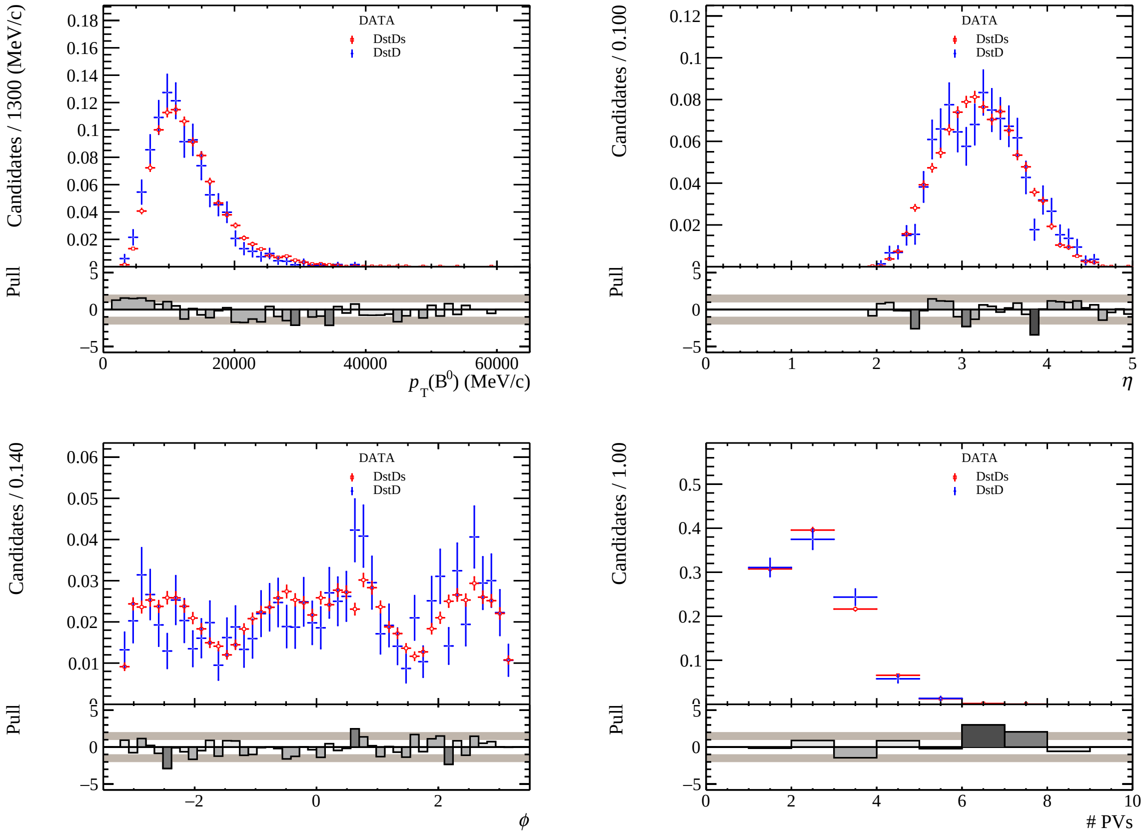


Figure 5.19: Comparison between $B^0 \to D^{*+} D_s^-$ Run 1 data sample (red) and $B^0 \to D^{*\pm} D^{\mp}$ (blue) for B^0 transverse momentum (top left), pseudorapidity (top right), azimuthal angle (bottom left) and number of primary vertex (bottom right).

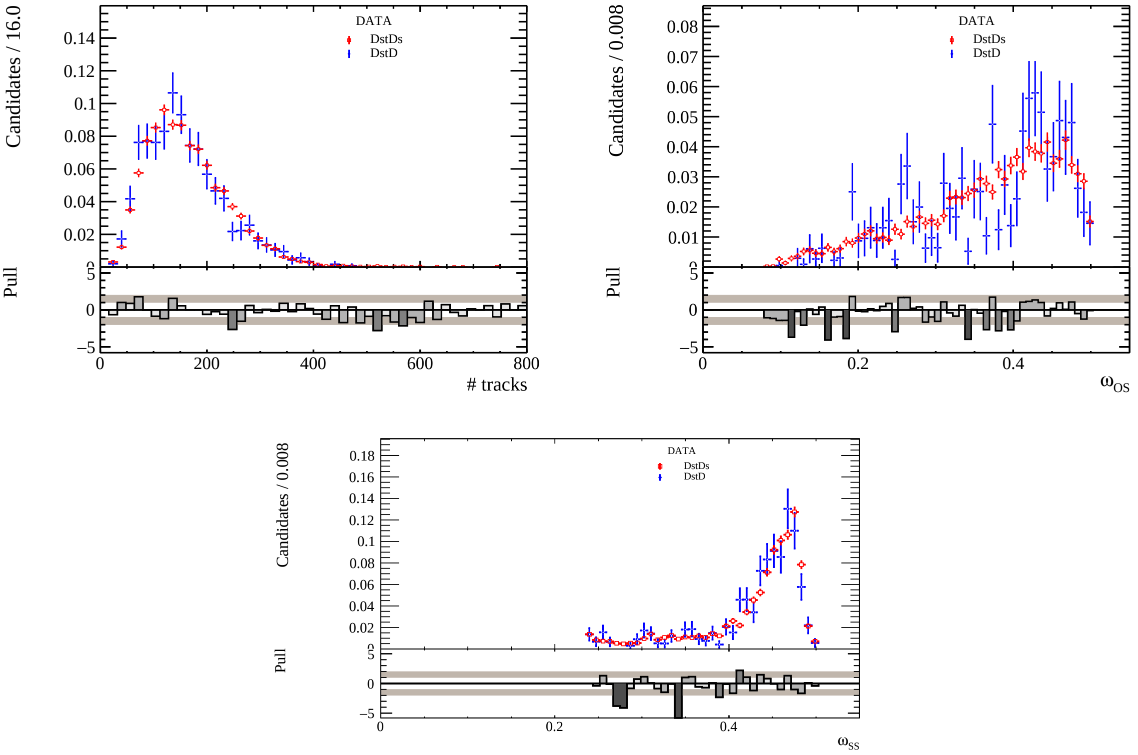


Figure 5.20: Comparison between $B^0 \to D^{*+} D_s^-$ Run 1 data sample (red) and $B^0 \to D^{*\pm} D^{\mp}$ (blue) for number of tracks (top left), mistag probability for the Opposite Side tagger combination (top right) and mistag probability for the Same side tagger combination (bottom).

5.4 Systematic uncertainties

5.4.1 Systematic uncertainties to $B^0 \to D^+D^-$ analysis

Several systematic effects related to the method of determining the calibration parameters are taken into account, these are determined separately for the OS and SS samples respectively. The systematic uncertainty associated to the ${}_s\mathcal{W}eights$ is assigned as the difference between the results obtained using the default ${}_s\mathcal{W}eights$, obtained from the mass fit to the invariant mass of B^0 candidates and the ones obtained performing a two-dimensional mass fit to the invariant mass distribution of the D^+ and D_s^- candidates.

An unbinned maximum likelihood fit in five category of mistag is used to assess the systematic effects related to the choice of the calibration method. The mistag category for the OS combination $\{0.20, 0.29, 0.34, 0.40, 0.50\}$ and for the SS combination $\{0.25, 0.34, 0.40, 0.45, 0.50\}$ are chosen in order to have the same number of events for each bin. The calibration parameters are obtained with a χ^2 fit using the linear function described in Eq. (5.4) on ω_i and η_i points. The full difference of the results obtained from the nominal method with respect to those obtained from the binned one, are used as systematic uncertainty.

In the nominal fit the acceptance function is implemented using a spline function, with parameters left free in the fit. In order to evaluate the effect of the choice of the spline definition, an alternative configuration is tested, while in the nominal fit the same number of knots at the same position as in the signal channel $B^0 \to D^+D^-$ is used, the alternative configuration make use of five knots at position $\{0.2, 0.5, 2.3, 9, 12\,\mathrm{ps}\}$.

The B^0 oscillation period of about 12 ps is much larger than the typical decay time resolution of LHCb of 50 fs, for this reason no larger influence is expected from an over or under estimation of the resolution. The associated systematic is evaluated replacing the function used for describing the resolution in the nominal configuration, with a double Gaussian model. The resolution parameters are fixed from $B^0 \to D_s^+D^-$ MC sample.

Finally in order to evaluate the systematic associated to the fixed parameters such as production asymmetry, detection asymmetry and the physics parameters, these are changed by 1σ and the determination of the calibration parameters is repeated.

All the systematic uncertainties on the calibration parameters are found to be small w.r.t their statistical uncertainties. The results for OS sample and SS sample are given in Table 5.10 and Table 5.11 respectively.

5.4.2 Systematic uncertainties to $B^0 \to D^{*\pm}D^{\mp}$ analysis

As for the $B^0 \to D^+D^-$ analysis also for the $B^0 \to D^{*\pm}D^{\mp}$ different sources of systematic uncertainties will be taken into account on the respective control channels. They will be determined separately for the OS and SS samples respectively. For the $B^0 \to D^{*\pm}D^{\mp}$ analysis two control channels have been used, for this reason an estimate of the systematic uncertainties will be performed separately on the two decay modes, and the results will combined in a final systematic uncertainty, which is expected to be of the same order of

Table 5.10: Systematic uncertainties on the flavour tagging calibration parameters for the OS sample to $B^0 \to D^+D^-$ analysis.

Origin	σ_{p_1}	σ_{p_0}	$\sigma_{\Delta p_1}$	$\sigma_{\Delta p_0}$
Portability	-	-	-	-
$_s\mathcal{W}eights$	-0.009	-0.003	-0.03	-0.01
Calibration method	-0.01	0.01	na	na
Decay time acceptance	-0.0001	-0.0009	0.001	-0.00002
Decay time resolution	0.0006	0.0001	-0.0002	0.00002
detection asymmetry			0.004	0.0005
production asymmetry	0.0001		0.0001	
A_{sl}^d	-0.0003	-0.0001	-0.0003	0.0002
Δm_d	-	-	-	-
τ	-	-	-	-
Sum	0.01	0.01	0.03	0.01

Table 5.11: Systematic uncertainties on the flavour tagging calibration parameters for the SS sample to $B^0 \to D^+D^-$ analysis.

Origin	σ_{p_1}	σ_{p_0}	$\sigma_{\Delta p_1}$	$\sigma_{\Delta p_0}$
Portability				
$_s\mathcal{W}eights$	0.002	-0.003	0.05	-0.001
Calibration method	-0.01	0.009	na	na
Decay time acceptance	0.00001	-0.000004	0.00008	0
Decay time resolution	-0.0005	0.0001	0.0003	-0.00001
detection asymmetry	0.0004	-	0.007	0.0006
production asymmetry	-	-	0.0001	-
A_{sl}^d	-0.0001	-	0.001	0.0001
Δm_d	-0.001	0.0001	-	-
τ	-0.001	0.009	0.05	0.001
Sum	0.01	0.009	0.05	0.001

the one found in $B^0 \to D^+D^-$ analysis.

5.5 Summary

In this section the results on the flavour tagging calibration parameters which are used in the decay time fit to signal channels, including their statistical and systematic uncertainties are reported.

Calibrations to $B^0 \to D^+D^-$: In the final fit to data two sets of calibration parameters are used, corresponding to OS and SS combination respectively.

$$\begin{aligned}
p_1^{\mathrm{OS},Run1} &= 1.069 \pm 0.072\,(\mathrm{stat}) \pm 0.01\,(\mathrm{syst})\ ,\\
p_0^{\mathrm{OS},Run1} &= 0.3691 \pm 0.0080\,(\mathrm{stat}) \pm 0.01\,(\mathrm{syst})\ ,\\
\langle \eta^{\mathrm{OS},Run1} \rangle &= 0.3627\ ,\\
\Delta p_1^{\mathrm{OS},Run1} &= 0.03 \pm 0.11\,(\mathrm{stat}) \pm 0.03\,(\mathrm{syst})\ ,\\
\Delta p_0^{\mathrm{OS},Run1} &= -0.009 \pm 0.012\,(\mathrm{stat}) \pm 0.001\,(\mathrm{syst})\ ,
\end{aligned} \tag{5.35}$$

$$\begin{aligned}
p_1^{\mathrm{SS},Run1} &= 0.842 \pm 0.090\,(\mathrm{stat}) \pm 0.01\,(\mathrm{syst})\ ,\\
p_0^{\mathrm{SS},Run1} &= 0.4296 \pm 0.0060\,(\mathrm{stat}) \pm 0.009\,(\mathrm{syst})\ ,\\
\langle \eta^{\mathrm{SS},Run1} \rangle &= 0.4282\ ,\\
\Delta p_1^{\mathrm{SS},Run1} &= 0.07 \pm 0.13\,(\mathrm{stat}) \pm 0.05\,(\mathrm{syst})\ ,\\
\Delta p_0^{\mathrm{SS},Run1} &= -0.0065 \pm 0.0087\,(\mathrm{stat}) \pm 0.001\,(\mathrm{syst})\ .
\end{aligned} \tag{5.36}$$

Calibrations to $B^0 \to D^{*+}D^-$: In the final fit to data different calibration parameters for Run 1 and Run 2 are used, and for each Run a single set of calibration parameters is used for the two D^0 modes. The OS and SS combinations used in the fit to signal, are obtained averaging the calibration parameters from the two control channels, namely $B^0 \to D_s^+D^-$ and $B^0 \to D^{*+}D_s^-$.

$$\begin{aligned}
p_1^{\mathrm{OS},Run1} &= 1.107 \pm 0.068\,(\mathrm{stat})\ ,\\
p_0^{\mathrm{OS},Run1} &= 0.3679 \pm 0.0071\,(\mathrm{stat})\ ,\\
\langle \eta^{\mathrm{OS},Run1} \rangle &= 0.3675\ ,\\
\Delta p_1^{\mathrm{OS},Run1} &= 0.0226 \pm 0.0993\,(\mathrm{stat})\ ,\\
\Delta p_0^{\mathrm{OS},Run1} &= 0.013 \pm 0.010\,(\mathrm{stat})\ ,\\
\Delta \epsilon_{tag}^{\mathrm{OS},Run1} &= 0.015 \pm 0.0136\,(\mathrm{stat})\ ,
\end{aligned} \tag{5.37}$$

$$
\begin{aligned}
p_1^{\mathrm{SS},Run1} &= 0.927 \pm 0.092\,(\mathrm{stat})\ , \\
p_0^{\mathrm{SS},Run1} &= 0.4288 \pm 0.0056\,(\mathrm{stat})\ , \\
\langle \eta^{\mathrm{SS},Run1} \rangle &= 0.4301\ , \\
\Delta p_1^{\mathrm{SS},Run1} &= -0.0042 \pm 0.1288\,(\mathrm{stat})\ , \\
\Delta p_0^{\mathrm{SS},Run1} &= -0.0037 \pm 0.0079\,(\mathrm{stat})\ , \\
\Delta \epsilon_{tag}^{\mathrm{SS},Run1} &= -0.0046 \pm 0.0071\,(\mathrm{stat})\ , \\
A_{prod}^{Run1} &= -0.0037 \pm 0.0062\,(\mathrm{stat})\ , \\
A_{det}^{Run1} &= 0.0060 \pm 0.0086\,(\mathrm{stat})\ ,
\end{aligned} \tag{5.38}
$$

$$
\begin{aligned}
p_1^{\mathrm{OS},Run2} &= 0.912 \pm 0.080\,(\mathrm{stat})\ , \\
p_0^{\mathrm{OS},Run2} &= 0.3729 \pm 0.0079\,(\mathrm{stat})\ , \\
\langle \eta^{\mathrm{OS},Run2} \rangle &= 0.3693\ , \\
\Delta p_1^{\mathrm{OS},Run2} &= 0.1796 \pm 0.1087\,(\mathrm{stat})\ , \\
\Delta p_0^{\mathrm{OS},Run2} &= 0.013 \pm 0.011\,(\mathrm{stat})\ , \\
\Delta \epsilon_{tag}^{\mathrm{OS},Run2} &= -0.046 \pm 0.012\,(\mathrm{stat})\ ,
\end{aligned} \tag{5.39}
$$

$$
\begin{aligned}
p_1^{\mathrm{SS},Run2} &= 1.110 \pm 0.084\,(\mathrm{stat})\ , \\
p_0^{\mathrm{SS},Run2} &= 0.4310 \pm 0.0052\,(\mathrm{stat})\ , \\
\langle \eta^{\mathrm{SS},Run2} \rangle &= 0.4286\ , \\
\Delta p_1^{\mathrm{SS},Run2} &= -0.0954 \pm 0.1156\,(\mathrm{stat})\ , \\
\Delta p_0^{\mathrm{SS},Run2} &= 0.009 \pm 0.007\,(\mathrm{stat})\ , \\
\Delta \epsilon_{tag}^{\mathrm{SS},Run2} &= 0.007 \pm 0.004\,(\mathrm{stat})\ , \\
A_{prod}^{Run2} &= -0.0053 \pm 0.0063\,(\mathrm{stat})\ , \\
A_{det}^{Run2} &= 0.0417 \pm 0.0094\,(\mathrm{stat})\ .
\end{aligned} \tag{5.40}
$$

5.6 Studies in the signal channel $B^0 \to D^+D^-$

5.6.1 Correlation of the mistag estimates

Correlation between OS and SS mistag estimates Using the $_s$*weighted* nominal data sample the linear Pearson correlation coefficient [1] ρ is calculated in order to check if the two dimensions of η_{OS} and η_{SS} factorize, it results to be $\rho = -0.047$. The bootstrap method with 10 000 repetitions is used to clarify the significance of this value, getting a 95 % confidence level interval for the correlation coefficient: $(-0.109, 0.019)$. The linear correlation resulting is small, so that the factorization holds. In Fig. 5.21 the 2-dimensional distribution of η_{OS} vs. η_{SS} in the overlap sample is shown, while in Fig. 5.22 the profile histograms are visible.

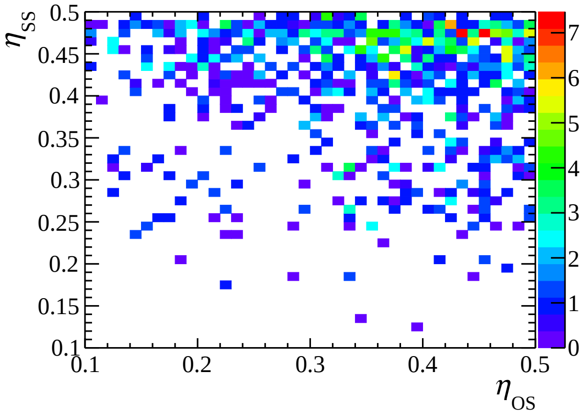


Figure 5.21: 2-dimensional histogram for η_{OS} and η_{SS} using an $_s$*weighted* data sample.

Correlation between decay time and OS/SS mistag estimates The correlation between the decay time distribution and the per-event mistag is also studied. A small correlation between the decay time and the OS/SS per-event mistag probability is shown in Table 5.12. In Fig. 5.23 the profile histogram of the OS and SS tagger are shown, while the first is rather flat and therefore no more action is taken the SS mistag seems to slowly increase with decay time. The validation of the correlation between the SS mistag and the

Table 5.12: $\eta_{\mathrm{OS/SS}}$ and B^0 decay time correlations and 95 % confidence levels.

	$\rho_{\eta,t}$	95 % CL
OS	-0.118	$-0.203, -0.033$
SS	0.056	$-0.005, \quad 0.116$

---

[1] `http://www.gnu.org/software/gsl/manual/html_node/Correlation.html`

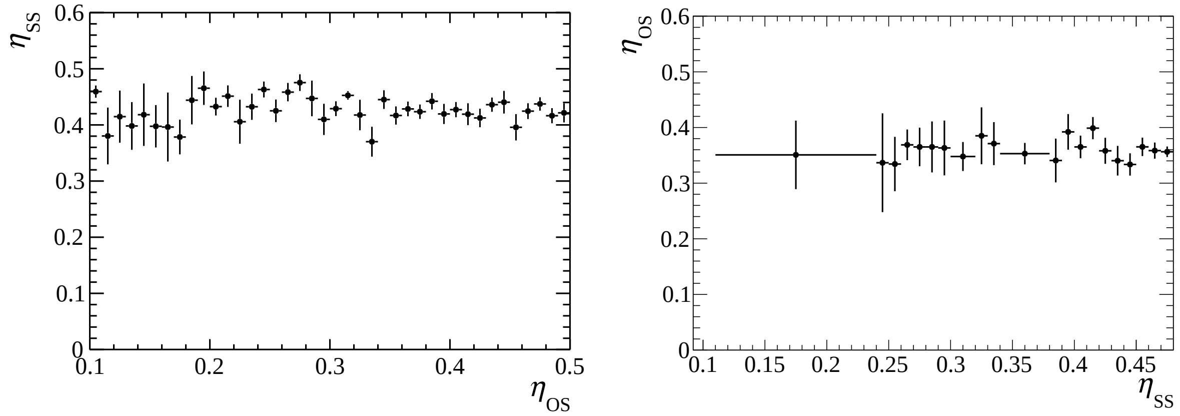


Figure 5.22: Profile histograms for η_{OS} and η_{SS}. Depicted are the mean value of η_{SS} (η_{OS}) and its error for each bin in η_{OS} and (η_{SS}) on the left (right) side using an *sweighted* data sample.

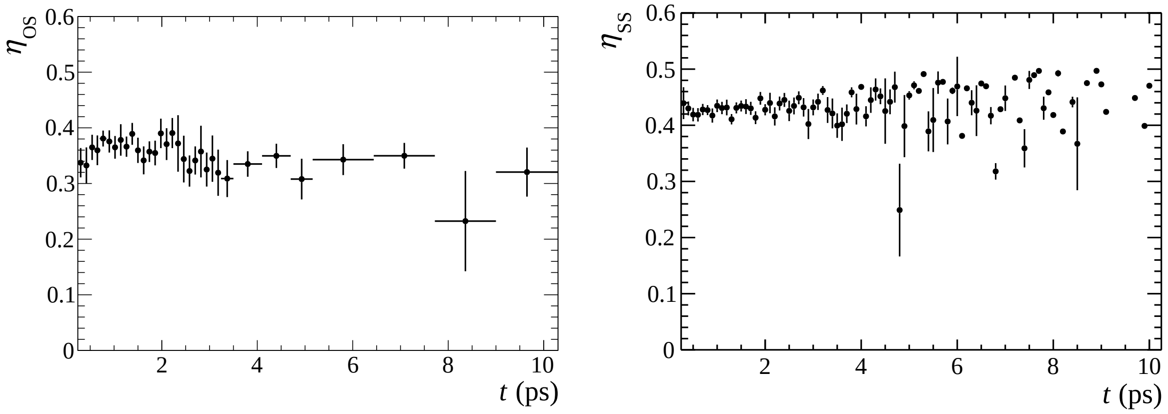


Figure 5.23: Profile histograms for the decay time dependence on η_{OS} and η_{SS}. Depicted are the mean value of η_{OS} (η_{SS}) and its error for each bin in t on the left (right) using an *sweighted* data sample.

decay time is done on the signal MC sample, performing a χ^2 fit using the linear function $\eta_{\mathrm{SS}} = a_{\eta_{\mathrm{SS}},t} t + b_{\eta_{\mathrm{SS}},t}$ in the the decay time range 0.25 ps to 8.25 ps. The slope is found to be $a_{\eta_{\mathrm{SS}},t} = (0.001\,50 \pm 0.000\,27)\ \mathrm{ps}^{-1}$. Even if a significant deviation from zero is found for the correlation, this is neglected in the nominal fit but included in the study of systematic uncertainties as described in Sec. 6.5.

5.6.2 Effective Tagging Efficiency

The effective tagging efficiency (see Eq. (5.3)) is calculated as sum of the effective tagging efficiencies that each signal candidate (including untagged candidates with $D = 0$) con-

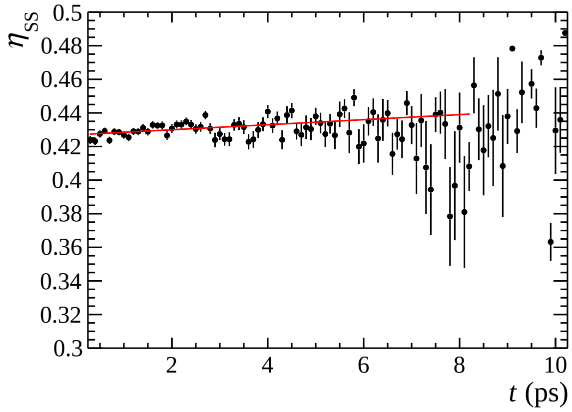


Figure 5.24: Profile histogram for the decay time dependence on η_{SS} for signal MC. The black datapoints represent the mean value of η_{SS} and its error for each bin in t. The red curve is the fitted linear function.

tributes. If more than one tagging algorithm is available, the calculation can be performed using the following formula:

$$\varepsilon_{\mathrm{eff}} = \frac{\sum_{i=0}^{\mathrm{N}} w_i \left(-d_{\mathrm{OS}}(1-2\omega_{\mathrm{OS}}) - d_{\mathrm{SS}}(1-2\omega_{\mathrm{SS}}) + d_{\mathrm{OS}} d_{\mathrm{SS}} \left(\omega_{\mathrm{OS}}^{B^0} - \omega_{\mathrm{OS}}^{\bar{B}^0} + \omega_{\mathrm{SS}}^{B^0} - \omega_{\mathrm{SS}}^{\bar{B}^0} - 2(\omega_{\mathrm{OS}}^{B^0}\omega_{\mathrm{SS}}^{B^0} - \omega_{\mathrm{OS}}^{\bar{B}^0}\omega_{\mathrm{SS}}^{\bar{B}^0}) \right) \right)^2}{\sum_{i=0}^{\mathrm{N}} w_i}, \tag{5.41}$$

where w_i are the ${}_s\mathcal{W}eights$. It expresses the effective dilution of the amplitude of the trigonometric terms. More details can be found in [78].

In Table 5.13, the effective tagging efficiency of the signal for exclusively OS tagged, exclusively SS tagged and exclusively OS and SS tagged candidates is given. These three exclusive samples sum up to the complete tagged data sample. The effective tagging efficiency of 8.1 % can be split into an overall tagging efficiency of $\varepsilon_{\mathrm{tag}} = (87.6 \pm 0.8)\,\%$ and an effective dilution of $D = 0.303 \pm 0.011$ which corresponds to an effective mistag probability of $\omega = 0.348 \pm 0.006$.

Table 5.13: Effective tagging efficiency

Tagger	$\varepsilon_{\mathrm{eff}}$ [%]
exclusively OS tagged	1.02 ± 0.09
exclusively SS tagged	1.36 ± 0.19
exclusively OS and SS tagged	5.7 ± 0.5
Total	8.1 ± 0.6

5.7 Studies in the signal channel $B^0 \to D^* + D^-$

5.7.1 Correlation of the mistag estimates

Two kind of possible correlations will be taken into account for the mistag. The first is the correlation between the OS and SS mistag in order to understand if they factorize, this check will be performed on ${}_s$*weighted* data after calibrating the mistag. Another correlation that will be evaluated is the one between the mistag and the decay time [79], and in case some correlation will be found, it will be included in the systematic uncertainties.

5.7.2 Effective Tagging Efficiency

Following what described in Sec. 5.6.2 the effective tagging efficiency on the signal $B^0 \to D^{*\pm}D^{\mp}$ for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ is calculated on Run2 sample, using the calibration parameters obtained from $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*+} D_s^-$ respectively. Three different samples are taken into account, composed by events which are OS tagged, SS tagged or tagged by their combination. The results are reported in Table 5.14 and Table 5.15, respectively.

Table 5.14: Effective tagging efficiency for D^0 to $K^-\pi^-\pi^+\pi^+$ Run 2, using the calibration parameters obtained from $B^0 \to D_s^+ D^-$

Tagger	ε_{eff} [%]
OS tagged	2.73 ± 0.35
SS tagged	3.81 ± 0.36
(OS + SS) tagged	5.92 ± 0.50

Table 5.15: Effective tagging efficiency for D^0 to $K^-\pi^-\pi^+\pi^+$ Run 2, using the calibration parameters obtained from $B^0 \to D^{*+} D_s^-$

Tagger	ε_{eff} [%]
OS tagged	3.81 ± 0.53
SS tagged	3.34 ± 0.34
(OS + SS) tagged	7.23 ± 0.64

Chapter 6

Measurement of CP Violation in $B^0 \to D^+D^-$ decay

In this chapter the measurement of CP Violation in the $B^0 \to D^+D^-$ decay will be described. The measurement is based on samples of pp collision data corresponding to a total integrated luminosity of 3 fb^{-1}, recorded by the LHCb experiment [80].

As described in Sec. 3.4, in the B0 meson system, CP violation is produced from the interference between the direct decay and the decay to the same final state after $B^0 - \bar{B}^0$ mixing. The observables associated to CP violation, namely S and C, are related to the B^0 mixing phase ϕ_d and a phase shift $\Delta\phi_d$ from the decay amplitude through Eq. (3.53). The measure of the decay-time-dependent decay rate expressed according to Eq. (3.55) gives access to these observables.

6.1 Selection

In this section the selection chain will be described. The data set collected by the LHCb experiment during 2011 and 2012 corresponds to an integrated luminosity of 1 and 2 fb^{-1}, at center-of-mass energies of 7 and 8 TeV respectively. For the reconstruction of the decay $B^0 \to D^+D^-$, between the possible final states for the D mesons (see Table 6.1)

Table 6.1: Branching fractions of possible $D^{\pm}$ final states.

Mode		$\mathcal{B}$ $[10^{-3}]$
D^+	$\to K^-\pi^+\pi^+$	94.6 ± 2.4
D^+	$\to K^+K^-\pi^+$	9.9 ± 0.3
D^+	$\to \pi^+\pi^+\pi^-$	3.3 ± 0.2

the two dominant ones are considered:

- $B^0 \to D^+D^- \to K^-\pi^+\pi^+K^+\pi^-\pi^-$

- $B^0 \to D^+D^- \to K^-K^+\pi^+K^+\pi^-\pi^-$

The selection chain starts with a two-layer trigger system which selects events from the proton-proton collisions. Then a first very general selection, called stripping, is applied to the triggered data which are written to tape. An offline cut-based selection is subsequently applied, after that vetoes on misidentified backgrounds, a multivariate classifier and a final cut-based selection in order to reduce the amount of background pollution.

Stripping Already during the process of tuples creation a set of conditions is applied. At first all tracks are required to be long tracks (see Sec. 4.3.2) and an event is only considered if it contains in total less than 500 long tracks. All the requirements applied on the daughter particles (kaons and pions) of the D mesons are listed in Table 6.2.

Table 6.2: Stripping cuts applied in the reconstruction and selection of the D meson daughters. $\Delta \ln \mathcal{L}_{K\pi}$ is the difference of the log-likelihood between a pion and a kaon hypothesis, which is used by the particle identification (PID) system to identify kaons from pions. IP χ^2 represents the difference between the vertex fit when including the track in the vertex and the vertex fit when excluding the track in the vertex.

track χ^2/ndf	< 3
p_{T}	$> 100\,\mathrm{MeV/c}$
p	$> 1000\,\mathrm{MeV/c}$
IP χ^2	> 4
ghost probability	< 0.3
π $\Delta \ln \mathcal{L}_{K\pi}$	< 20
K $\Delta \ln \mathcal{L}_{K\pi}$	> -10

Three of these hadron tracks have to form a common vertex. The transverse momentum of one of the daughter hadrons has to exceed (100 MeV/c) 1 GeV/c and its track χ^2/ndf has to be less than 3. The criteria that $D^\pm$ combination have to fulfill are given in Table 6.3. Finally the two D mesons have to form a common vertex, the requirement on the reconstructed B^0 are listed in Table 6.4

Preselection During this stage of tuple processing many of the cuts used in the stripping are reimplemented (see Table 6.5) using the DecayTreeFitter (DTF) [81]. This function allow to correctly comprise uncertainties and correlation on vertex position, invariant masses, decay times, flight distances and particle momenta. Constraints on the primary vertex position and on the invariant mass of both D mesons to the PDG values are used. Moreover the $D^\pm$ mass region is tightened to be ± 50 MeV/c^2 around its nominal value and a set of additional cuts to remove outlier events are applied (see Table 6.6). The total number of candidates after the preselection is about 110 000, in Fig. 6.1 the invariant D^+D^- mass distribution is shown.

Table 6.3: Stripping cuts on the $D^\pm$ combination. DOCA is the distance of the closest approach between all possible pairs of particles forming the vertex, vertex distance χ^2 represents the difference between the χ^2 of fitting both vertices as one and the sum of the χ^2 of vertex fit one and vertex fit two, while DIRA is the cosine of the angle between the momentum of the D meson and the direction from the best PV to the decay vertex.

$m_{D^\pm}$	$1769.62\,\mathrm{MeV/c^2} - 2068.49\,\mathrm{MeV/c^2}$
$\sum p_\mathrm{T}\ (hhh)$	$> 1800\,\mathrm{MeV/c}$
DOCA	$< 0.5\,\mathrm{mm}$
$D^\pm\ \chi^2_\mathrm{vtx}/\mathrm{ndf}$	< 10
vertex distance χ^2 to any PV	> 36
DIRA	> 0

Table 6.4: Stripping cuts on B^0 combination.

$B^0\ \chi^2_\mathrm{vtx}/\mathrm{ndf}$	< 10
m_{B^0}	$4750\,\mathrm{MeV/c^2} - 6000\,\mathrm{MeV/c^2}$
$p_\mathrm{T}(D^+) + p_\mathrm{T}(D^-)$	$> 5\,\mathrm{GeV/c}$
t	$> 0.2\,\mathrm{ps}$
IP χ^2	< 25
DIRA	> 0.999
p	$> 10\,\mathrm{GeV/c}$
BDT($p_\mathrm{T}(B^0)$, FD $\chi^2(B^0)$)	> 0.05

Table 6.5: Cuts reimplemented using DTF.

$p_\mathrm{T}(K,\pi)$	$> 100\,\mathrm{MeV/c}$
$p(K,\pi)$	$> 1000\,\mathrm{MeV/c}$
IP $\chi^2(K,\pi)$	> 4
$\sum p_\mathrm{T}$ (D daughters)	$> 1800\,\mathrm{MeV/c}$
t	$> 0.2\,\mathrm{ps}$
$p_\mathrm{T}(D^+) + p_\mathrm{T}(D^-)$	$> 5\,\mathrm{GeV/c}$
$p(B^0)$	$> 10\,\mathrm{GeV/c}$
IP $\chi^2(B^0)$	< 25
m_{B^0} (D^+D^- mass + PV constraint)	$4750\,\mathrm{MeV/c^2} - 6000\,\mathrm{MeV/c^2}$

Vetos Misidentification of kaons and protons as pions for the final state particles of D mesons leads to backgrounds, which are suppressed by applying explicit vetos.

For $D^\pm$ mesons which are reconstructed in the $D^\pm \to K^\mp\pi^\pm\pi^\pm$ final state, a $K \to \pi$ mis-ID can lead to background contributions from $D_s^\pm \to K^\mp K^\pm\pi^\pm$. To reduce them the kaon mass hypothesis is assigned to the pion with the higher transverse momentum. The

Table 6.6: Offline cuts to remove outlier events.

σ_m	$< 30\,\mathrm{MeV/c^2}$
σ_t	$< 0.2\,\mathrm{ps}$
$\lvert z_{PV}\rvert$	$< 250\,\mathrm{mm}$

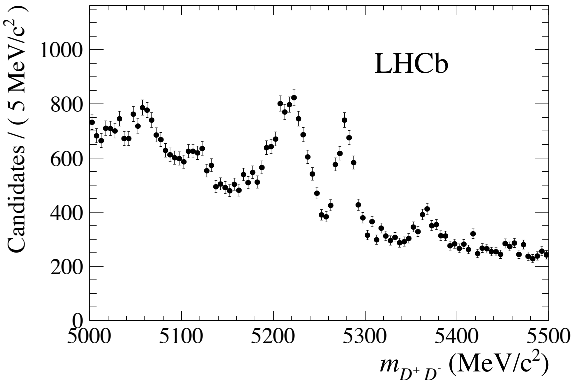


Figure 6.1: Plot of the invariant D^+D^- mass of the candidates which survive the preselection.

invariant mass of the three hadrons and the invariant mass of the hypothetical kaon pair are recalculated. The candidate is rejected if the invariant mass of the kaon pair is compatible with the ϕ mass of $M_\phi = 1019.461\,\mathrm{MeV/c^2}$ [28], within $\pm 10\,\mathrm{MeV/c^2}$. Additionally, the ProbNN variables from the TupleToolPid package [82], which deliver a probability for a particle to be a kaon, pion, proton, muon, electron or a ghost, is used. If for the pion with the higher p_T the kaon probability is larger than the pion probability and the invariant mass of the three hadrons is compatible with the $D_s^\pm$ mass of $M_{D_s^\pm} = 1968.30\,\mathrm{MeV/c^2}$ [28], within $\pm 25\,\mathrm{MeV/c^2}$ the candidate is also rejected. A plot showing the distributions before and after applying this veto is given in Fig. 6.2.

To reduce $p \to \pi$ mis-ID the proton mass hypothesis is assigned to the pion with the higher p_T of the $D^\pm \to K^\mp\pi^\pm\pi^\pm$ final state and the invariant mass of the $Kp\pi$ combination is calculated. The candidate is rejected if the proton probability `ProbNNp` for the pion with the higher p_T is larger than the pion probability and the invariant mass of the $Kp\pi$ combination is compatible with the $\Lambda_c^\pm$ mass of $M_{\Lambda_c^\pm} = 2286.46\,\mathrm{MeV/c^2}$ [28], within $\pm 25\,\mathrm{MeV/c^2}$. As can be seen from Fig. 6.3 this veto reduces the excess of candidates around the $\Lambda_c^\pm$ mass.

In order to estimate the presence of possible single or double charmless background, a study in the mass region $\pm$ 40 MeVc^2 around nominal D mass is performed. The nominal part of the D mass window ($\pm$ 25 MeVc^2) is vetoed for one or both D meson and a mass fit is performed considering the DTF mass variable with PV constraint but without D mass constraints and applying all the DTF cuts. From this fit is observed that the only source of charmless background is due to $B^0 \to K^-K^+\pi^+D^-(K^+\pi^-\pi^-)$, the estimated yield is

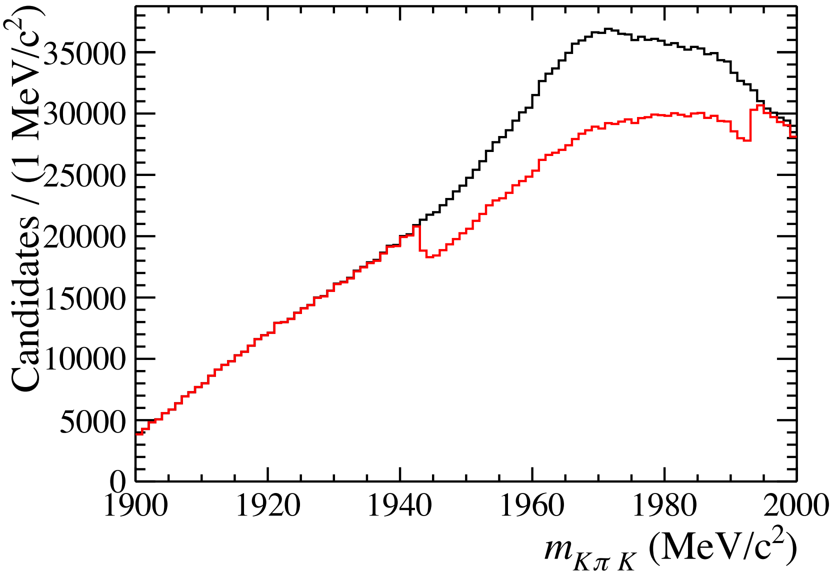


Figure 6.2: Plot of the invariant mass of the $K\pi K$ combination. The distribution is given without (black) and with (red) the $D_s^{\pm}$ veto described in the text.

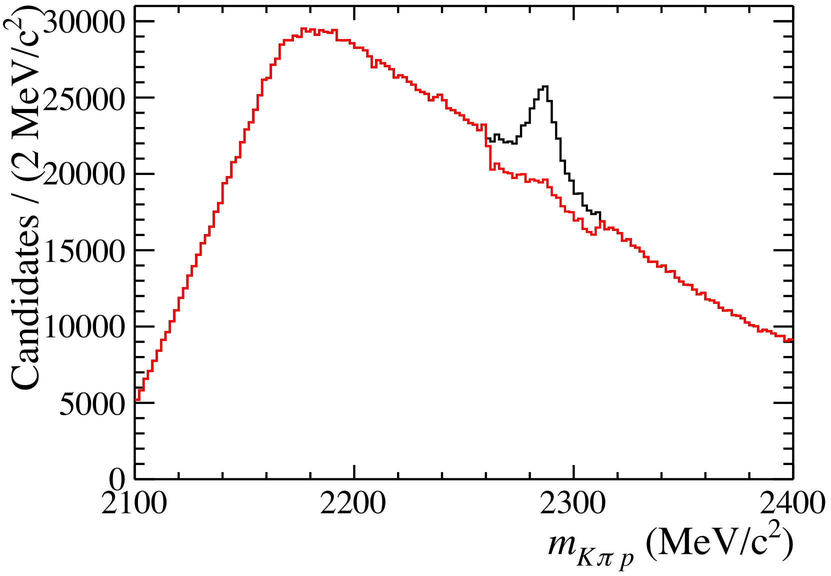


Figure 6.3: Plot of the invariant mass of the $K\pi p$ combination. The distribution is given without (black) and with (red) the $\Lambda_c^{\pm}$ veto described in the text.

$N_{KK\pi D} = 28.7 \pm 19.5$. To reduce the amount of this partially charmless contribution further selection requirement are applied, which are listed in Table 6.7.

Table 6.7: Cuts to reject charmless background contributions.

$\lvert m_{hhh} - m_{D^\pm} \rvert$	$< 25\,\mathrm{MeV/c^2}$
D vertex separation	> 2
$t_{D^\pm}/\sigma_{t_{D^\pm}}$	> 0 ($K^-\pi^+\pi^+K^+\pi^-\pi^-$)
$t_{D^\pm}/\sigma_{t_{D^\pm}}$	> 3 ($K^-K^+\pi^+K^+\pi^-\pi^-$)

Multivariate Classifier The number of combinatorial background candidates is suppressed making use of a multivariate algorithm, in this case a Boosted Decision Tree (BDT) [83, 84]. A signal MC sample serves as signal template, while the upper mass sideband above $5500\,\mathrm{MeV/c^2}$ of the data sample is used to model the combinatorial background, for the mass fit the range considered is 5150-5500 $\mathrm{MeV/c^2}$, so that the sample used for the training and the one used for the mass fit are independent. The BDT is trained on half of these samples while its performances are tested on the other half. The previous cut-based preselection and the vetoes, are already applied to the training samples. Since PID variables are used in the BDT, two separate trainings, for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ and the $K^-K^+\pi^+K^+\pi^-\pi^-$ final states have to be performed. The 21 input variables, ordered according to their ranking are listed in Table 6.8. Each BDT is built out of 700 trees, and the depth of the trees is limited to three. At each node at least 3 % of the training events have to be present. The variables are scanned at 40 points to find the optimal cut value. For the boosting the AdaBoost method [85] with a boost factor of $\beta = 0.1$ is deployed. The BDT outputs, with training and testing samples superimposed are presented in Figs. 6.4 and 6.5.

The optimization of the cut value to apply on the BDT output classifier is done on data. At first, the optimal cut value for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state BDT is determined. For each cut point on the BDT output classifier fits to the invariant D^+D^- mass spectrum are performed. Using the yields of the mass fits ${}_s\mathcal{W}eights$ are calculated. These are then used to perform fits on the signal decay time distributions, from which the uncertainties on *CP* observables $S_{D^+D^-}$ and $C_{D^+D^-}$ are determined. The result of this scanning procedure, showing the dependency of uncertainties on *CP* observables as a function of the BDT cut is shown in Fig. 6.6. The cut on the BDT classifier is chosen greater than -0.10, this should be a good compromise between both observables as the uncertainties of $S_{D^+D^-}$ and $C_{D^+D^-}$ are almost the same here and close to their minima. The chosen BDT cut has a signal efficiency of $(96.5 \pm 0.5)\,\%$ and rejects $(84.18 \pm 0.34)\,\%$ of the combinatorial background. The same optimization cannot be performed for the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state because the sample is too small for a reliable fit at each cut value. For this reason, a simultaneous fit to both final states is performed and the BDT cut for the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state is varied. For the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ subsample the previously determined BDT cut is applied at the optimized value. In this case the uncertainties on *CP* observables show a minimum at around -0.05 which is chosen as cut value. This cut removes $(90.75 \pm 0.33)\,\%$ of the combinatorial background at a signal efficiency of $(87.2 \pm 1.9)\,\%$.

In Fig. 6.8 the invariant D^+D^- mass distribution of candidates that remain after

Table 6.8: List of input variables used in the training of the BDT.

BDT for $K^-\pi^+\pi^+K^+\pi^-\pi^-$	BDT for $K^-K^+\pi^+K^+\pi^-\pi^-$
min($D^\pm$ τ significance)	PID ratio of $K^\pm$
B direction angle	B direction angle
log(DTF χ^2/ndof)	PID ratio of K^+
PID ratio of K^-	log(DTF χ^2/ndof)
PID ratio of K^+	PID ratio of K^-
min p_T of $K^\pm$	min($D^\pm$ τ significance)
log(B impact parameter χ^2)	log(min(h Velo χ^2/ndof))
log(min($\pi^\pm$ Velo χ^2/ndof))	p_T of $K^\pm$
p_T of π^- with lower p_T	log(min($K^\pm$ T-track χ^2/ndof))
log(min($K^\pm$ T-track χ^2/ndof))	log(B impact parameter χ^2)
log(min($\pi^\pm$ T-track χ^2/ndof))	PID ratio of $\pi^\pm$ with lower p_T
PID ratio of π^- with higher p_T	log(min(h VELO-T-Match χ^2))
p_T of π^+ with lower p_T	log(min($K^\pm$ Velo χ^2/ndof))
PID ratio of π^- with lower p_T	PID ratio of single $\pi^\pm$
PID ratio of π^+ with higher p_T	p_T of $\pi^\pm$ with higher p_T
p_T of π^+ with higher p_T	log(min(h T-track χ^2/ndof))
PID ratio of π^+ with lower p_T	p_T of $\pi^\pm$ with lower p_T
log(min($K^\pm$ Velo χ^2/ndof))	min p_T of K^+ and K^-
log(min($\pi^\pm$ VELO-T-Match χ^2))	p_T of single $\pi^\pm$
log(min($K^\pm$ VELO-T-Match χ^2))	log(min($K^\pm$ VELO-T-Match χ^2))
p_T of π^- with higher p_T	PID ratio of $\pi^\pm$ with higher p_T

applying the BDT cuts is shown, while in Fig. 6.9 and Fig. 6.10 the D meson mass distributions splitted by decay mode are presented.

Final Selection The final selection step reduces the range of the invariant D^+D^- mass from 5150 MeV/c^2 to 5500 MeV/c^2 and the range of the decay time from 0.25 ps to 10.25 ps. The signal efficiency of these cuts is nearly 100 %. An analysis of all candidates surviving the previously described selection reveals that in 49 events there is more than one candidate, since they have a-priori the same possibility to be the true signal candidate one of the candidates is chosen randomly. After all selection steps 6209 candidates remain in the data sample, of which about 25 % are signal candidates (see Sec. 6.2). In Table 6.9 and Table 6.10 the signal efficiency for the different selection steps are quoted, these are estimated on MC samples.

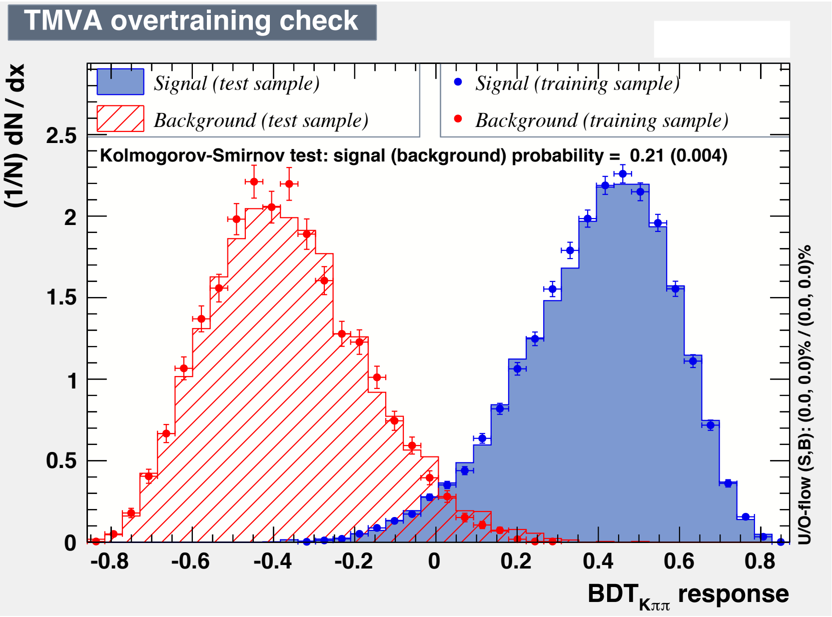


Figure 6.4: Comparison of BDT response on training and test sample for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state.

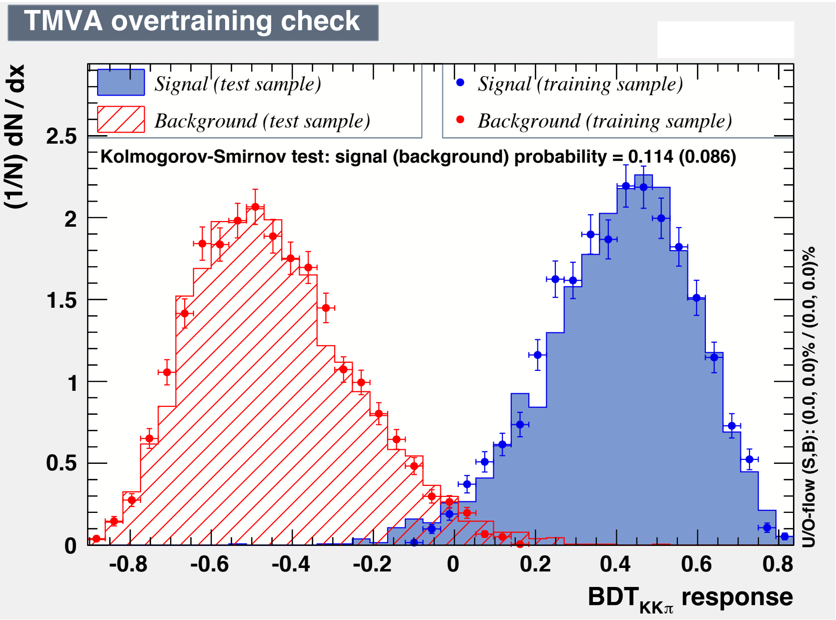


Figure 6.5: Comparison of BDT response on training and test sample for the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state.

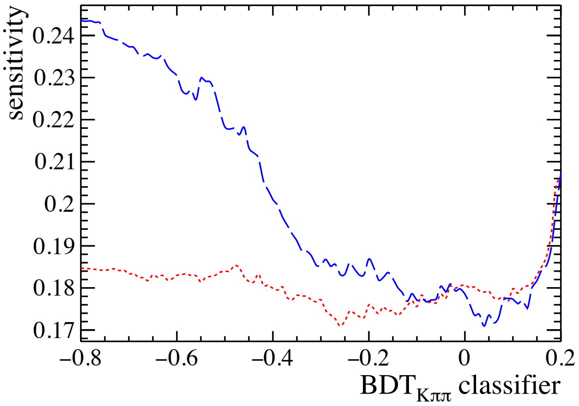


Figure 6.6: Sensitivity of $S_{D^+D^-}$ (red short-dashed) and $C_{D^+D^-}$ (blue long-dashed) as a function of the BDT output classifier for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state.

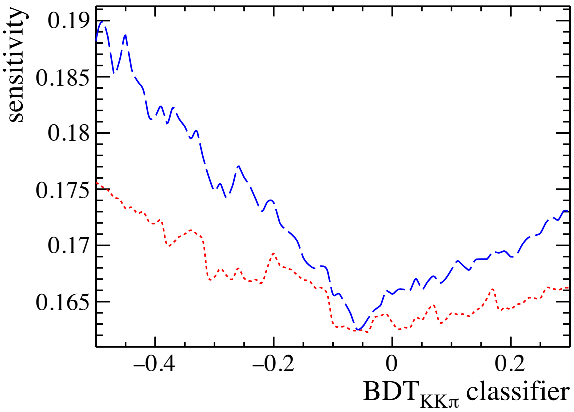


Figure 6.7: Sensitivity of $S_{D^+D^-}$ (red short-dashed) and $C_{D^+D^-}$ (blue long-dashed) as a function of the BDT output classifier for the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state.

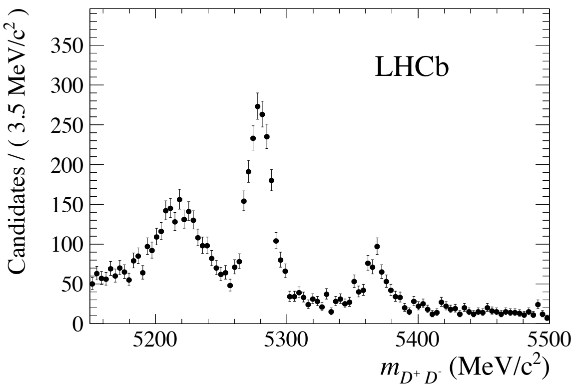


Figure 6.8: Plot of the invariant D^+D^- mass of the candidates after applying the BDT cuts.

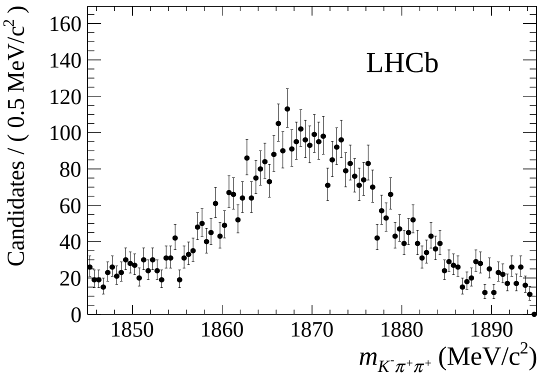


Figure 6.9: Mass distribution of the positive (negative) D meson on the left (right) in the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state, after applying the BDT cut.

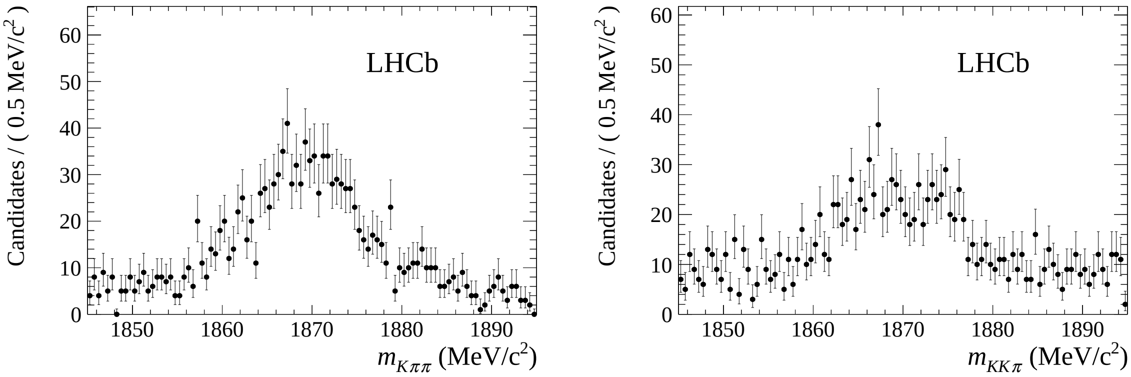


Figure 6.10: Mass distribution of D mesons reconstructed in the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state, after applying the BDT cut.

Table 6.9: Overall offline selection performance of $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state.

	$\varepsilon_{\text{Sig}}(\%)$
Preselection	82.17 ± 0.19
Vetoes	95.74 ± 0.10
BDT	96.5 ± 0.5
Total	75.92 ± 0.33

Table 6.10: Overall offline selection performance of $K^-K^+\pi^+K^+\pi^-\pi^-$ final state.

	$\varepsilon_{\text{Sig}}(\%)$
Preselection	67.5 ± 0.5
Vetoes	98.21 ± 0.14
BDT	87.2 ± 1.9
Total	57.8 ± 1.4

6.2 Mass Fit

The fit to the invariant mass, described in this section, is used to discriminate between signal and background candidates, from it ${}_s\mathcal{W}eights$ are calculated. These can be applied to other observables under the condition that these are uncorrelated with the invariant mass. In order to check if signal ${}_s\mathcal{W}eights$ can be used in the decay time fit, on signal MC the linear Pearson correlation coefficient between the decay time and the invariant mass is calculated. It results to be $\rho = 0.007$, so the ${}_s\mathcal{W}eights$ can be applied in the decay time fit from which the *CP* observables are extracted. The decay time fit is described in Sec. 6.3. For the mass fit the extended maximum likelihood method is used (Eq. (5.8)).

The mass fit is performed simultaneously in four disjoint categories: the two years of data-taking 2011 and 2012 and the two final states $K^-\pi^+\pi^+K^+\pi^-\pi^-$ and $K^-K^+\pi^+K^+\pi^-\pi^-$ ($s = \{2011, K\pi\pi\}, \{2011, KK\pi\}, \{2012, K\pi\pi\}, \{2012, KK\pi\}$).

The invariant D^+D^- mass is restricted to candidates in the range $5150\,\mathrm{MeV/c^2}$ to $5500\,\mathrm{MeV/c^2}$, leaving enough candidates in the upper mass sideband to determine the shape of the combinatorial background, while eliminating the background at low masses like $B^0 \to D^{*\pm}D^{\mp}$.

The probability density function (PDF) $\mathcal{P}^s$ consists of five components $\mathcal{P}^s_j$: $B^0 \to D^+D^-$ signal ($j = B^0$), $B^0_s \to D^+D^-$ background ($j = B^0_s$), background from $B^0 \to D^+_sD^-$ ($j = B^0 \to D^+_sD^-$), background from $B^0_s \to D^-_sD^+$ ($j = B^0_s \to D^-_sD^+$) and combinatorial background ($j = \mathrm{Comb}$):

$$N^s\mathcal{P}^s = N^s_{B^0}\mathcal{P}^s_{B^0} + N^s_{B^0_s}\mathcal{P}^s_{B^0_s} + N^s_{B^0 \to D^+_sD^-}\mathcal{P}^s_{B^0 \to D^+_sD^-} + N^s_{B^0_s \to D^-_sD^+}\mathcal{P}^s_{B^0_s \to D^-_sD^+} + N^s_{\mathrm{Comb}}\mathcal{P}^s_{\mathrm{Comb}} \quad . \tag{6.1}$$

B^0 signal The reconstructed mass of the B^0 signal component is parametrized with the sum of three crystal-ball PDFs [74], having the same mean μ_{B^0} but different width parameters σ_i and tails to opposite directions ($\alpha_1, \alpha_2 > 0$, $\alpha_3 < 0$). The exponent of the power law part n is fixed to 10. To determine the shape parameters, a maximum-likelihood fit to the invariant D^+D^- mass distribution of $B^0 \to D^+D^-$ signal MC in the range $4800\,\mathrm{MeV/c^2}$ to $5400\,\mathrm{MeV/c^2}$ is performed. The MC sample is made of a mixture of both D final states generated in the ratio of the current world averages and the full selection is applied. The fit results are listed in Table 6.11 and a plot of the distribution with the projection of the PDF is given in Fig. 6.11.

$B^0_s \to D^+D^-$ background The heavier B^0_s decaying via the same decay channel as the signal is also included. It is parametrized with the same PDF as the B^0 signal component sharing the width and tail parameters. The difference between the peak positions is fixed to the world average $\Delta m_{B^0_s - B^0} = \mu_{B^0_s} - \mu_{B^0} = 87.35\,\mathrm{MeV/c^2}$ [28].

$B \to D_sD$ background A significant amount of candidates from $B^0 \to D^+_sD^-$ remains in the data sample, although an explicit veto is applied in the selection to remove this background. This component is described with the sum of two crystal-ball PDFs. The power law exponent n is fixed to 10. While the widths and the mean of the PDF are

Table 6.11: Fit results of the mass fit to $B^0 \to D^+D^-$ signal MC.

Parameter		Value
$\mu_{B^0}^{\rm MC}$	(MeV/c^2)	5279.70 ± 0.09
$\sigma_1^{\rm MC}$	(MeV/c^2)	8.5 ± 0.4
$\sigma_2^{\rm MC}$	(MeV/c^2)	16 ± 5
$\sigma_3^{\rm MC}$	(MeV/c^2)	9.0 ± 0.4
$f_1^{\rm MC}$		0.48 ± 0.06
$f_2^{\rm MC}$		0.0098 ± 0.0011
$\alpha_1^{\rm MC}$		1.18 ± 0.08
$\alpha_2^{\rm MC}$		0.12 ± 0.04
$\alpha_3^{\rm MC}$		-1.46 ± 0.08

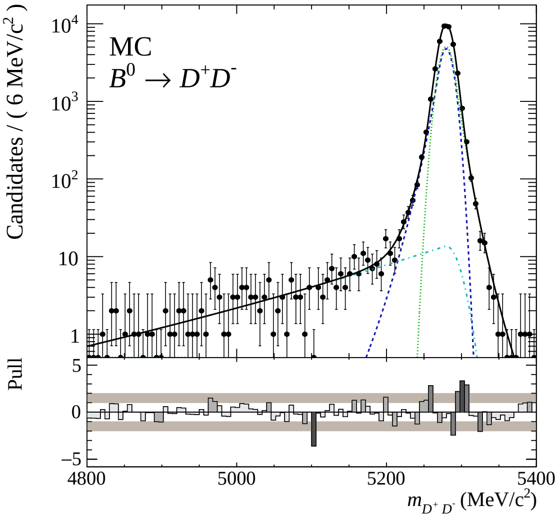


Figure 6.11: Mass distribution of the $B^0 \to D^+D^-$ signal MC sample overlaid with the projection of the fitted PDF. The y-axis has a logarithmic scale.

floating parameters in the fit to the invariant mass distribution. The fraction parameter and the tail parameters are determined on simulated $B^0 \to D_s^+D^-$ events which have been reconstructed as $B^0 \to D^+D^-$. On this MC sample the full selection is applied. The fit results are listed in Table 6.12, while the corresponding plot is shown in Fig. 6.12. In order to check that the $B^0 \to D_s^+D^-$ background does not have a correlation between mass and decay time, a comparison between decay time distributions for the low mass region ($m_{D^+D^-} < 5240$ MeV/c^2) and high mass region ($m_{D^+D^-} > 5320$ MeV/c^2) is done, showing that no correlation is present.

Table 6.12: Fit results of the mass fit to $B^0 \to D_s^+ D^-$ MC.

Parameter		Value
$\mu^{\rm MC}_{B^0\to D_s^+ D^-}$	(MeV/c^2)	5222.2 ± 0.9
$\sigma^{\rm MC}_{1,D_s^+ D^-}$	(MeV/c^2)	15.0 ± 1.5
$\sigma^{\rm MC}_{2,D_s^+ D^-}$	(MeV/c^2)	20.7 ± 2.1
$f^{\rm MC}_{1,D_s^+ D^-}$		0.78 ± 0.13
$\alpha^{\rm MC}_{1,D_s^+ D^-}$		0.60 ± 0.09
$\alpha^{\rm MC}_{2,D_s^+ D^-}$		-1.8 ± 0.4

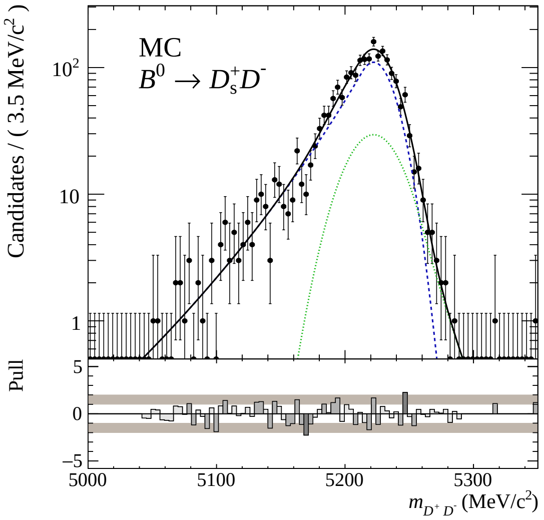


Figure 6.12: Mass distribution of the $B^0 \to D_s^+ D^-$ MC sample reconstructed as $B^0 \to D^+ D^-$ overlaid with the projection of the two crystal-ball PDFs. The y-axis has a logarithmic scale.

$B_s^0 \to D_s^- D^+$ background This decay channel is included in the nominal fit even if only few candidates are expected. It is parametrized with the sum of two crystal-ball PDFs, the mean parameter is constrained to the sum of the mean parameter of the B^0 signal component and the mass difference $\Delta m_{B_s^0 - B^0}$. All the other shape parameters are shared with the ones used for the B^0 component.

Combinatorial background The combinatorial background component is parametrized using an exponential function. Individual exponent $\beta^{K\pi\pi}$ and $\beta^{KK\pi}$ are used for the two different D final states and they are left free in the fit.

The fit results for the floating shape parameters and for the yields of the mass fit are listed in Table 6.13 and Table 6.14 respectively. The total number of B^0 signal candidates results to be $N_{B^0} = 1610 \pm 50$. In Fig. 6.13 the full data sample is plotted along with the PDF projections and its components.

Table 6.13: Results of the floating shape parameters in the mass fit.

Parameter		Value
μ_{B^0}	(MeV/c^2)	5279.26 ± 0.29
R_{B^0}		0.995 ± 0.032
$\mu_{D_s^+ D^-}$	(MeV/c^2)	5218.2 ± 1.1
$\sigma_{1,D_s^+ D^-}$	(MeV/c^2)	19.2 ± 2.7
$\sigma_{2,D_s^+ D^-}$	(MeV/c^2)	14.3 ± 3.1
$\beta_{K^-\pi^+\pi^+K^+\pi^-\pi^-}$	(1/(MeV/c^2))	-0.0031 ± 0.0005
$\beta_{K^-K^+\pi^+K^+\pi^-\pi^-}$	(1/(MeV/c^2))	-0.0041 ± 0.0006

Table 6.14: Results for the yields in the mass fit.

Sample		Parameter	Fitted Value
2011	$K\pi\pi$	N_{B^0}	397 ± 24
		$N_{B_s^0}$	112 ± 14
		$N_{B^0 \to D_s^+ D^-}$	354 ± 37
		$N_{B_s^0 \to D_s^- D^+}$	0 ± 23
		N_{Bkg}	664 ± 48
	$KK\pi$	N_{B^0}	85 ± 12
		$N_{B_s^0}$	20 ± 7
		$N_{B^0 \to D_s^+ D^-}$	107 ± 19
		$N_{B_s^0 \to D_s^- D^+}$	0 ± 4
		N_{Comb}	249 ± 26
2012	$K\pi\pi$	N_{B^0}	950 ± 38
		$N_{B_s^0}$	234 ± 20
		$N_{B^0 \to D_s^+ D^-}$	936 ± 65
		$N_{B_s^0 \to D_s^- D^+}$	79 ± 34
		N_{Bkg}	1104 ± 89
	$KK\pi$	N_{B^0}	178 ± 17
		$N_{B_s^0}$	33 ± 9
		$N_{B^0 \to D_s^+ D^-}$	221 ± 30
		$N_{B_s^0 \to D_s^- D^+}$	0 ± 38
		N_{Comb}	486 ± 40

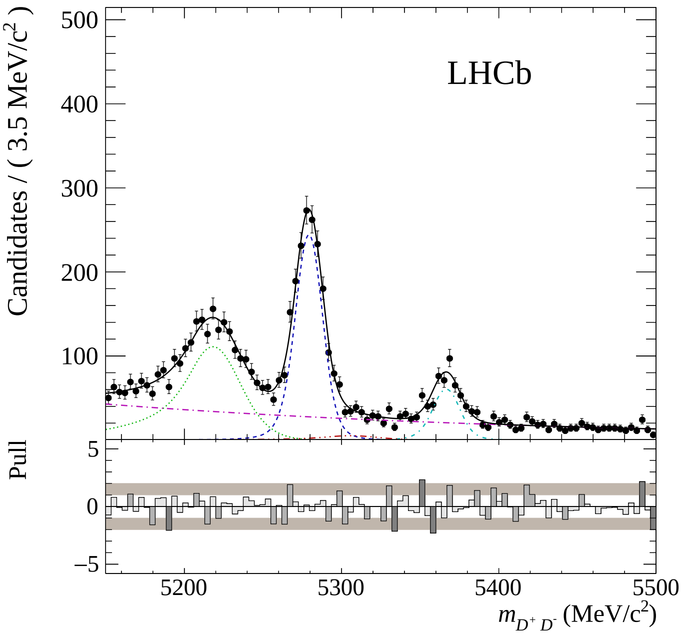


Figure 6.13: Plot of the reconstructed mass of the $B^0 \to D^+D^-$ data sample with the projected PDF and pull distributions. Besides the data points and the full PDF (solid black) the projections of the B^0 signal (dashed blue), the $B_s^0 \to D^+D^-$ background (short-dash-dotted turquoise), the $B^0 \to D_s^+D^-$ background (dotted green), the $B_s^0 \to D_s^-D^+$ background (long-dash- three-dotted red) and the combinatorial background (long-dash-dotted purple) are shown.

6.3 Decay Time Fit

The results of the mass fit described in Sec. 6.2, allow to determine signal weights using the *sPlot* method. These weights are then used in an sFit with a PDF that describes the signal distribution of observables related to the reconstructed decay time and to the flavour tagging output. Through this fit the CP observables can be obtained.

The conditional PDF describing the distribution of reconstructed decay time t and tag decisions $\vec{d} = (d_{\mathrm{OS}}, d_{\mathrm{SS}})$, given a per-event resolution σ_t and the mistag probability estimates $\vec{\eta} = (\eta_{\mathrm{OS}}, \eta_{\mathrm{SS}})$ can be expressed as

$$P\left(t, \vec{d} \mid \sigma_t, \vec{\eta}\right) = \epsilon(t) \cdot \left(\mathcal{P}(t', \vec{d} \mid \vec{\eta}) \otimes \mathcal{R}(t - t' \mid \sigma_t)\right) \ . \tag{6.2}$$

The B physics PDF $\mathcal{P}$ describes the distribution in flavour tags $\vec{d}$ and true decay time t', and is convolved with the resolution function $\mathcal{R}(t - t' \mid \sigma_t)$ (see Sec. 6.3), while the function $\epsilon(t)$ called decay time acceptance, reflects the relative efficiency change of all reconstruction and selection steps as a function of the reconstructed decay time t (see Sec. 6.3).

To construct $\mathcal{P}(t', \vec{d}|\vec{\eta})$, it is useful to start from the theoretical decay rates written in Eq. (6.3)

$$\begin{aligned}
\frac{\mathrm{d}\Gamma_{B^0}(t)}{\mathrm{d}t} &= \frac{\tilde{A}_{B^0}\,\mathrm{e}^{-t/\tau}}{2}\left(\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D\sinh\left(\frac{\Delta\Gamma t}{2}\right) - S\sin\left(\Delta m t\right) + C\cos\left(\Delta m t\right)\right),\\
\frac{\mathrm{d}\Gamma_{\overline{B}^0}(t)}{\mathrm{d}t} &= \frac{\tilde{A}_{\overline{B}^0}\,\mathrm{e}^{-t/\tau}}{2}\left(\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D\sinh\left(\frac{\Delta\Gamma t}{2}\right) + S\sin\left(\Delta m t\right) - C\cos\left(\Delta m t\right)\right),
\end{aligned} \tag{6.3}$$

and neglect *CP* violation in the mixing, while introducing the "true tag" d', which takes the value $+1$ (-1) for a true B^0 ($\overline{B}^0$) flavour at production. Experimentally, possible differences in the production rates of B^0 ($R_{\overline{B}^0}$) and $\overline{B}^0$ (R_{B^0}), which can be expressed via the production asymmetry A_P, have to be taken into account (see Eq. (6.4))

$$A_\mathrm{P} = \frac{R_{\overline{B}^0} - R_{B^0}}{R_{\overline{B}^0} + R_{B^0}}. \tag{6.4}$$

The resulting PDF for the true distribution of tag and decay time can be written as in Eq. (6.5)

$$\mathcal{P}(t', d') = \mathcal{N}_{t',d'}\,\frac{1 - d'\,A_\mathrm{P}}{2}\,\mathrm{e}^{-t'/\tau}\big(H(t') - d'\,T(t')\big), \tag{6.5}$$

where

$$\begin{aligned}
H(t') &= \cosh\left(\frac{\Delta\Gamma t'}{2}\right) + D\sinh\left(\frac{\Delta\Gamma t'}{2}\right),\\
T(t') &= S\sin\left(\Delta m t'\right) - C\cos\left(\Delta m t'\right),
\end{aligned} \tag{6.6}$$

and the normalization factor $\mathcal{N}_{t',d'}$ is chosen so that

$$\sum_{d'=-1,+1}\int_{t'_\mathrm{min}}^{t'_\mathrm{max}}\mathrm{d}t'\,\frac{1 - d'\,A_\mathrm{P}}{2}\,\mathrm{e}^{-t'/\tau}\big(H(t') - d'\,T(t')\big) = 1. \tag{6.7}$$

Moreover the PDF $\mathcal{P}(t, \vec{d})$ used in the fit to data needs to describe the observed tags $\vec{d}$. The effect of imperfect flavour tagging can be explained following the description given in [78].

The tagging efficiency, that describes the fraction of events that have a tag decision $d_i \neq 0$, can depend on the initial flavour, so it can be written as

$$\tilde{\varepsilon}_{\mathrm{tag},i}(d') = \varepsilon_{\mathrm{tag},i} - d'\,\frac{\Delta\varepsilon_{\mathrm{tag},i}}{2}. \tag{6.8}$$

The mistag estimates η_i can be transformed into the true mistag probabilities $\tilde{\omega}_i$ by applying a calibration function

$$\tilde{\omega}_i(\eta_i, d') = \omega_i(\eta_i) + d'\,\frac{\Delta\omega_i(\eta_i)}{2}, \tag{6.9}$$

that depends on the true initial flavour of the reconstructed B^0 meson. To simplify the notation in the following, $\omega_{i,d'} = \tilde{\omega}_i(\eta_i, d')$ and $\varepsilon_{i,d'} = \tilde{\varepsilon}_{\mathrm{tag},i}(d')$ will be used. For multiple taggers, the probability to observe a specific combination of tag decisions $\vec{d}$, assuming that the decisions and mistag estimates of the different tagging algorithms are independent, is given by

$$\mathcal{P}(\vec{d}\,|\,d') = \mathcal{P}(d_1, \ldots, d_N \mid d') = \prod_{i=1}^{N} \mathcal{P}(d_i \mid d'), \tag{6.10}$$

where

$$\begin{aligned} \mathcal{P}(d_i \mid d') &= \varepsilon_{i,d'} \Big(\delta_{d_i,d'} \, (1 - \omega_{i,d'}) + \delta_{d_i,-d'} \, \omega_{i,d'} \Big) + \delta_{d_i,0} (1 - \varepsilon_{i,d'}) \\ &= \delta_{d_i,|d'|} \, \varepsilon_{i,d'} \left(\frac{1 + d' d_i (1 - 2\omega_{i,d'})}{2} \right) + \delta_{d_i,0} \, (1 - \varepsilon_{i,d'}), \end{aligned} \tag{6.11}$$

here δ is the Kronecker delta, d_i is the observed tag, while d' is the true tag, so that $\delta_{d_i,d'}$ is equal to 1 if $d_i = d'$ and 0 otherwise. The PDF $\mathcal{P}(t', \vec{d})$ can then be written as

$$\begin{aligned} \mathcal{P}(t', d) &= \sum_{d'=-1,+1} \mathcal{P}(\vec{d}\,|\,d') \mathcal{P}(t', d') \\ &= \mathcal{P}(\vec{d}\,|\,\overline{B}^0) \mathcal{P}(t', \overline{B}^0) + \mathcal{P}(\vec{d}\,|\,B^0) \mathcal{P}(t', B^0), \end{aligned} \tag{6.12}$$

where $\overline{B}^0$ (B^0) represents $d' = -1$ ($d' = +1$) respectively.

Using the definitions given in Eq. (6.13)

$$\Delta_P(\vec{d}) = \mathcal{P}(\vec{d}\,|\,\overline{B}^0) - \mathcal{P}(\vec{d}\,|\,B^0), \quad \Sigma_P(\vec{d}) = \mathcal{P}(\vec{d}\,|\,\overline{B}^0) + \mathcal{P}(\vec{d}\,|\,B^0), \tag{6.13}$$

the final PDF used in the fit to data can be written as in Eq. (6.14)

$$\mathcal{P}(t', \vec{d}) = \frac{\mathcal{N}_{t',d'}}{2} \, \mathrm{e}^{-t'/\tau} \left(\left[\Sigma_P(\vec{d}) + A_{\mathrm{P}} \, \Delta_P(\vec{d}) \right] \, H(t') + \left[\Delta_P(\vec{d}) + A_{\mathrm{P}} \, \Sigma_P(\vec{d}) \right] \, T(t') \right). \tag{6.14}$$

Decay Time Resolution The finite vertex and momentum resolution leads to a finite decay time resolution which dilutes the observed *CP* asymmetry, and can be expressed as described in [75]

$$\mathcal{D} = e^{\frac{-\Delta m_d{}^2 \, \sigma^2}{2}} \; . \tag{6.15}$$

Even though for B^0 mesons the dilution induced by the decay time resolution has a small effect on the measurement of *CP* observables, it is worthwhile to include an accurate description in the decay time fit. An unbinned maximum likelihood fit is performed to the simulated events where the difference between true and reconstructed decay time is below 0.4 ps, and the prediction of the DTF on the decay time error is used to determine the decay time resolution. These predictions have to be calibrated, this is done using a linear function with parameters b and c. To account for different sources introducing the decay time resolution an effective model consisting of two Gaussians with per-event widths is used. A third component is added to describe the effect of events matched to the wrong PV. The wrong PV component is parametrized with a broad Gaussian distribution with a common mean μ_t to the other two Gaussians and one width parameter σ_{PV}. The complete parametrization of the resolution model is given in Eq. (6.16)

$$\begin{aligned}\mathcal{R}(t - t_{\text{true}}|\sigma_t) = &\sum_{i=1}^{2} g_i \cdot \frac{1}{\sqrt{2\pi}(c_i + b_i \cdot \sigma_t)} \exp\left(-\frac{(t - t_{\text{true}} - \mu_t)^2}{2(c_i + b_i \cdot \sigma_t)^2}\right) \\ &+ f_{\text{PV}} \frac{1}{\sqrt{2\pi}\sigma_{\text{PV}}} \exp\left(-\frac{(t - t_{\text{true}} - \mu_t)^2}{2\sigma_{\text{PV}}^2}\right) \ .\end{aligned} \tag{6.16}$$

The first two Gaussian components have different calibration parameters b_i and c_i and thus different widths. Together with the fraction f_{PV} of the wrong PV component the fractions of the two Gaussian components g_1 and g_2 sum up to unity. The result of the fit is shown in Table 6.15 and the corresponding plot is presented in Fig. 6.14.

Table 6.15: Fit parameters of decay time resolution function.

Parameter		Value
μ_t	(ps)	-0.00156 $\pm$ 0.00023
b_1		1.022 $\pm$ 0.031
c_1	(ps)	0.0036 $\pm$ 0.0012
b_2		1.24 $\pm$ 0.08
c_2	(ps)	0.0127 $\pm$ 0.0035
g_2		0.23 $\pm$ 0.12
σ_{PV}	(ps)	0.16 $\pm$ 0.04
f_{PV}		0.0024 $\pm$ 0.0014

Decay Time Acceptance The trigger requirements along with some input variables used in the BDT result in a non-flat acceptance. Additionally, the VELO reconstruction inefficiency [86] causes a decay time acceptance for events with larger decay times. In order to correctly describe these effects a data-driven approach is employed. In the fit to data the B^0 lifetime is constrained to its PDG value of $\tau = (1.519 \pm 0.005)\,\text{ps}$ [28]. Any deviation from the pure exponential behavior is described using a cubic spline model. The knot positions for the splines are chosen to be at $\{0.25, 0.8, 2.0, 10.25\,\text{ps}\}$ and the second to last spline parameter is fixed to 1.0 for normalization. Exploiting the truth information, available in MC sample, the shape of the decay time acceptance can be separated from the pure exponential decay. This shape can be used to make a comparison with the spline model described above. However the spline acceptance is just an effective model and in MC, where the statistics is 25 times higher than on the data sample, two more knots are needed to describe the shape properly, the configuration $\{0.25, 0.7, 1.0, 1.5, 2.5, 10.25\,\text{ps}\}$

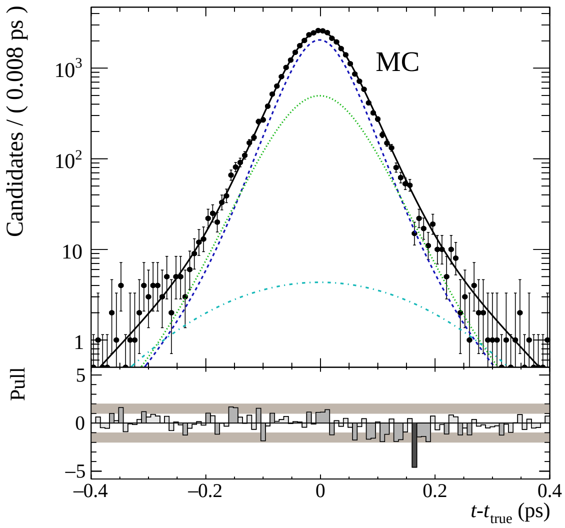


Figure 6.14: Fit of per event resolution model to the difference between true and reconstructed decay time in signal MC. The black solid line is the projection of the full PDF. The blue dashed and the green dotted line represent the two per-event components and the turquoise dashed-dotted line shows the wrong PV component.

for the knots position is used. The plots obtained from MC samples, for the two final states are shown in Fig. 6.15. As expected, due to the usage of two different BDT in the

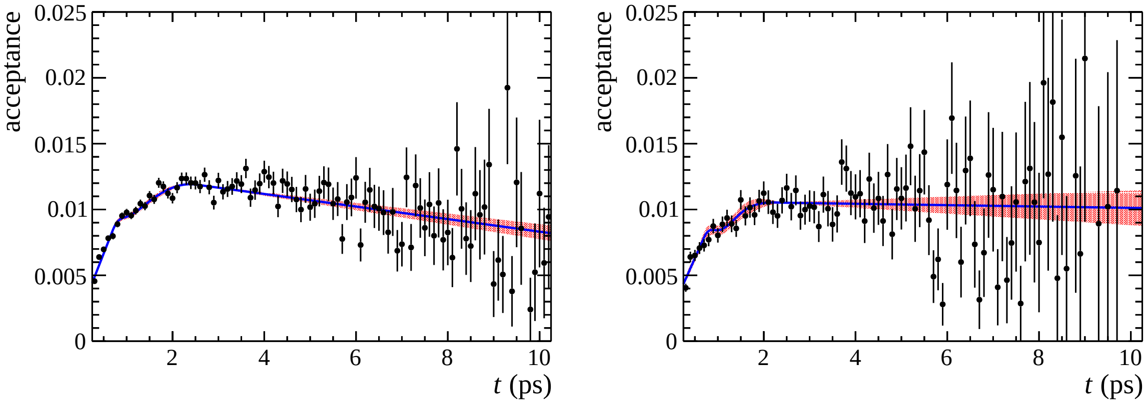


Figure 6.15: Decay time acceptance of truth-matched signal MC for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state (left) and the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state (right). The black data points show the true decay time acceptance calculated dividing the reconstructed by the true decay time distribution. The blue line is the spline acceptance function with six knots and the red stripes indicate the $1\,\sigma$ error band taking into account the statistical uncertainties.

selection, there are some differences visible between the decay time acceptances, however the low statistics of the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state on data does not allow for individual spline coefficients for these two categories, this is included in the study of systematics.

6.4 Results

Fitter Validation In order to validate the nominal fit model, two different strategies are exploited. First a fit on signal MC is performed and the fit results are compared to the generation values of $S_{D^+D^-}^{\mathrm{Gen}} = -0.7$, $C_{D^+D^-}^{\mathrm{Gen}} = 0$. In Eq. (6.17) the results are shown, where both *CP* parameters are compatible with values used in generation.

$$\begin{aligned} S_{D^+D^-}^{\mathrm{SigMC}} &= -0.697 \pm 0.032 \ , \\ C_{D^+D^-}^{\mathrm{SigMC}} &= \quad -0.035 \pm 0.031 \ . \end{aligned} \tag{6.17}$$

A second test using ToyMC is performed to check whether the likelihood fit estimation is unbiased. To this aim the full fit model, including the mass model and decay time model, is used for the generation of the samples and for the following fit. In the generation the *CP* parameters $S_{D^+D^-}$ and $C_{D^+D^-}$ are each varied in 5 steps of 0.1 from -0.6 to -0.95 and from -0.2 to 0.2, respectively, resulting in a total of 25 different configurations. All other parameter values are taken from the nominal fit including the constraints and the number of candidates to generate. For each configuration 1000 toys were run. As an example the pull distributions for the toy produced with $S_{D^+D^-} = -0.8$ and $C_{D^+D^-} = -0.1$ are shown in Fig. 6.16. From this study it is observed that the bias of the pull distributions for $S_{D^+D^-}$ is always positive, this means that the fitted value is on average larger than the generated one, however the maximal deviation is below 10 % of the statistical uncertainty.

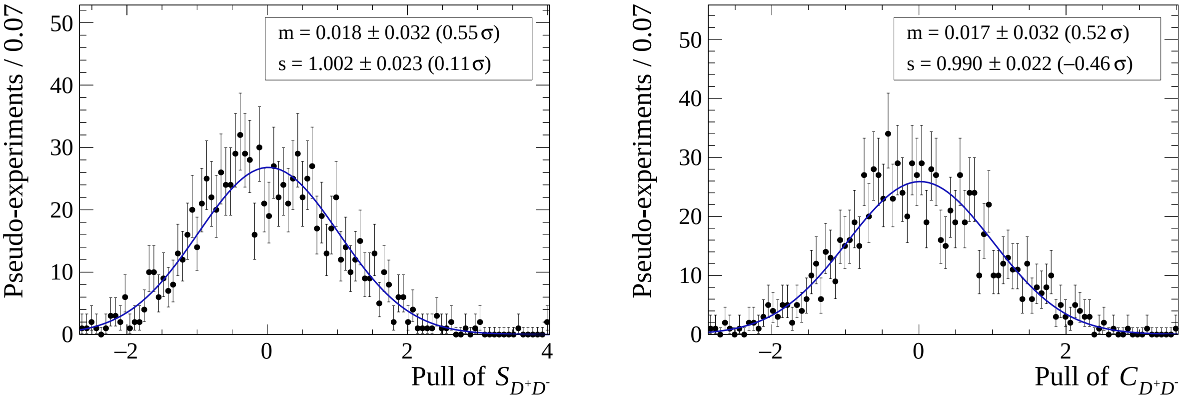


Figure 6.16: Pull distributions of $S_{D^+D^-}$ (left) and $C_{D^+D^-}$ (right) for generation values of $S_{D^+D^-} = -0.8$ and $C_{D^+D^-} = -0.1$ in the fitter validation using ToyMC.

Blinding Before performing the fit, a blinding transformation is applied on the fitted *CP* parameters $S_{D^+D^-}$ and $C_{D^+D^-}$. This transformation adds an hidden offset to the fit parameters. In this way the uncertainty on the extracted parameter does not change and can still be used for understanding the outcome of the fit. The blinded offset is drawn from a uniform distribution with a range $\pm s$, using a random number generator with known seed, called *blinding string*. In Table 6.16 are listed the blinding strings and s parameters used.

Table 6.16: Blinding strings and scales s for RooUnblindUniform.

Parameter	Blinding string	s
$S_{D^+D^-}$	SB02DD3fb	2.0
$C_{D^+D^-}$	CB02DD3fb	2.0

Input parameters In the decay time fit, some values are fixed, namely the decay time resolution parameters (see Table 6.17), the flavour tagging parameters $\langle\eta^{OS}\rangle = 0.3627$ and $\langle\eta^{SS}\rangle = 0.4282$ (which are taken from the $B^0 \to D_s^+ D^-$ calibration on Run 1 sample Sec. 5.5) and the B^0 lifetime difference $\Delta\Gamma = 0\ \text{ps}^{-1}$. On the other hand some parameters are constrained, in order to account for the uncertainties associated to these, Gaussian PDFs with mean μ and width σ are used (see Eq. (6.18))

Table 6.17: Fixed decay time resolution parameters.

Parameter		Fixed Value
b_1		1.061
c_1	(ps)	0.0038
b_2		1.38
c_2	(ps)	0.014
g_2		0.14
μ_t	(ps)	0.0
σ_{PV}	(ps)	0.160
f_{PV}		0.0029

$$\mathcal{G}(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}. \tag{6.18}$$

The variable represents the parameter to be constrained, while the Gaussian's mean and width are fixed to the parameter's value and uncertainty respectively. In Table 6.18 the list of all constrained parameters along with their uncertainties, is summarized. The value of the production asymmetry is obtained from the LHCb measurement [76], applying a

Table 6.18: Constrained parameters in the fit.

Parameter	Value and uncertainty	Source
A_{P}^{11}	-0.0047 ± 0.0106	Ref. [76]
ΔA_{P}	-0.0024 ± 0.0018	Ref. [87]
Δm ($\hbar\,\mathrm{ps}^{-1}$)	0.510 ± 0.004	Ref. [6]
τ (ps)	1.519 ± 0.005	Ref. [6]
p_0^{OS}	0.369 ± 0.008	Sec. 5.5
p_1^{OS}	1.07 ± 0.07	Sec. 5.5
Δp_0^{OS}	-0.009 ± 0.012	Sec. 5.5
Δp_1^{OS}	0.03 ± 0.11	Sec. 5.5
p_0^{SS}	0.429 ± 0.006	Sec. 5.5
p_1^{SS}	0.84 ± 0.09	Sec. 5.5
Δp_0^{SS}	-0.007 ± 0.009	Sec. 5.5
Δp_1^{SS}	0.07 ± 0.13	Sec. 5.5

weighting procedure based on transverse momentum and pseudorapidity of the nominal data set. Since the measurement has been performed on 2011 data only, the numbers for A_{P}^{11} and A_{P}^{12} are highly correlated. For this reason in the fit the constrain $A_{\mathrm{P}}^{12} = A_{\mathrm{P}}^{11} + \Delta A_{\mathrm{P}}$ with $\Delta A_{\mathrm{P}} = -0.0024 \pm 0.0018\,(\mathrm{syst})$ is used. Finally the world average value of the B^0 oscillation frequency Δm and that of the B^0 lifetime τ are used [28], while the tagging parameters p_0, p_1, Δp_0 and Δp_1 are constrained four-dimensionally taking into account their correlations.

Fit results The CP observables resulting from the decay time fit are

$$\begin{aligned} S_{D^+D^-} &= -0.54 \pm {}^{0.17}_{0.16} \,, \\ C_{D^+D^-} &= 0.26 \pm {}^{0.18}_{0.17} \,, \\ \rho(S_{D^+D^-}, C_{D^+D^-}) &= 0.48 \,. \end{aligned}$$

In Fig. 6.17 the fit to the decay time distribution of the $B^0 \to D^+D^-$ data sample is presented. The correlation matrix is shown in Fig. 6.18, while the parameters of the acceptance splines, listed in Table 6.19, show a quite high positive correlation among themselves. $S_{D^+D^-}$ and $C_{D^+D^-}$ have a correlation of about 0.5, but apart from these only quite small correlations are present, also between $S_{D^+D^-}$ and Flavour Tagging calibration parameters. Moreover since the constraint applied on Δm is a lot tighter than the sensitivity available from the data sample, almost no correlation with $C_{D^+D^-}$ is observed. In Fig. 6.19 the time-dependent signal asymmetry is shown, while in Fig. 6.20, the two dimensional likelihood profile scan for $S_{D^+D^-}$ and $C_{D^+D^-}$ is presented.

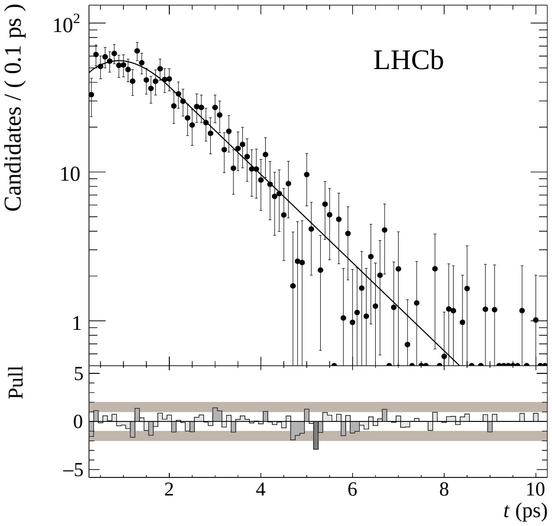
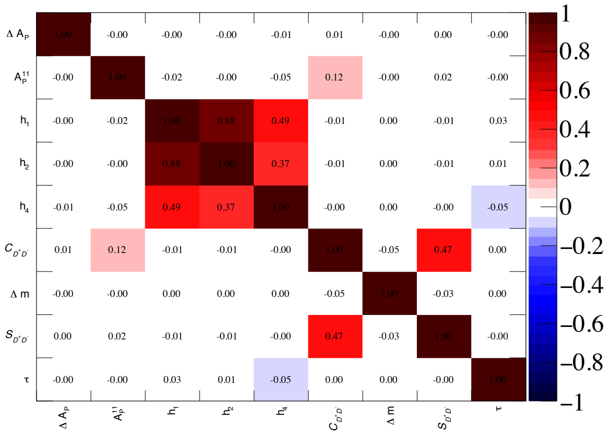


Figure 6.17: Plot of the decay time distribution of the $_s weighted$ $B^0 \to D^+D^-$ signal data sample with the projection of the PDF and the pull distribution. The y-axis is plotted in logarithmic scale.



	ΔA_P	A_P^{11}	h_1	h_2	h_4	$C_{D^+D^-}$	Δm	$S_{D^+D^-}$	τ
ΔA_P	1.00	-0.00	-0.00	-0.00	-0.01	0.01	-0.00	0.00	-0.00
A_P^{11}	-0.00	1.00	-0.02	-0.00	-0.05	0.12	-0.00	0.02	-0.00
h_1	-0.00	-0.02	1.00	0.88	0.49	-0.01	0.00	-0.01	0.03
h_2	-0.00	-0.00	0.88	1.00	0.37	-0.01	0.00	-0.01	0.01
h_4	-0.01	-0.05	0.49	0.37	1.00	-0.00	0.00	-0.00	-0.05
$C_{D^+D^-}$	0.01	0.12	-0.01	-0.01	-0.00	1.00	-0.05	0.47	0.00
Δm	-0.00	-0.00	0.00	0.00	0.00	-0.05	1.00	-0.03	0.00
$S_{D^+D^-}$	0.00	0.02	-0.01	-0.01	-0.00	0.47	-0.03	1.00	-0.00
τ	-0.00	-0.00	0.03	0.01	-0.05	0.00	0.00	-0.00	1.00

Figure 6.18: Correlation matrix of the fit parameters in the decay time fit to data. Positive correlations are represented by the red palette on the z axis, while negative correlations are presented by the blue palette.

6.5 Systematic uncertainties

The possible sources of systematic uncertainties are described in the following. At first, in order to check for possible systematic effects, fits are performed dividing the data set in different subsamples according to the tagging algorithms, the years of data-taking, the magnet polarities, the D final states and the BDT classifier range for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$

Table 6.19: Acceptance spline parameters from the decay time fit to $B^0 \to D^+D^-$ data.

Parameter	Value
h_1	0.67± 0.06
h_2	0.87± 0.11
h_4	0.992± 0.010

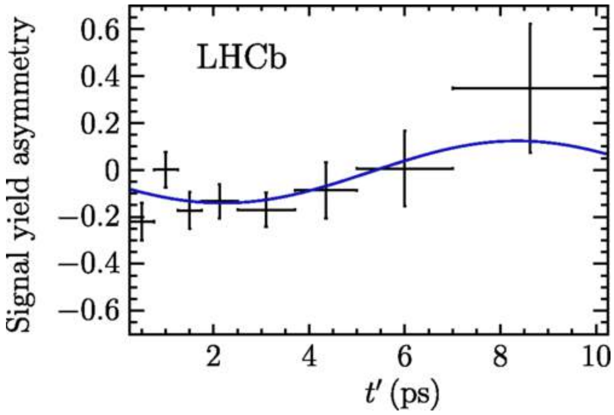


Figure 6.19: Time-dependent signal yield asymmetry $(N_{\overline{B}^0} - N_{B^0})/(N_{\overline{B}^0} + N_{B^0})$. Here $N_{B^0}(N_{\overline{B}^0})$ is the number of $B^0 \to D^+D^-$ decays with a B^0 ($\overline{B}^0$) flavour tag. The data points are obtained with the sPlot technique, assigning signal weights to the events based on a fit to the reconstructed mass distribution. The solid curve is the projection of the signal PDF.

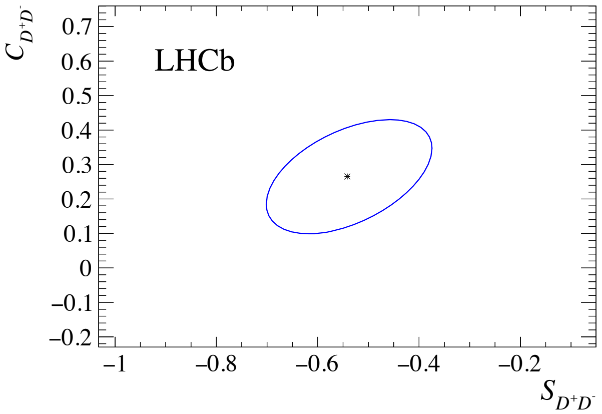


Figure 6.20: Two dimensional likelihood profile scan for $S_{D^+D^-}$ and $C_{D^+D^-}$. The contour line shows the 1σ confidence level.

final state. In Fig. 6.21 the results for the possible configurations are illustrated. Almost all subsamples show compatible results, the only difference can be observed between the 2011 and the 2012 subsample for $S_{D^+D^-}$, which can be interpreted as a a statistical fluctuation. The results of the difference between the weighted average of the subsamples and the nominal fit result are listed in Table 6.20.

Table 6.20: Difference between weighted average of subsamples and nominal result for the splits defined above.

Split	$S_{D^+D^-}$	$C_{D^+D^-}$
Tagging algorithm	−0.01	−0.01
Year of data-taking	−0.02	−0.02
Tagging + year	0.01	0.02
Magnet polarity	−0.02	0.01
BDT	−0.01	0.02
Final state	0.01	−0.02

Moreover dedicated studies were performed to estimate the systematic uncertainties related to the mass model, the correlation between decay time and the SS mistag, the decay time resolution and acceptance models used in the fit, the z-scale, the production asymmetry, the decay width and mass difference, as described in the following.

Decay Time Fit Bias To estimate the bias due to the likelihood fit, 10000 pseudo-experiments are performed. For the generation the nominal fit results of CP observables are used. In Fig. 6.22 the pull distributions are presented, here a very small deviation of the mean value from zero is visible. The resulting systematic uncertainties, obtained by multiplying the statistical uncertainty with the bias are:

$$S_{D^+D^-} = 0.004, \qquad C_{D^+D^-} = 0.0025. \tag{6.19}$$

This correction on the decay time fit bias is included in all the following studies on systematic uncertainties, in order to correctly interpret all these effects.

Mass Model In order to estimate the systematic uncertainties associated to the mass model two possible sources are taken into account. At first the effect of mismodelling the mass components is studied. In $B^0 \to D^+D^-$ analysis the BDT training is performed using a MC sample where the PID variables are not corrected in order to obtain an optimal data/MC agreement. Since some of the shape parameters used in the mass fit are extracted from MC samples, these could be distorted due to data/MC differences. For this reason different parametrizations of the mass components are tested in order to extract *sWeights* which are then used in the decay time fit. The resulting CP observables are then compared with the ones extracted from the nominal fit. In the first configuration the $B^0 \to D^+D^-$ signal and the $B_s^0 \to D^+D^-$ are modeled using a single Gaussian, in the second configuration a Chebychev polynomial of first kind of second order is used to describe the combinatorial background while in the third configuration the $B^0 \to D_s^+D^-$

tail parameters are once extracted from a MC sample where the BDT is not applied and once extracted from a MC sample where the BDT is applied and a tighter cut with respect to the nominal one is chosen.

The largest deviations obtained for $S_{D^+D^-}$ and $C_{D^+D^-}$ are

$$S^{mass}_{D^+D^-} = 0.004, \qquad C^{mass}_{D^+D^-} = 0.006. \tag{6.20}$$

The second source of systematic considered is due to the missing component for the $B^0_s \to D^{*+}D^-$ in the nominal mass fit. In principle both the $B^0 \to D^{*\pm}D^{\mp}$ and $B^0_s \to D^{*+}D^-$ contributions should be taken into account, but since the mass range considered starts at 5150 MeV/c^2, only the second one enters the fit region. However since the expected number of $B^0_s \to D^{*+}D^-$ is low this contribution is neglected in the nominal mass fit. Another contribution which is neglected in the nominal mass fit is the partially charmless background, where one of the hadron triplets is not originating from a D decay. In this case the systematic uncertainty is estimated by making use of 1000 pseudo-experiment, in which the two mass components ($B^0_s \to D^{*+}D^-$ and partially charmless background) are included in the generation but excluded from the fit procedure. Two single Gaussian functions centered around 5150 MeV/c^2 and 5200 MeV/c^2 are used to parametrize the $B^0_s \to D^{*+}D^-$ component, while for the partially charmless background a single Gaussian distribution with the mean set to the same position as the signal B^0 and a width of 10 MeV/c^2 is chosen. The $B^0_s \to D^{*+}D^-$ is generated without any tagging asymmetry, for the partially charmless background the worst case scenario of maximal CP violation with opposite CP eigenvalue is used. The resulting pull distributions are shown in Fig. 6.23, the biases on the mean parameter of $S_{D^+D^-}$ and $C_{D^+D^-}$ translate into the following systematic uncertainties

$$S^{mass2}_{D^+D^-} = 0.05, \qquad C^{mass2}_{D^+D^-} = 0.013. \tag{6.21}$$

Correlation between decay time and SS mistag In the nominal fit no correlation is observed between OS mistag and decay time (see Sec. 5.6.1), while the linear correlation observed on signal MC between the SS mistag and the decay time, in the nominal fit is neglected. In order to take into account the effect of this neglect, 1000 pseudo-experiments are generated, using a gaussian distribution for the SS mistag with mean drawn from the linear function written in Eq. (6.22)

$$\eta_{SS} = a_{\eta SS,t}t + b_{\eta SS,t}t. \tag{6.22}$$

In the fitting the correlation is again neglected, the outcoming pull distributions (see Fig. 6.24) show no significant bias. The resulting systematic uncertainties are

$$S_{D^+D^-} = 0.0007, \qquad C_{D^+D^-} = 0.007. \tag{6.23}$$

Decay Time Resolution Model To evaluate the impact of an inaccurate determination of the resolution parameters, 1000 pseudo-experiments are produced. In the generation the parameters obtained from the simulation sample and listed in Table 6.15 are enlarged by a 15% and the mean parameter of the Gaussians is fixed to zero, while in the fitting the resolution parameters are fixed to their nominal values and the mean is set to zero. The pull distributions are shown in Fig. 6.25, while the resulting uncertainties, obtained as product of the biases on the mean parameter and the statistical uncertainty are

$$S_{D^+D^-} = 0.0020, \qquad C_{D^+D^-} = 0.0023. \tag{6.24}$$

Decay Time Acceptance Model Due to the low statistic of the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state, in the nominal fit it is not possible to use an individual spline model for each D final state. In order to quantify the effect of neglecting this difference 1000 pseudo-experiments are performed, where the decay time distribution is generated using the histograms (one for each D final state) of the true decay time acceptance from signal MC and fitted using a single spline acceptance as done in the nominal fit. To cover uncertainties arising from the choice of the number and position of the knots, two histograms with 100 bins each are used. The resulting pull distributions are shown in Fig. 6.26 and the uncertainties, determined as the product of the shift in the pull distribution and the statistical uncertainty are

$$S_{D^+D^-} = 0.007, \qquad C_{D^+D^-} = 0.0027. \tag{6.25}$$

z-scale Since the decay time is measured from the distance between PV and decay vertex, any uncertainty on the position of detector elements can lead to biased decay times, in particular as the main contribution to the flight distance are the z coordinates, the scale uncertainty in z direction has to be studied. The impact of this uncertainty on CP observables, is investigated using 1000 pseudo-experiments, for each a new value for the uncertainty on the z-scale is drawn from a Gaussian distribution with width $\sigma_{z-scale}$ and mean set to zero. The product of this value with the decay time is then added to the width parameter of the Gaussian function modelling the decay time resolution. For the fit a nominal width corresponding to the effective resolution estimated on signal MC and corresponding to 50 fs is used. The pull distributions of the pseudo-experiments are shown in Fig. 6.27, while the nominal statistical uncertainty is taken as systematic uncertainty

$$S_{D^+D^-} = 0.0031, \qquad C_{D^+D^-} = 0.0028. \tag{6.26}$$

Production Asymmetry The systematic uncertainty on the production asymmetry is estimated using 1000 pseudo-experiments. In the nominal fit this parameter is constrained, so for the toy generation the nominal value is used. For the fitting procedure the mean parameters of the Gaussian constraints are varied. At first the mean is shifted by one systematic uncertainty, then the resulting Gaussian distribution is used to draw a new value for the mean. Finally the new Gaussian distribution is used to constrain the different parameters in the fit. The largest deviation is taken as systematic uncertainty, after testing both possible variations, namely up and down, for the mean:

$$S_{D^+D^-} = 0.0015, \qquad C_{D^+D^-} = 0.004. \tag{6.27}$$

The resulting pull distributions are presented in Fig. 6.28. The systematic uncertainty for the production asymmetry difference ΔA_P is already included in the Gaussian constraint of the nominal fit.

Decay Width Difference In the nominal fit the decay width difference $\Delta\Gamma_d$ is fixed to zero, however experimentally the uncertainty associated to this parameter is relatively large. To study the effect of this choice for the nominal fit, 1000 pseudo-experiments are used, where for the generation the actual statistical precision is exploited $\sigma(\Gamma_d) = \pm 0.007$ ps^{-1} [28], while for the fit the nominal set is chosen. The pull distributions are shown in Fig. 6.29, while the resulting systematic uncertainties are:

$$S_{D^+D^-} = 0.014, \qquad C_{D^+D^-} = 0.0021. \tag{6.28}$$

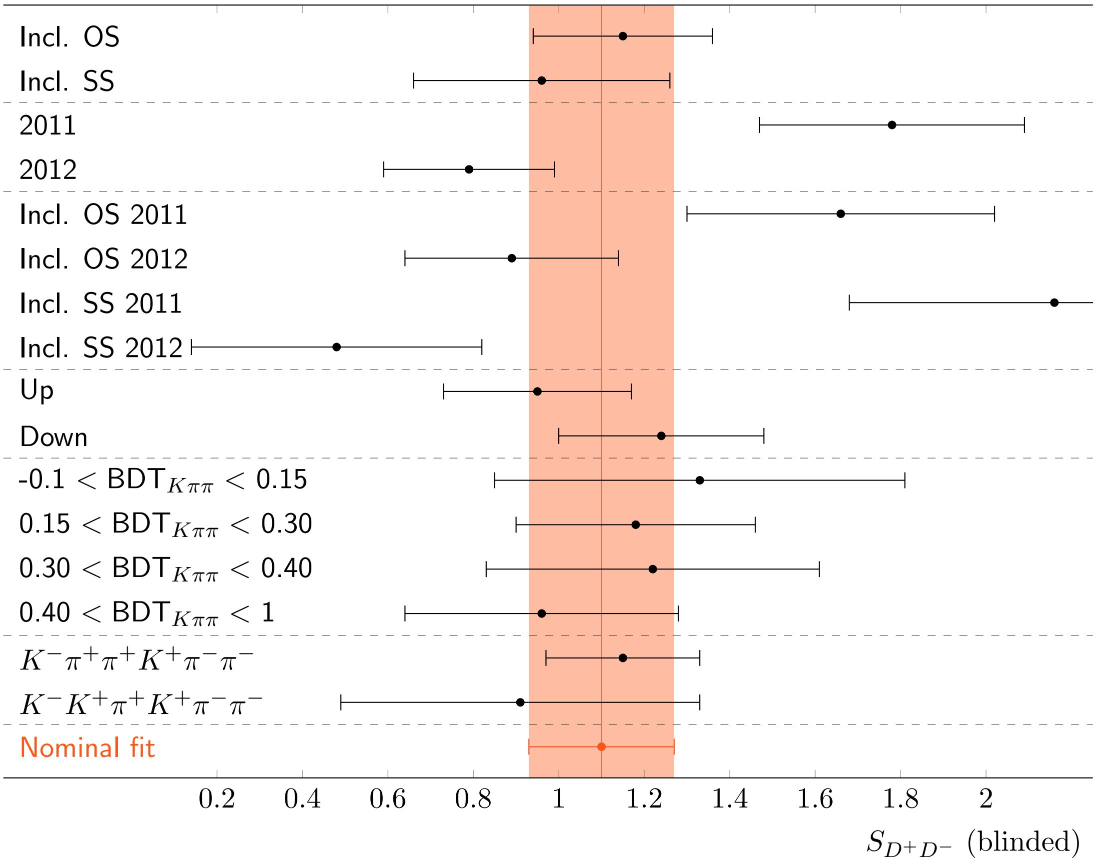
Δm_d In the nominal fit the mass difference Δm_d is Gaussian constrained but the effect of systematic uncertainty is not included. For this reason a set of 1000 pseudo-experiments is exploited, in which for the generation the nominal value is chosen. In the fitting, at first the mean of the Gaussian distribution, that has as width the statistical precision of the world average, is shifted by one systematic uncertainty (once up and once down), then a new value is drawn from the distribution, and finally this new constraint is used for the fit. In Fig. 6.30 the resulting pull distributions are presented, the assigned systematic uncertainties are:

$$S_{D^+D^-} = 0.0025, \qquad C_{D^+D^-} = 0.006. \tag{6.29}$$

The full systematic uncertainty is calculated by summing the single uncertainties in quadrature, the result is summarized in Table 6.21.

Table 6.21: Systematic uncertainties on the CP observable $S_{D^+D^-}$ and $C_{D^+D^-}$.

Origin	$\sigma_{S_{D^+D^-}}$	$\sigma_{C_{D^+D^-}}$
Neglecting components in mass model	0.05	0.013
$\Delta\Gamma_d$	0.014	0.0021
Decay time acceptance	0.007	0.0027
Correlation between mass and decay time	0.0007	0.007
Parametrization of PDFs in mass model	0.004	0.006
Δm_d	0.0025	0.006
Fit bias	0.004	0.0025
z-scale	0.0031	0.0028
Decay time resolution	0.0020	0.0023
Production asymmetry	0.0015	0.004
Sum	0.05	0.018



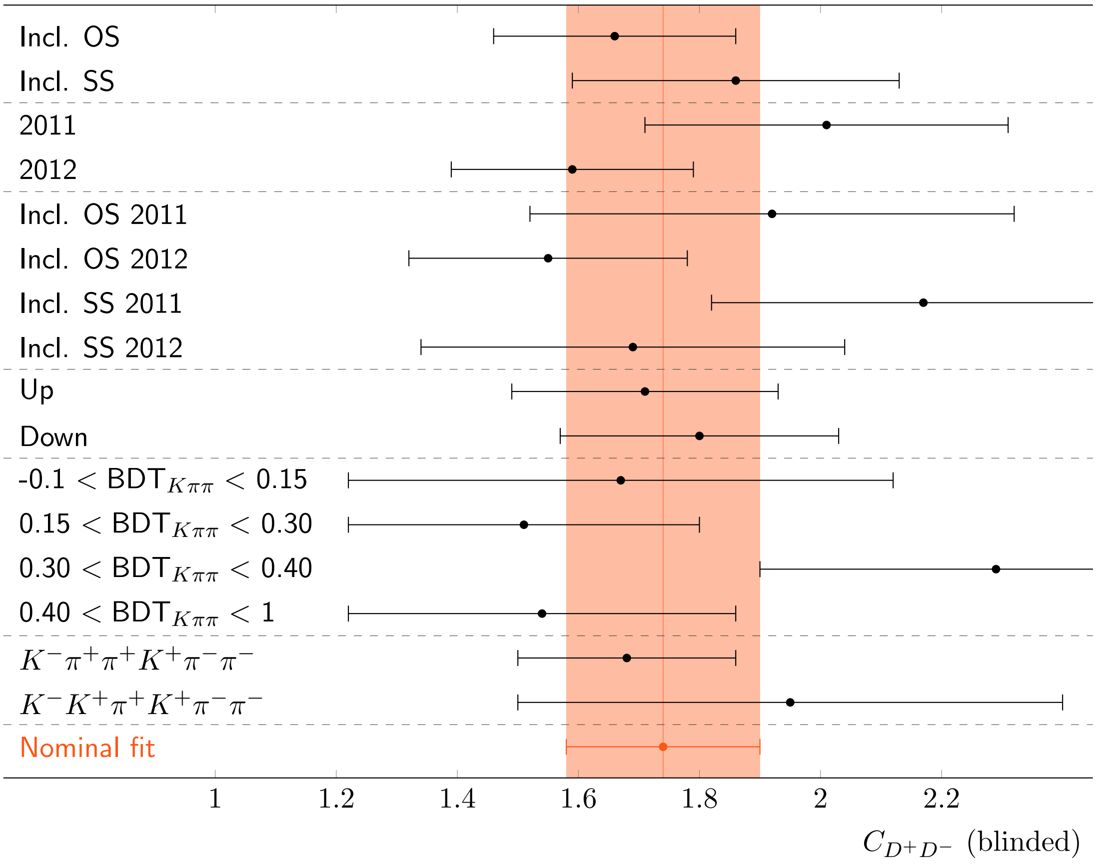


Figure 6.21: Comparison of fit results of $S_{D^+D^-}$ and $C_{D^+D^-}$ for fits on various subsamples.

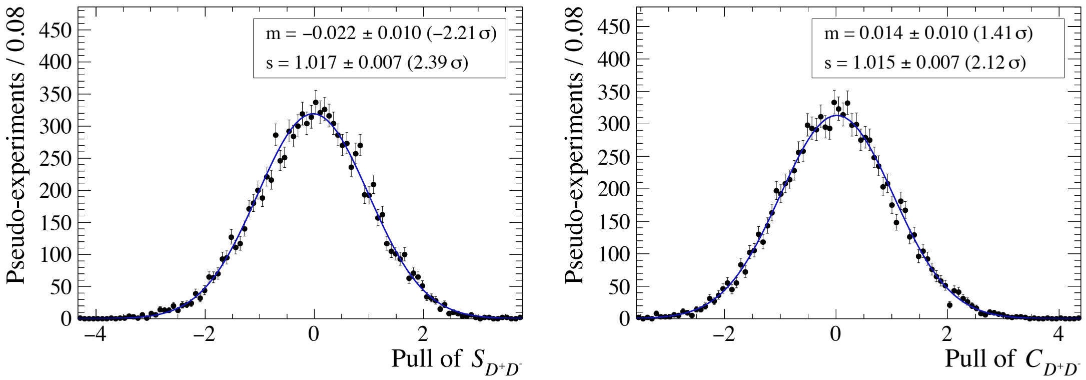


Figure 6.22: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainty due to the likelihood fitter.

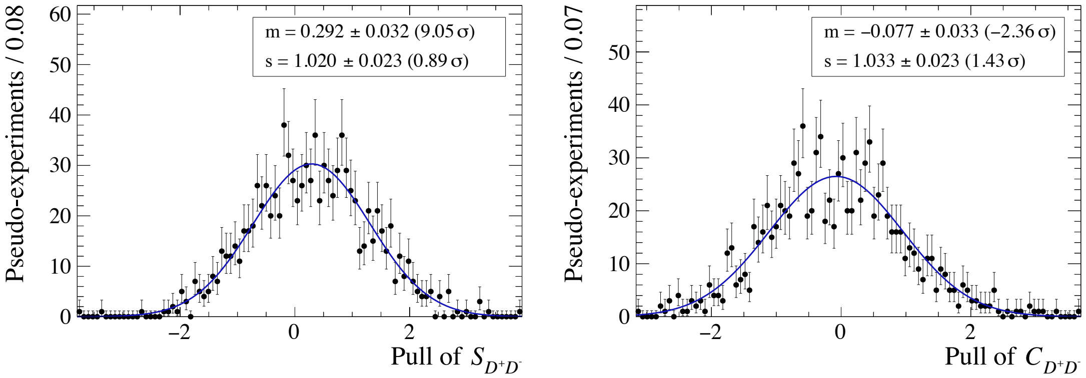


Figure 6.23: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainty due to neglecting $B_s^0 \to D^{*+}D^-$ and partially charmless contributions.

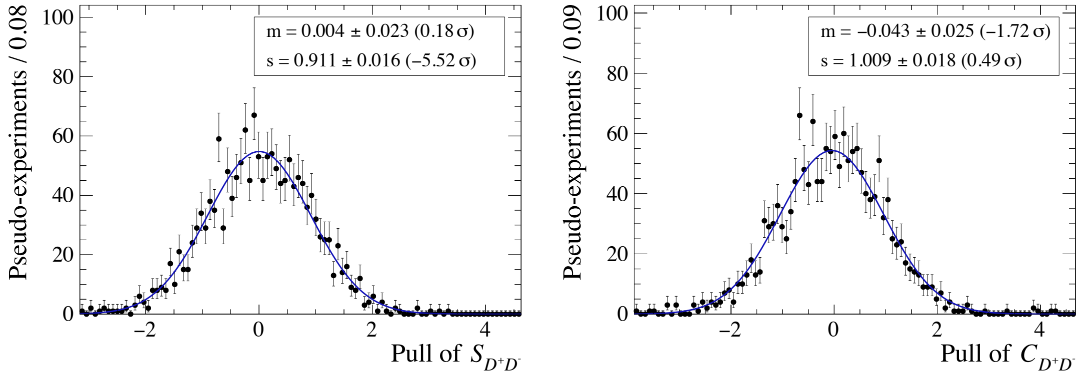


Figure 6.24: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainties due to neglecting the correlation between the SS mistag and the decay time.

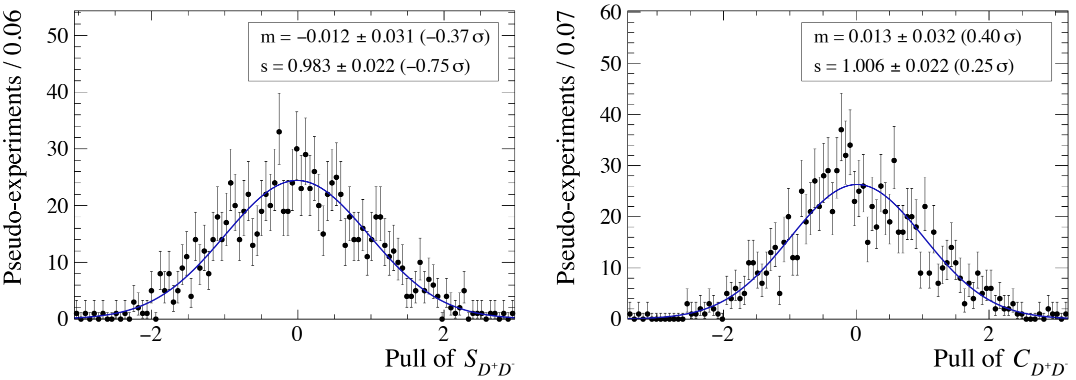


Figure 6.25: Pull distribution of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainties coming from imperfect modelling of the decay time resolution.

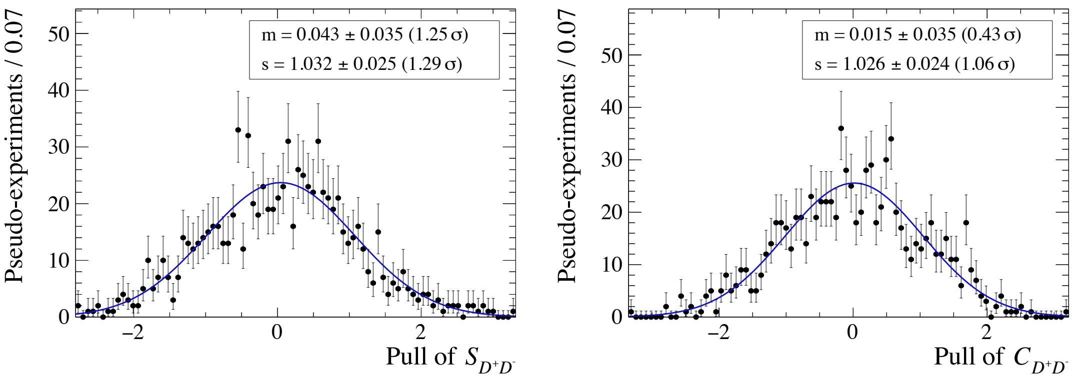


Figure 6.26: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainties due to the decay time acceptance model.

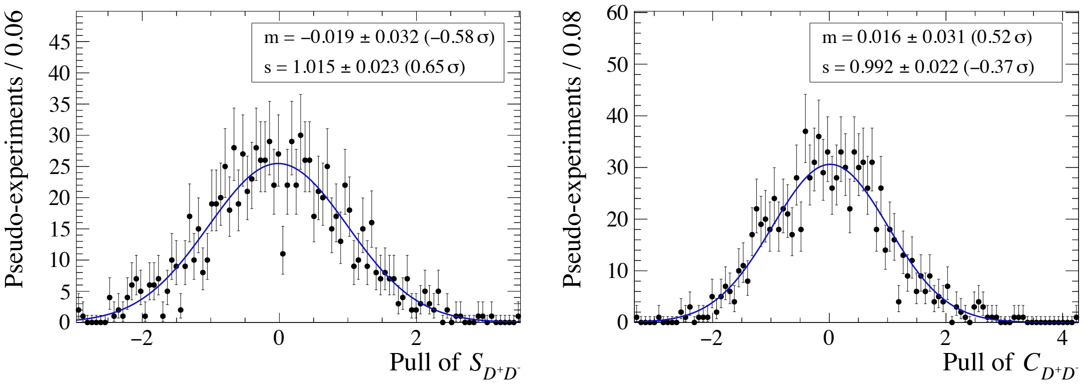


Figure 6.27: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainties due to the relative uncertainty of the z-scale.

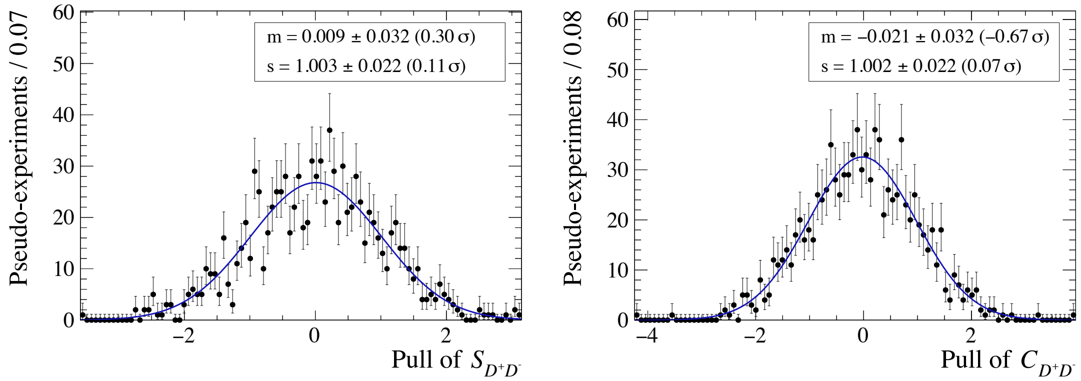


Figure 6.28: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainties associated to the production asymmetry.

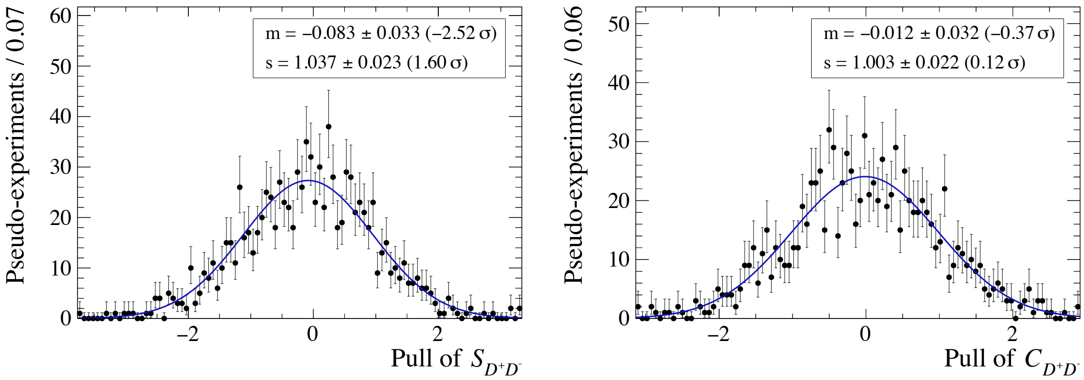


Figure 6.29: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainties due to a possible variation of $\Delta\Gamma_d$ from zero.

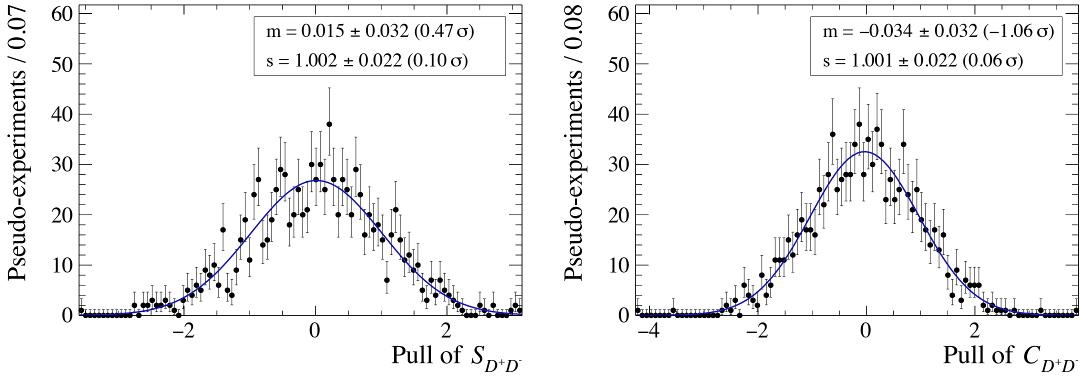


Figure 6.30: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainties associated to Δm_d.

Chapter 7

Measurement of CP violation in $B^0 \to D^{*+} D^-$

In this chapter the measurement of CP violation in $B^0 \to D^{*\pm} D^{\mp}$ decay will be described. The measurement is based on data samples of pp collisions corresponding to a total integrated luminosity of 5 fb^{-1}, recorded by the LHCb experiment during Run 1 and Run 2. As described in Sec. 3.5, CP violation in this decay arises from the interference between the direct decay and the decay after B^0-$\overline{B}^0$ mixing, and the measurement of the decay-time dependent decay rate (Eq. (3.61)) , gives access to the CP related observables S_{D^*D}, ΔS_{D^*D}, C_{D^*D} and ΔC_{D^*D}.

7.1 Selection

In this section the selection steps used in $B^0 \to D^{*\pm} D^{\mp}$ analysis will be described. For the reconstruction of the decay $B^0 \to D^{*\pm} D^{\mp}$, between the possible final states (see Table 7.1), the analysis is restricted, for the $D^{\pm}$ mesons, to the decay

Table 7.1: Branching fractions of possible D^0 and $D^{\pm}$ final states.

Mode		$\mathcal{B}$ $[10^{-3}]$
D^0	$\to K^-\pi^-\pi^+\pi^+$	80.6 ± 2.3
D^0	$\to K^-\pi^+$	39.3 ± 0.4
D^0	$\to K^-K^+$	4.0 ± 0.1
D^0	$\to \pi^-\pi^+$	1.42 ± 0.03
D^+	$\to K^-\pi^+\pi^+$	94.6 ± 2.4
D^+	$\to K^+K^-\pi^+$	9.9 ± 0.3

- $D^{\mp} \to K^{\pm}\pi^{\mp}\pi^{\mp}$

while the $D^{*\pm}$ mesons are reconstructed as $D^{*\pm} \to D^0\pi^\pm$ where for the D^0 the two following decays are considered

- $D^0 \to K^\pm\pi^\pm\pi^\mp\pi^\mp$
- $D^0 \to K^\pm\pi^\mp$

The branching fractions reported in Table 7.1 don't enter any calculation in the analysis, for this reason their uncertainties will not be propagated as systematics.

The selection chain starts with a two-layer trigger system which selects events from the proton-proton collisions, followed by a first very general selection (stripping), and the writing of data to tape. The selection then proceeds with an offline cut-based preselection, the application of a multivariate classifier to suppress combinatorial background, the use of a set of vetoes on misidentified backgrounds, a random suppression of the multiple candidates and finally a mass fit which is employed to separate signal events from the remaining background ones. Very similar selection criteria are applied to the two D^0 final states, in order to reduce possible differences in the decay time fit (see Sec. 7.3).

7.1.1 Stripping

Starting from the process of tuples creation a set of conditions is applied.

At first all tracks are required to be long tracks and an event is only considered if it contains in total less than 500 long tracks. All the requirements applied on the daughter particles (kaons and pions) of the D mesons are listed in Table 7.2.

Table 7.2: Stripping cuts applied in the reconstruction and selection of the D^0 and $D^\pm$ meson daughters.

track χ^2/ndf	< 2.5
p_T	$> 500\,\mathrm{MeV/c}$
p	$> 5000\,\mathrm{MeV/c}$
IP χ^2	> 4
ghost probability	< 0.4

These hadron tracks have to form a common vertex. The criteria that $D^{*\pm}$ and $D^\pm$ combination have to fulfill are given in Table 7.3.

Finally the $D^{*\pm}$ and $D^\pm$ mesons have to form a common vertex, the requisite on the reconstructed B^0 are listed in Table 7.4.

7.1.2 Preselection

During this stage of tuple processing many of the cuts used in the stripping are reimplemented (see Table 7.5) using the DecayTreeFitter (DTF) [81].

Table 7.3: Stripping cuts on the $D^{*\pm}$ and $D^{\pm}$ combination.

$m_{D^{\pm}}$	$1769.62\,\mathrm{MeV/c^2} - 2068.49\,\mathrm{MeV/c^2}$
$m_{D^{*\pm}}$	$1410.26\,\mathrm{MeV/c^2} - 2610.26\,\mathrm{MeV/c^2}$
m_{D^0}	$1664.83\,\mathrm{MeV/c^2} - 2064.83\,\mathrm{MeV/c^2}$
$\sum p_{\mathrm{T}}$ (hhh)	$> 1800\,\mathrm{MeV/c}$
DOCA	$< 0.5\,\mathrm{mm}$
$D^{\pm}$ $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$	< 10
$D^{*\pm}$ $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$	< 10
vertex distance χ^2 to any PV	> 36
DIRA	> 0

Table 7.4: Stripping cuts on B^0 combination.

B^0 $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$	< 10
m_{B^0}	$4750\,\mathrm{MeV/c^2} - 6000\,\mathrm{MeV/c^2}$
$p_{\mathrm{T}}(D^+) + p_{\mathrm{T}}(D^-)$	$> 5\,\mathrm{GeV/c}$
t	$> 0.2\,\mathrm{ps}$
IP χ^2	< 25
DIRA	> 0.999
p	$> 10\,\mathrm{GeV/c}$

Table 7.5: Cuts reimplemented using the DTF.

$p_{\mathrm{T}}(K, \pi)$	$> 100\,\mathrm{MeV/c}$
$p(K, \pi)$	$> 1000\,\mathrm{MeV/c}$
IP $\chi^2(K, \pi)$	> 4
$\sum p_{\mathrm{T}}$ (D, D^0 daughters)	$> 1800\,\mathrm{MeV/c}$
t	$> 0.2\,\mathrm{ps}$
$p_{\mathrm{T}}(D^{*+}) + p_{\mathrm{T}}(D^-)$	$> 5\,\mathrm{GeV/c}$
$p(B^0)$	$> 10\,\mathrm{GeV/c}$
IP $\chi^2(B^0)$	< 25
m_{B^0} (D^{*+}, D^0, D^- mass + PV constraints)	$4750\,\mathrm{MeV/c^2} - 6000\,\mathrm{MeV/c^2}$

Moreover a set of requirements on the D meson masses are applied, as reported in Table 7.6.

In Fig. 7.1 the invariant $D^{*+}D^-$ mass distributions after applying the preselection

Table 7.6: Offline cuts applied in the preselection.

D^*, D^0 masses	$\mid m_{D^*} - m_{D^0} \mid < 155.$ MeV/c^2
D^0 mass	$\mid m_{D^0} - 1864.83 \mid < 40.$ MeV/c^2
$D^\pm$ mass	$\mid m_{D^\pm} - 1869.58 \mid < 50.$ MeV/c^2

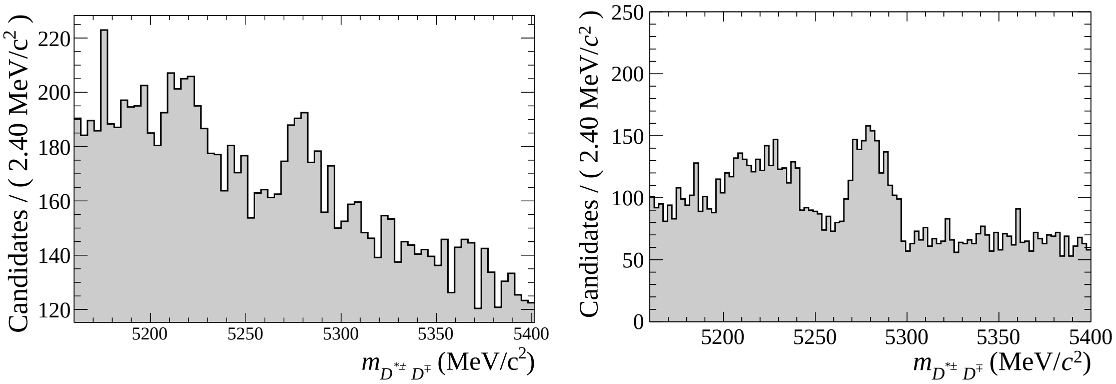


Figure 7.1: Plot of the invariant $D^{*+}D^-$ mass of the candidates which survive the preselection cuts for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ (left) and to $K^-\pi^+$ (right) in Run 1 data sample.

cuts are shown for the two D^0 final states, where the signal peak and the D^*Ds peak are already clearly visible.

7.1.3 Vetoes against misidentified $D^\pm$ final states

Misidentification of kaons and protons as pions for the final state particles of $D^\pm$ meson leads to backgrounds, which are suppressed by applying explicit vetos.

$\Lambda_c^\pm$ To reduce $p \to \pi$ mis-ID the proton mass hypothesis is assigned to the pion with the higher p_T of the $D^\pm \to K^\mp\pi^\pm\pi^\pm$ final state and the invariant mass of the $Kp\pi$ combination is calculated. The candidate is rejected if the ratio of the pion probability over the sum of the pion and proton probabilities $\texttt{ProbNN}\pi/(\texttt{ProbNNp}+\texttt{ProbNN}\pi)$ is lower than 0.7 and the invariant mass of the $Kp\pi$ combination is compatible with the $\Lambda_c^\pm$ mass of $M_{\Lambda_c^\pm} = 2286.46\,\mathrm{MeV/c^2}$ [28], within $\pm 25\,\mathrm{MeV/c^2}$. As can be seen from Fig. 7.2 this veto reduces the excess of candidates around the $\Lambda_c^\pm$ mass.

ϕ The $K \to \pi$ mis-ID can lead to background contributions from $\phi \to KK$. To reduce it the kaon mass hypothesis is assigned to the pion with the higher transverse momentum. The invariant mass of the kaon pair is then calculated. The candidate is rejected if the invariant

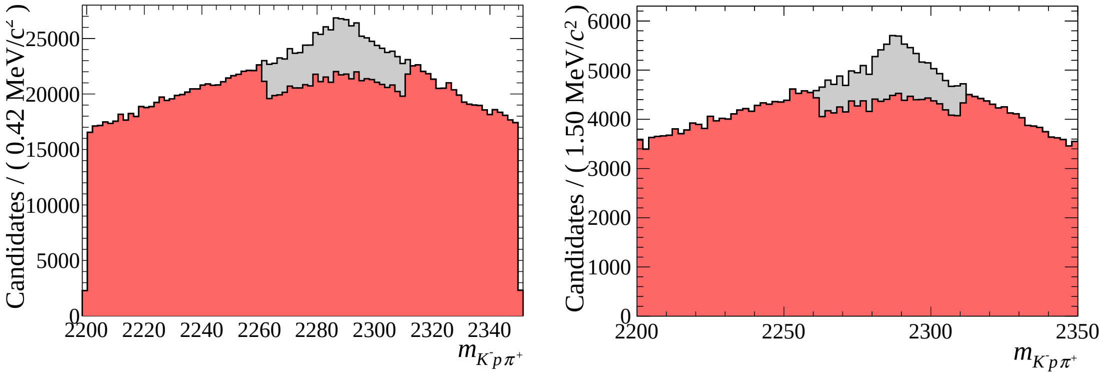


Figure 7.2: Plots of the invariant mass of the $Kp\pi$ combination. The distributions are given without (grey) and with (red) the $\Lambda_c^\pm$ veto described in the text applied, for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ on the left and $K^-\pi^+$ on the right in Run 1 data sample.

mass of the kaon pair is compatible with the ϕ mass of $M_\phi = 1019.461\,\mathrm{MeV/c^2}$ [28], within $\pm 10\,\mathrm{MeV/c^2}$. The plots showing the distributions before and after applying this veto, for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ and $K^-\pi^+$ final states, are given in Fig. 7.3.

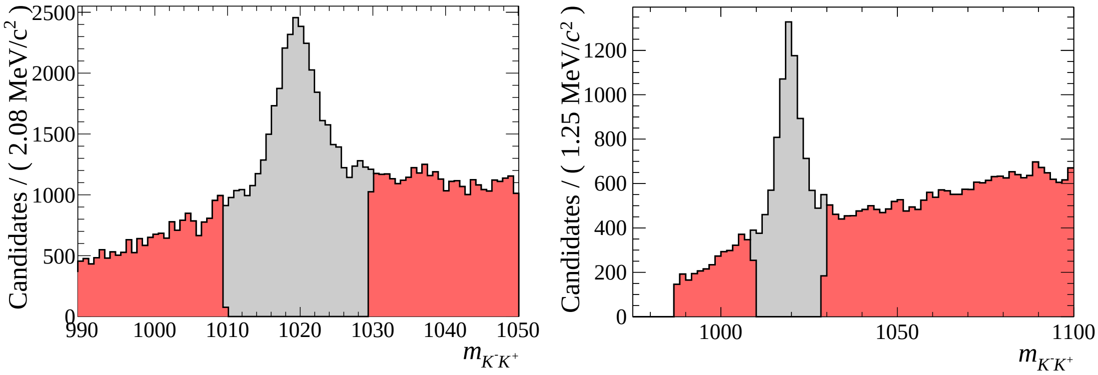


Figure 7.3: Plots of the invariant mass of the KK combination. The distributions are given without (grey) and with (red) the ϕ veto described in the text, for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ on the left and $K^-\pi^+$ on the right for Run 1 data sample.

$D_s^\pm$ For $D^\pm$ mesons which are reconstructed in the $D^\pm \to K^\mp\pi^\pm\pi^\pm$ final state, a $K \to \pi$ mis-ID can lead to background contributions from $D_s^\pm \to K^\mp K^\pm \pi^\pm$. To reduce this background, the kaon mass hypothesis is assigned to the pion with the higher transverse momentum. The invariant mass of the three hadrons is then recalculated. The candidate is rejected if the kaon probability is higher than the pion probability and the invariant mass of the $KK\pi$ combination is compatible with the mass of $M_{D_s^\pm} = 1968.30\,\mathrm{MeV/c^2}$ [28], within ±25 MeV/c^2. The plots showing the distributions before and after applying the $D_s^\pm$ veto are given in Fig. 7.4.

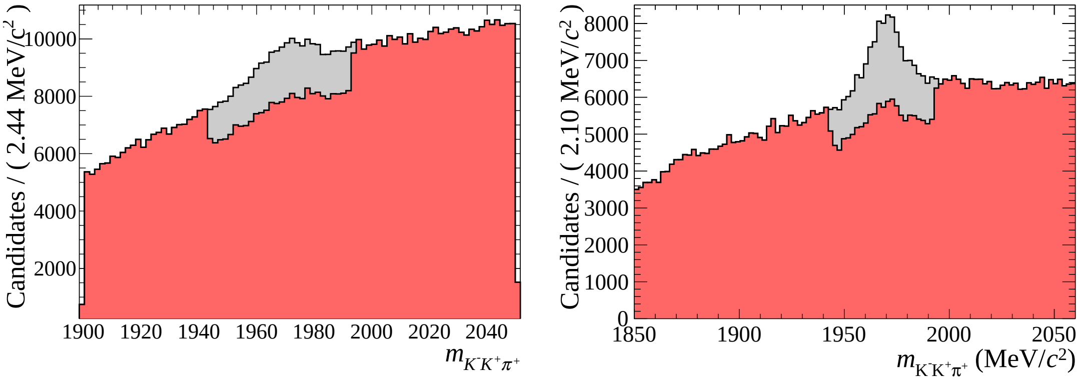


Figure 7.4: Plots of the invariant mass of the $KK\pi$ combination. The distribution is given without (grey) and with (red) the $D_s^\pm$ veto described in the text applied, for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ on the left and $K^-\pi^+$ on the right for Run 1 data sample.

7.1.4 Vetoes against single charm background

To understand the effect of possible contamination from decays as $B^0_{(s)} \to Dhhh$ whose branching ratios can be comparable or even higher (e.g. $\mathcal{B}(B^0 \to D^{*+}\pi^+\pi^-\pi^-) = (7.21\pm0.29)\cdot10^{-3}$ [28]) with respect to the signal one ($\mathcal{B}(B^0 \to D^{*+}D^-) = (6,1\pm0.6)\cdot10^{-4}$ [28]), a study exploiting an enhanced background sample is performed. This sample corresponds to events for which the $D^\pm$ decay time significance w.r.t. the B^0 decay vertex is lower than zero or the χ^2 of the flight distance of the $D^\pm$ w.r.t. the B^0 decay vertex is lower than two. The DTF (DecayTreeFitter) mass variable with constraints on PV, $D^{*\pm}$ mass and D^0 mass is compared with the nominal one, i.e. the one that is used in the mass fit, with constraints on PV, $D^{*\pm}$ mass, D^0 mass and $D^\pm$ mass, in order to check for excesses under the signal peak or in other part of the analyzed mass region.

$B_s^0 \to D^{*+}K^+\pi^-\pi^-$ The first source of single charm background observed is the one coming from $B_s^0 \to D^{*+}K^+\pi^-\pi^-$. The comparison between the nominal mass and the one obtained removing the $D^\pm$ constraint is shown in Fig. 7.5. A clear excess under the B_s^0 peak position is visible, showing the presence of the above described background. In order to suppress this background, candidates are rejected if the invariant mass of the $D^*K\pi\pi$ is compatible with the nominal mass of the B_s^0 ($M_{B_s^0} = 5366.89$ MeV/c^2 [28]) within ±25 MeV/c^2 and the candidates belong to the enhanced background sample. This veto is applied only to D^0 decaying into $K^-\pi^-\pi^+\pi^+$, as no excess is observed for the $K^-\pi^+$ sample. The plot showing the distributions before and after applying this veto is shown in Fig. 7.6.

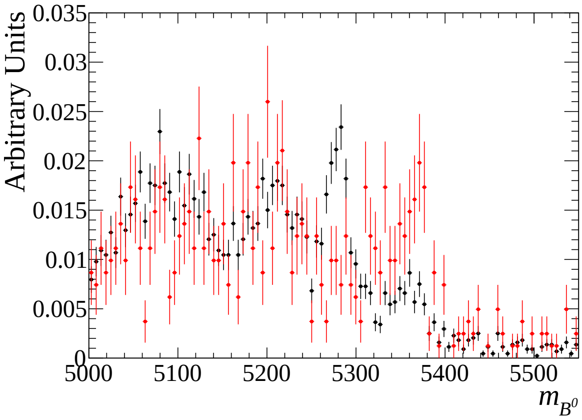


Figure 7.5: Comparison between the nominal mass variable with PV,$D^{*\pm}$,D^0 and $D^\pm$ constraints (black) and the mass variable removing the $D^\pm$ constraint in the background enhanced sample (red), that show an excess under the B_s^0 peak position.

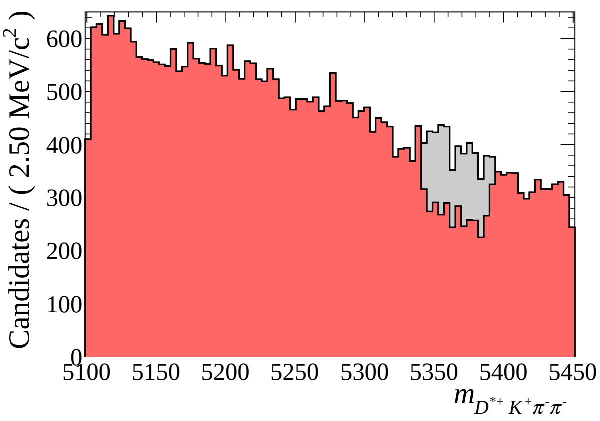


Figure 7.6: Plots of the invariant mass of the D^* $K\pi\pi$ combination. The distributions are given without (grey) and with (red) the $B_s^0 \to D^* K\pi\pi$ veto described in the text, for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ in Run 1 data sample.

$B^0 \to D^{*+}\pi^+\pi^-\pi^-$ The second source of single charm background found is the one coming from $B^0 \to D^{*+}\pi^+\pi^-\pi^-$, induced by $K \to \pi$ mis-ID. The pion mass hypothesis is assigned to the kaon and the invariant mass of the three hadrons plus the D^* meson is calculated. The comparison between the nominal mass distribution with the one without $D^\pm$ mass constraint is shown in Fig. 7.7. The presence of this kind of background is visible as an excess under the B^0 signal peak.

Candidates are rejected if the invariant mass of the $D^*\pi\pi\pi$ is compatible with the nominal mass of the B^0 ($M_{B^0} = 5279.62\,\mathrm{MeV/c^2}$ [28]) within $\pm 40\,\mathrm{MeV/c^2}$ or the candidates

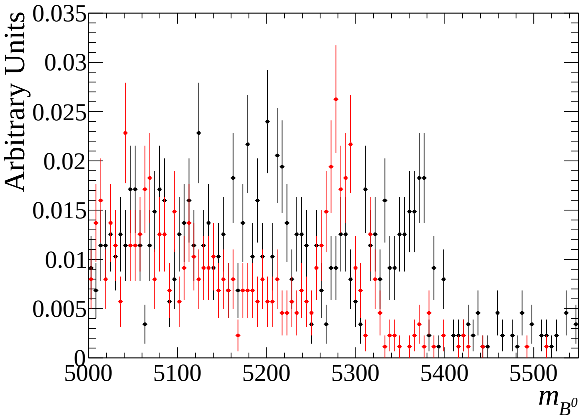


Figure 7.7: Comparison between the nominal mass variable (black) and the mass variable removing the $D^{\pm}$ constraint in the background enhanced sample (red), an excess in correspondence of the B^0 signal peak is visible.

belong to the enhanced background sample and the ratio of the kaon probability over the sum of the kaon and pion probabilities `ProbNNK`/(`ProbNN`π+`ProbNNK`) is lower than 0.4. The plots showing the $D^{*}\pi\pi\pi$ mass distribution before and after applying this veto are presented in Fig. 7.8.

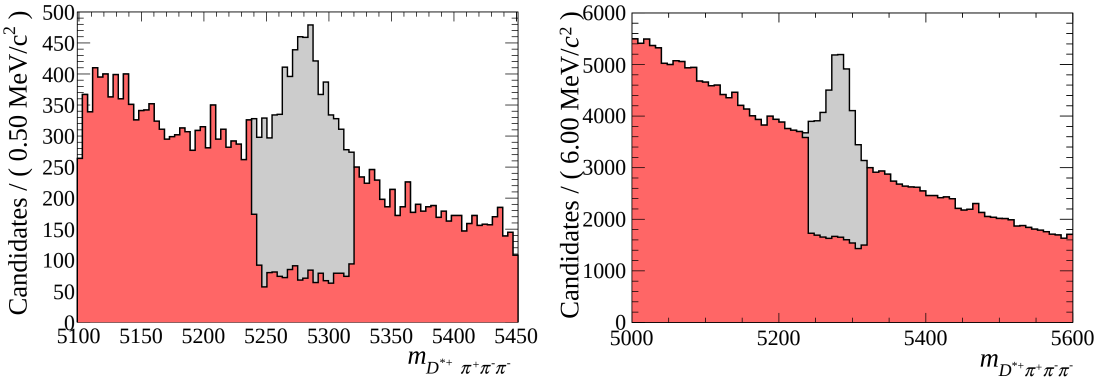


Figure 7.8: Plots of the invariant mass of the D^{*} $\pi\pi\pi$ combination. The distributions are given without (grey) and with (red) the $B^0 \to D^{*}\pi\pi\pi$ veto described in the text, for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ (left) and to $K^-\pi^+$ (right) on Run 1 data sample.

7.1.5 Multivariate Classifier

In order to suppress the combinatorial background, a multivariate algorithm is used, in particular a Boosted Decision Tree (BDT) [83, 84] is chosen.

7.1.6 PID Response Transformation on MC sample

A MC correction approach is applied on PID variables, in order to improve the data-MC agreement. Even if a measurement of CP violation, doesn't need calculation of efficiencies extracted from MC, any data-MC disagreement could affect the result of BDT output. Since PID variables are used as input for the BDT training in both D^0 final states samples, a PID response transformation is employed, this is described in detail in [88]. The basic idea behind is to transform a PID variable in such a way that its distribution matches the one seen in data in the calibration sample (for any kinematics of the track), and that the transformed variable is strongly correlated with the output of simulation. In that case, the correlations of the PID variable present in simulation (with other PID variables for the same track, or other parameters of the track and the event in general) will be preserved. In this approach the simulated variable x_{MC} is used to obtain the sampling variable ξ from the cumulative distribution $P_{MC}(x)$ corresponding to the simulated PID response:

$$\xi = P_{MC}(x_{MC} \mid p_T, \eta, N_{tr}) = \int_{-\infty}^{x} p_{MC}(y \mid p_T, \eta, N_{tr}) dy, \tag{7.1}$$

where $p_{MC}(x \mid p_T, \eta, N_{tr})$ is the distribution of simulated PID variable x as a function of p_T, η and N_{tr}. As a result the transformation of variable x_{MC} into x_{corr} that depends on the parameters p_t, η and N_{tr} is obtained:

$$x_{corr} = P_{exp}^{-1}(P_{MC}(x_{MC} \mid p_T, \eta, N_{tr}) \mid p_T, \eta, N_{tr}). \tag{7.2}$$

An important property of the dependence $x_{corr}(x_{MC})$ is that, if the difference between two distributions $p_{exp}(x)$ and $p_{MC}(x)$ is small, x_{MC} and x_{corr} are highly correlated. The strong correlation of the corrected and original PID response has important consequences: correlations of the PID variable with other properties of the event are preserved via the correlations in simulation, correlations between different PID variables for the same track are also preserved via correlations with simulated variables which allows to use the simulated samples after PID correction in multivariate classifier trainings, moreover small variations of the calibration PDFs will in turn result in small variations of the corrected PID variables, helping in estimating systematic uncertainties due to PID correction. The comparison of ProbNN distributions used in the BDT training for standard MC , corrected MC and sweighted data samples is presented in Fig. 7.9 for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ and $K^-\pi^+$ respectively. Moreover a study performed on D^0 decaying to $K^-\pi^+$ final state, comparing the BDT distribution between sweighted data and signal MC when the BDT has been trained using corrected PID variables with respect to uncorrected PID variables, showed that a better agreement is achieved.

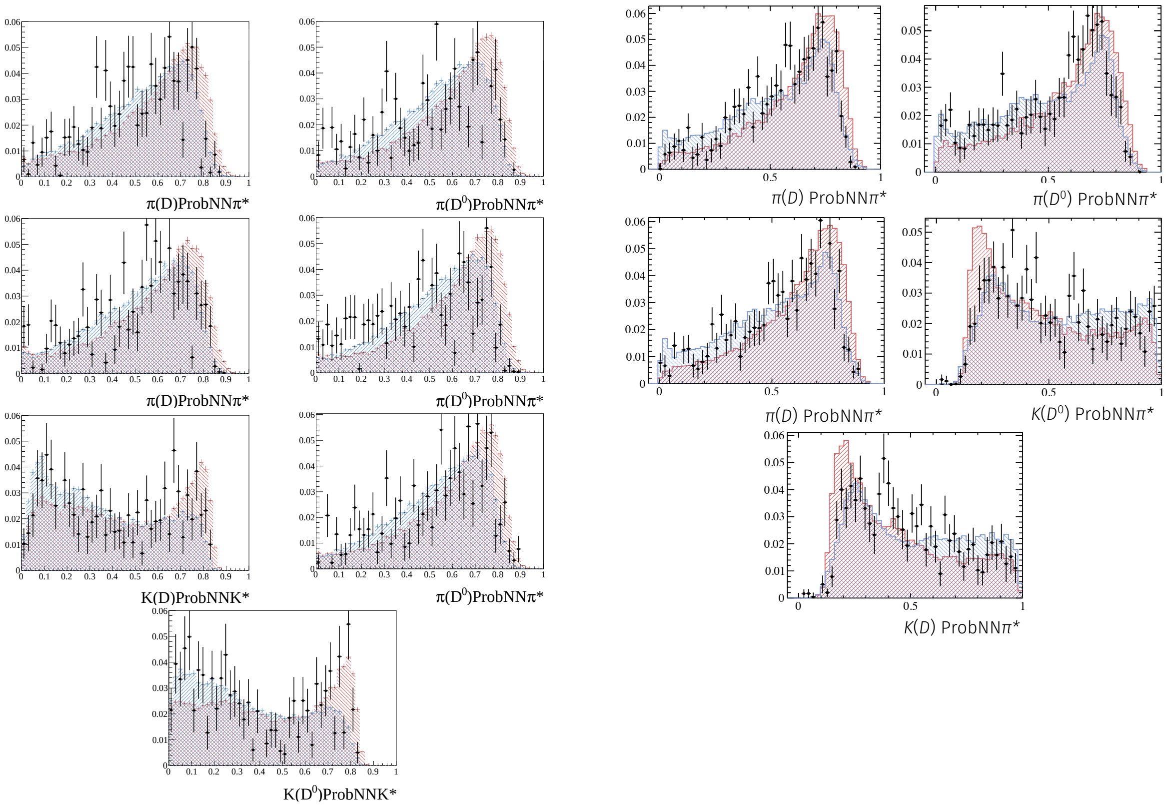


Figure 7.9: Comparison between PID variable distributions for Run 1 data (black), nominal MC sample (red) and corrected MC samples (blue). Plots for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ are shown on the left while those for D^0 decaying into $K^-\pi^+$ are shown on the right.

7.1.7 BDT training

For the BDT training, a signal MC sample is used as signal template, while the data sample corresponding to the mass range above $5600\,\mathrm{MeV/c^2}$ for the $K^-\pi^-\pi^+\pi^+$ and above $5400\,\mathrm{MeV/c^2}$ for $K^-\pi^+$ serves to model the shape of the combinatorial background. The BDT is trained on half of these samples while its performances are tested on the other half. The previously described cut-based preselection and the vetoes, are already applied to the training samples. Two separate trainings, for the $K^-\pi^-\pi^+\pi^+$ and the $K^-\pi^+$ final states have to be performed, since PID variables are used in the BDT, on these variables the PID correction procedure described above has been applied. The input variables used in $K^-\pi^-\pi^+\pi^+$ BDT training along with the ones used for $K^-\pi^+$ final state, ordered according to their ranking are listed in Table 7.7.

Table 7.7: List of input variables used in the training of the BDT for $K^-\pi^-\pi^+\pi^+$ (left) and $K^-\pi^+$ (right) final states, order according to variables importance.

Variable
K from D^0 ProbNNk
K from D^+ ProbNNk
D^+ decay time significance
π_3 from D^0 ProbNNπ
cos($\angle$(B,D^*))
DTF χ^2
D^0 decay time significance
π_2 from D^+ ProbNNπ
B^0 IP χ^2 wrt own PV
cos($\angle$(D^0,π_1))
π_1 from D^+ ProbNNπ
π_2 from D^0 ProbNNπ
B^0 χ^2 wrt own PV
K from D^+ p_T
K from D^0 p_T
cos($\angle$(D^0,π_2))
π_1 from D^0 ProbNNπ
D^0 FD χ^2 wrt own PV
D^+ decay time
D^0 decay time

Variable
DTF χ^2
B^0, D^0, D^+ IP χ^2 wrt own PV
D^0, D^+ FD χ^2 wrt own PV
D^0 decay time significance
D^+ decay time significance
K from D^+ p_T
π_1 from D^+ p_T
π_2 from D^+ p_T
D^*p_T
D^+p_T
B^0p_T
$m(D^*) - m(D^0)$
K,π_1,π_2 from D^+ ProbNN
K,π from D^0 ProbNN
π from D^* ProbNN
cos($\angle$(B,D^*))
cos($\angle$(D^0,π))
cos($\angle$(D^+,K))
cos($\angle$(D^+,π_1))
cos($\angle$(D^+,π_2))

The BDTs are built out of 850 trees for $K^-\pi^-\pi^+\pi^+$ and 800 for $K^-\pi^+$ respectively, and the depth of the trees is limited to three. At each node at least 2.5 % of the training events have to be present. The variables are scanned at 40 points to find the optimal cut value. For the boosting the AdaBoost method [85] with a boost factor of $\beta = 0.5$ is deployed. The BDT outputs, with training and testing samples superimposed are presented in Fig. 7.10.

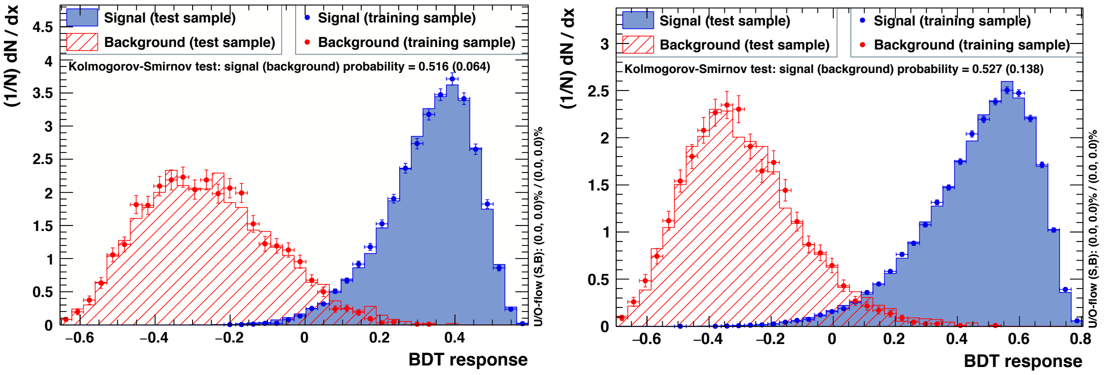


Figure 7.10: Comparison of the BDT response for training and testing samples considering the D^0 decaying into $K^-\pi^-\pi^+\pi^+$ (left) and $K^-\pi^+$ (right) on Run 1 sample.

7.1.8 Figure of merit for optimizing the BDT cut point

For a decay-time dependent and flavour-tagged measurement of CP violation the figure of merit that should be maximized is the inverse variance of the parameter that is of interest, in this case $\sin\phi_d^{eff}$ [89]. There are several properties of the data sample which directly influence the statistical power in this relation. Besides a high effective signal size, small decay time error estimates σ_t, as well as small mistag probabilities ω will lead to a more precise measurement. Also the decay time t of the B^0 candidates impacts the quality of the measurement. The distribution of these observables is likely to change depending on the BDT cut that is chosen, so that a proper optimization of the working point in terms of the BDT response needs to take this into account. The figure of merit (FOM) used in $B^0 \to D^{*\pm}D^{\mp}$ analysis is based on what described in [90] and extended including signal ${}_s\mathcal{W}eights$ to obtain the according signal distributions as well as the effective signal size itself. It is defined as

$$\mathrm{FOM} \equiv \mathrm{Q}_{mod} = \frac{(\sum_i s_{wi})^2}{\sum_i s_{wi}^2}\bar{\mathrm{D}}_{mod}, \tag{7.3}$$

where $\bar{\mathrm{D}}_{mod}$ is given by

$$\bar{\mathrm{D}}_{mod} = \frac{1}{\sum_i s_{wi}}\sum_i (1-2\omega_i)^2 e^{-(\Delta m_d \sigma(t_i))^2} \cdot \mathrm{X}_i \cdot s_{wi}, \tag{7.4}$$

with X_i being

$$\mathrm{X}_i = \Big[\frac{2d_i\lambda\sin(\Delta m_d t_i)}{1+\lambda^2+d_i(1-2\omega_i)e^{-(\Delta m_d\sigma(t_i))^2/2}\Big(-2\lambda sin(\Delta m_d t_i)\sin\phi_d^{eff}-(1-\lambda^2)\cos(\Delta m_d t_i)\Big)}\Big]^2. \tag{7.5}$$

The time resolution is described by $\sigma(t_i)$ and d_i denotes the decision on the production flavour of the B^0 meson. The FOM is at first calculated only in parts for a better understanding of the influence of the single contributions, and then calculated completely. The tagging power term is given by

$$\mathrm{FOM}_{\epsilon D^2} \equiv \frac{1}{\sum_i s_{wi}}\sum_i (1-2\omega_i)^2\cdot s_{wi}, \tag{7.6}$$

the power of the decay time resolution is given by

$$\mathrm{FOM}_{\sigma_t} \equiv \frac{1}{\sum_i s_{wi}}\sum_i e^{-(\Delta m_d\sigma(t_i))^2}\cdot s_{wi}. \tag{7.7}$$

The effective signal size, which is comparable to the simple signal significance $\mathrm{FOM}_{sig} = \frac{S}{\sqrt{S+B}}$, is given by

$$\mathrm{FOM}_{S_{eff}} \equiv \frac{(\sum_i s_{wi})^2}{\sum_i s_{wi}^2}, \tag{7.8}$$

and

$$\mathrm{FOM}_{B^0} \equiv \frac{1}{\sum_i s_{wi}} X_i\cdot s_{wi}, \tag{7.9}$$

denotes the term associated to $\sin 2\phi_d^{eff}$. The optimization of the cut value to apply on the BDT output classifier is done on data. At each cut point of the BDT classifier the FOM is newly calculated including the event-by-event mistag rate ω and the resolution.

Fig. 7.11 shows the curves for the complete figure of merit along with its components, obtained from a scan in the BDT response, for D^0 decaying to $K^-\pi^-\pi^+\pi^+$ and to $K^-\pi^+$ respectively. Mass fits at selected cut points are shown in Fig. 7.12, for the D^0 decaying into $K^-\pi^-\pi^+\pi^+$ as an example. The optimal cut points are found to be at 0.18 for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ and at 0.16 for D^0 decaying into $K^-\pi^+$.

7.1.9 Multiple Candidates

In order to keep only one of the multiple candidates that remains in the selected events, a random suppression is performed. The study of the mass distribution of candidates in data showed that between multiple candidates a mass correlation is present (see Fig. 7.13), which leads to an overestimation of signal yield. On data the fraction of events where multiple candidates are present is found to be 7%. In order to estimate the effect of removing these multiple candidates, the signal MC sample is exploited. The estimate reduction of the signal yield in D^0 decaying into $K^-\pi^-\pi^+\pi^+$ is found to be 6.4%. In order

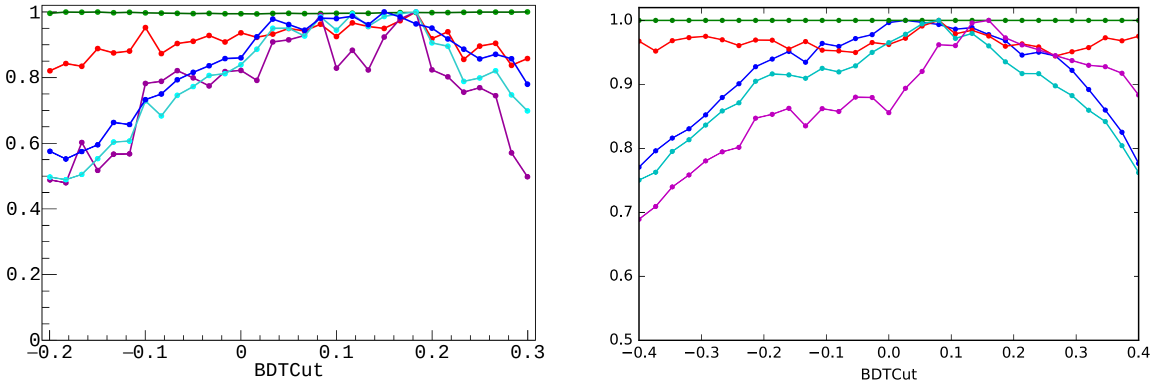


Figure 7.11: Scan in BDT response for the FOM with its studied components: in blue the effective signal term, in green the decay time resolution power, in red the tagging power term, in turquoise the FOM without the ϕ_d^{eff} term and in purple the complete FOM. On the left the FOM scan for D^0 going to $K^-\pi^-\pi^+\pi^+$ and on the right for D^0 decaying into $K^-\pi^+$ on Run 1 sample.

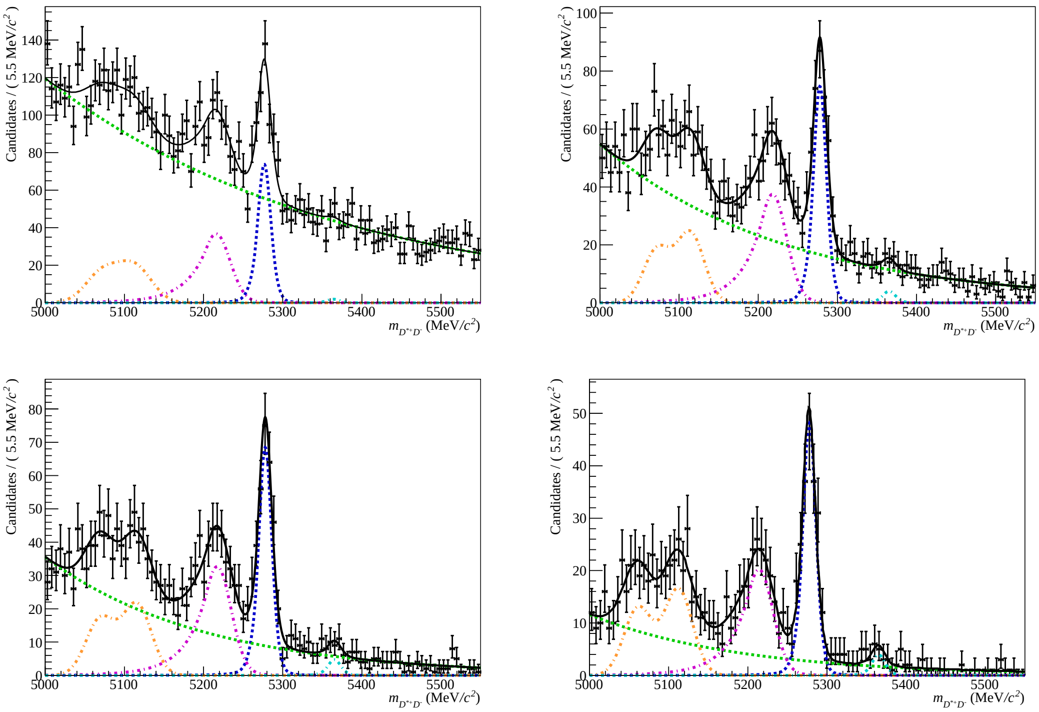


Figure 7.12: Fits to the D^0 to $K^-\pi^-\pi^+\pi^+$ data sample with different cuts on the BDT response. BDT cut points are at -0.2 (top left), 0.07 (top right), 0.24 (bottom left) and 0.3 (bottom right).

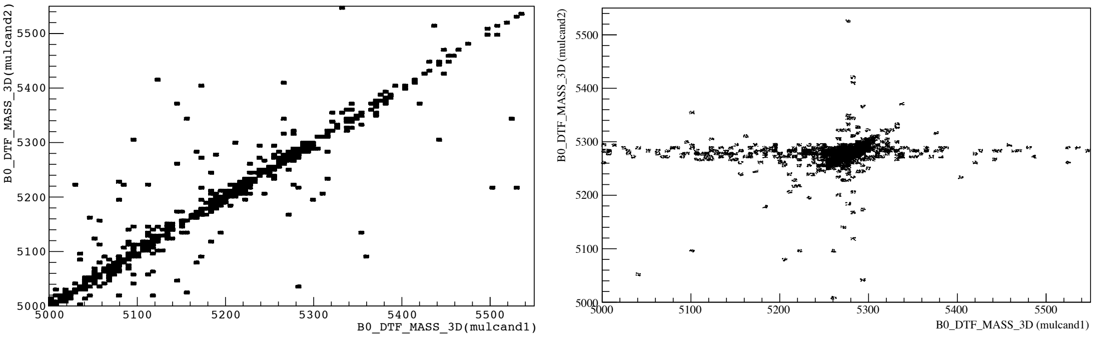


Figure 7.13: 2-dimensional histogram for the nominal mass of the first multiple candidate versus the second, for Run 1 data sample (left) and signal MC (right), where D^0 goes into $K^-\pi^-\pi^+\pi^+$.

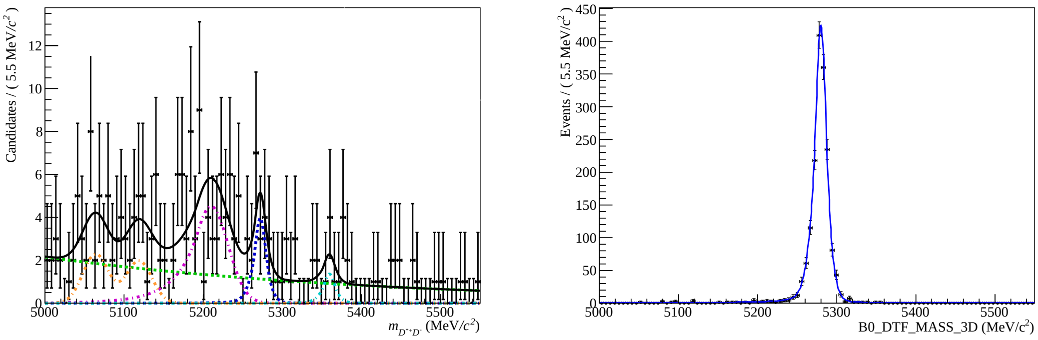


Figure 7.14: Mass fit performed on suppressed candidates considering the Run 1 data sample (left) and signal MC (right) of D^0 decaying into $K^-\pi^-\pi^+\pi^+$.

to cross check this, a mass fit to candidates which survive the random suppression and to candidates which are eliminated is performed (see Fig. 7.14).

The results of the cross check described above are summarized in Table 7.8, it shows an agreement between the expected reduction and measured one, based on the estimate done on MC signal. The resulting signal yield with and without multiple candidates suppression however are compatible.

7.1.10 Final Selection

The final selection step is reducing the range of the invariant $D^{*+}D^-$ mass from 5150 MeV/c^2 up to 5550 MeV/c^2 and the range of the decay time from 0.3 ps to 10.3 ps, these cuts have nearly a 100% signal efficiency. Plots of the invariant D^* D mass distribution after these final cuts are shown in Fig. 7.15 for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ and

Table 7.8: Estimated signal yield for data and MC sample after random suppression of multiple candidates.

	Signal MC Sample	Data Sample
$N_{D^*D}^{\text{no suppression}}$	27296±165	316±22
$N_{D^*D}^{\text{suppression measured}}$	25663±160	303±21
$N_{D^*D}^{\text{suppression expected}}$	25665	296

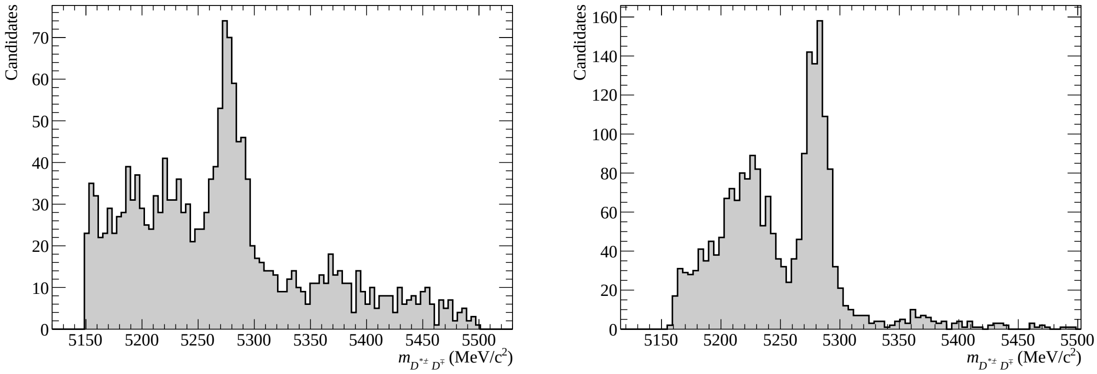


Figure 7.15: Plots of the invariant D^* D mass after all selection steps have been applied to D^0 decaying into $K^-\pi^-\pi^+\pi^+$ (left) and $K^-\pi^+$ (right) Run 1 samples respectively.

$K^-\pi^+$ respectively.

On the Run 2 samples the very same selection steps are applied. The overall offline selection performances for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ are quoted in Table 7.9 . These numbers represent the efficiencies for the 2011/2012 and 2015/2016 merged samples respectively.

Table 7.9: Overall offline selection performance of D^0 decaying into $K^-\pi^-\pi^+\pi^+$ final state sample in Run 1 and Run2.

	$\varepsilon_{\text{Sig}}^{\text{Run1}}(\%)$	$\varepsilon_{\text{Sig}}^{\text{Run2}}(\%)$
Preselection	89.23±0.71	88.61±0.66
BDT	94.42±0.84	97.83±0.83
Multiple Candidates	89.48±0.76	88.39±0.72
Vetoes	97.04±0.79	95.89±0.75
Charmless Vetoes	99.16±0.82	99.34±0.78
Total	72.54±1.36	72.99±1.30

7.2 Mass Fit

In this section, the fit of the invariant mass distribution is described. The invariant mass is used to discriminate between signal and background candidates, using the ${}_s\mathcal{P}lot$ method. On signal MC the linear Pearson correlation coefficient between the decay time and the invariant mass is calculated as $\rho = 0.05$. So the ${}_s\mathcal{W}eights$ can be applied in a decay time fit in which the *CP* observables are determined. The decay time fit is described in Sec. 7.3.

For the mass fit the extended maximum likelihood method is used (Eq. (5.8)).

The mass fit is performed into two stages, at first a wide mass range is considered, corresponding to the invariant $D^{*+}D^-$ mass between 5000 and 5500 MeV/c^2. This leaves enough candidates in the upper mass sideband to determine the shape of the combinatorial background.

The probability density function (PDF) $\mathcal{P}^s$ consists of five components $\mathcal{P}^s_j$, $B^0 \to D^{*\pm}D^{\mp}$ signal ($j = B^0$), $B^0_s \to D^{*+}D^-$ background ($j = B^0_s$), background from $B^0 \to D^{*+}D^-_s$ ($j = B^0 \to D^{*+}D^-_s$), background from $B^0 \to D^{*+}D^{*-}$ ($j = B^0 \to D^{*+}D^{*-}$) and combinatorial background ($j = \text{comb}$):

$$N^s\mathcal{P}^s = N^s_{B^0}\mathcal{P}^s_{B^0} + N^s_{B^0_s}\mathcal{P}^s_{B^0_s} + N^s_{B^0 \to D^{*+}D^-_s}\mathcal{P}^s_{B^0 \to D^{*+}D^-_s} + N^s_{B^0 \to D^{*+}D^{*-}}\mathcal{P}^s_{B^0 \to D^{*+}D^{*-}} + N^s_{\text{comb}}\mathcal{P}^s_{\text{comb}} \ . \quad (7.10)$$

From this fit the fraction of each background component with respect to the total background is calculated in a narrow mass range corresponding to 5150-5500 MeV/c^2. A second mass fit is then repeated, fixing the background fractions and the parameters that were floating in the previous fit. The probability density function at this stage is made of two terms

$$N^s\mathcal{P}^s = N^s_{B^0}\mathcal{P}^s_{B^0} + N^s_{\text{Bkg}}\mathcal{P}^s_{\text{Bkg}}, \quad (7.11)$$

where $\mathcal{P}^s_{\text{Bkg}}$ is the sum of all the background components. The mass fits are performed separately, distinguishing between the two runs of data-taking Run 1 and Run2, and between the two D^0 final states ($K\pi\pi\pi$ and $K\pi$), for both only the $D^\pm$ final state $K\pi\pi$ is considered.

7.2.1 Reconstructed mass PDF

B^0 signal: The reconstructed mass of the B^0 signal component is parametrized with the sum of two crystal-ball PDFs [74], with the same mean μ_{B^0} but different width parameters σ_i and tails to opposite directions ($\alpha_1 > 0$, $\alpha_2 < 0$).

In the first mass fit, the exponent of the power law part n, the parameters α_1 and α_2 of the power law functions and the fraction f of the two crystal-ball PDFs are taken from simulation. The widths σ_1 and σ_2 are also determined on signal MC.

$B^0_s \to D^{*+}D^-$ background: Apart from the B^0 signal the heavier B^0_s decaying via the same decay channel is also present. In the first mass fit it is parametrized with the same PDF as the B^0 signal component, sharing the width and tail parameters. The difference between the peak positions is fixed to the world average $\Delta m_{B^0_s - B^0} = \mu_{B^0_s} - \mu_{B^0} = 87.35\,\text{MeV/c}^2$ [28].

$B^0 \to D^{*+}D_s^-$ background: Although the contribution from misidentified kaons is suppressed in the selection a significant amount of candidates from $B^0 \to D^{*+}D_s^-$ remains in the data sample. These candidates can be described with the sum of two crystal-ball PDFs. The power law exponent n is fixed to 10, while the fraction, the tail parameters and the widths are determined on simulated $B^0 \to D^{*+}D_s^-$ events which have been reconstructed as $B^0 \to D^{*\pm}D^{\mp}$. The mean of the PDF is a floating parameter in the first fit to the invariant mass distribution then it is fixed in the second fit.

$B^0 \to D^{*+}D^{*-}$ background: In this analysis only the decay of a $D^{*\pm} \to D^0\pi^\pm$ is considered. However, the $D^{*\pm}$ meson can also decay as $D^{*\pm} \to D^\pm\pi^0$. If the π^0 is not detected, the final states can be combined to form $B^0 \to D^{*\pm}D^{\mp}$, whose position is displaced downwards by a little more than the mass of a pion. To describe the shape of this background component, the sum of two Gaussians is used, with different means to take into account the characteristic double peak structure. All the shape parameters are left floating in the first mass fit, while they are fixed in the second.

Combinatorial background: The reconstructed mass PDF of the combinatorial background component is modelled as an exponential function (see Eq. (5.9)). The exponent parameter β is left free in the first fit and fixed in the second one.

7.2.2 MC Mass Fits

A maximum-likelihood fit to the invariant $D^{*+}D^-$ mass distribution of $B^0 \to D^{*\pm}D^{\mp}$ signal MC in the range 5000 MeV/c^2 to 5500 MeV/c^2 is performed to determine the shape parameters. For each run of data-taking, two MC samples are used for the different D^0 final states, where the full selection is applied. The background categories 0 (signal) and 50 (LowMassBackground, *i.e.* missing photons) are considered as signal template. The fit results from D^0 decaying to $K^-\pi^-\pi^+\pi^+$ are listed in Table 7.10, while the corresponding plots of the mass distribution overlaid with the projection of the PDF are given in Fig. 7.16.

Additionally, a fit to the invariant mass distribution of $B^0 \to D^{*+}D_s^-$ MC sample reconstructed as $B^0 \to D^{*\pm}D^{\mp}$ is performed. Also in this case for each run of data-taking, two MC samples are used for the different D^0 final states, again the full selection is applied. Here, the MC-truth information of the three D mesons is used instead of the background category to select signal events. The fit results of the two crystal-ball PDFs for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ are listed in Table 7.11 , where all the parameters are floating except for the exponent of the power law which is fixed to 10. The corresponding plots are presented in Fig. 7.17.

7.2.3 Fit Results

The fixed parameters used in the mass fit to data for D^0 decaying to $K^-\pi^-\pi^+\pi^+$ are listed in Table 7.12.

Table 7.10: Fit results of the mass fit to $B^0 \to D^{*\pm}D^{\mp}$ ($D^0 \to K^-\pi^-\pi^+\pi^+$) signal MC.

Sample	Parameter	Fitted Value
Run 1	$\mu_{B^0}^{\rm MC}$ (MeV/c^2)	5279.984 ± 0.071
	$\sigma_1^{\rm MC}$ (MeV/c^2)	6.86 ± 0.24
	$\sigma_2^{\rm MC}$ (MeV/c^2)	10.90 ± 0.33
	$f^{\rm MC}$	0.50 ± 0.05
	$\alpha_1^{\rm MC}$	1.54 ± 0.10
	$\alpha_2^{\rm MC}$	-2.36 ± 0.06
	$n^{\rm MC}$	2.16 ± 0.11
Run 2	$\mu_{B^0}^{\rm MC}$ (MeV/c^2)	5280.03 ± 0.14
	$\sigma_1^{\rm MC}$ (MeV/c^2)	8.39 ± 0.21
	$\sigma_2^{\rm MC}$ (MeV/c^2)	18.18 ± 1.40
	$f^{\rm MC}$	0.85 ± 0.03
	$\alpha_1^{\rm MC}$	2.15 ± 0.12
	$\alpha_2^{\rm MC}$	-2.51 ± 0.18
	$n^{\rm MC}$	1.74 ± 0.22

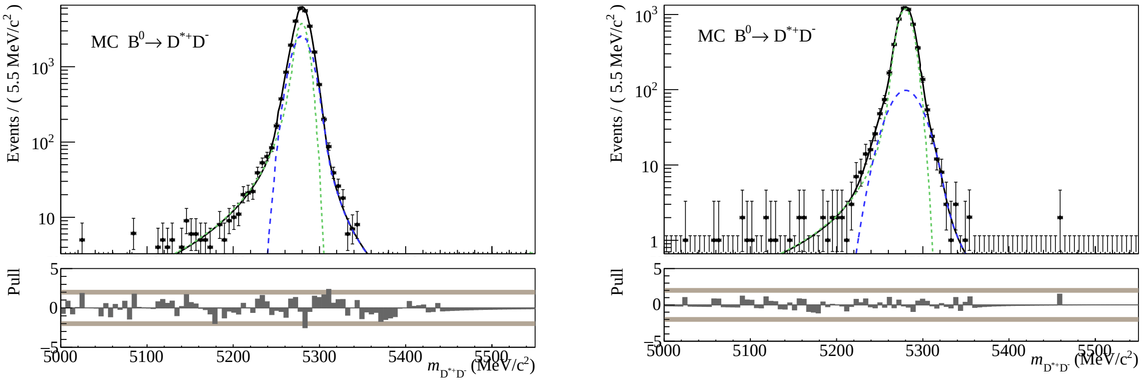


Figure 7.16: Mass distribution of the $B^0 \to D^{*\pm}D^{\mp}$ ($D^0 \to K^-\pi^-\pi^+\pi^+$) signal MC sample overlaid with the projection of the fitted PDF, for Run 1 (left) and Run 2 (right). The y-axis has a logarithmic scale.

In Table 7.13 the results of the floating shape parameters in the wide mass fit are shown, for $D^0 \to K^-\pi^-\pi^+\pi^+$ in Run1 and Run2 respectively. These shape parameters are fixed in the narrow mass fit, while the yields listed in Table 7.14, are used to estimate the background fractions, which are also fixed in the narrow fit.

The results obtained in the narrow fit are shown in Table 7.15. The total number of B^0 signal candidates of D^0 decaying into $K^-\pi^-\pi^+\pi^+$ are $N_{B^0}^{D^0 \to K^-\pi^-\pi^+\pi^+} = 348 \pm 23$ for Run1 and $N_{B^0}^{D^0 \to K^-\pi^-\pi^+\pi^+} = 441 \pm 28$ for Run2, while for the D^0 decaying into $K^-\pi^+$ the

Table 7.11: Fit results of the mass fit to $B^0 \to D^{*+} D_s^-$ $(D^0 \to K^-\pi^-\pi^+\pi^+)$ MC.

Sample	Parameter	Fitted Value
Run 1	$\mu_{B^0}^{\rm MC}$ (MeV/c^2)	5220.61 ± 0.61
	$\sigma_1^{\rm MC}$ (MeV/c^2)	16.88 ± 0.72
	$\sigma_2^{\rm MC}$ (MeV/c^2)	26.96 ± 5.06
	$f^{\rm MC}$	0.92 ± 0.05
	$\alpha_1^{\rm MC}$	0.63 ± 0.04
	$\alpha_2^{\rm MC}$	-1.36 ± 0.35
Run 2	$\mu_{B^0}^{\rm MC}$ (MeV/c^2)	5219.58 ± 1.01
	$\sigma_1^{\rm MC}$ (MeV/c^2)	18.63 ± 0.87
	$\sigma_2^{\rm MC}$ (MeV/c^2)	26.85 ± 8.52
	$\alpha_1^{\rm MC}$	0.81 ± 0.06
	$\alpha_2^{\rm MC}$	-1.26 ± 0.45

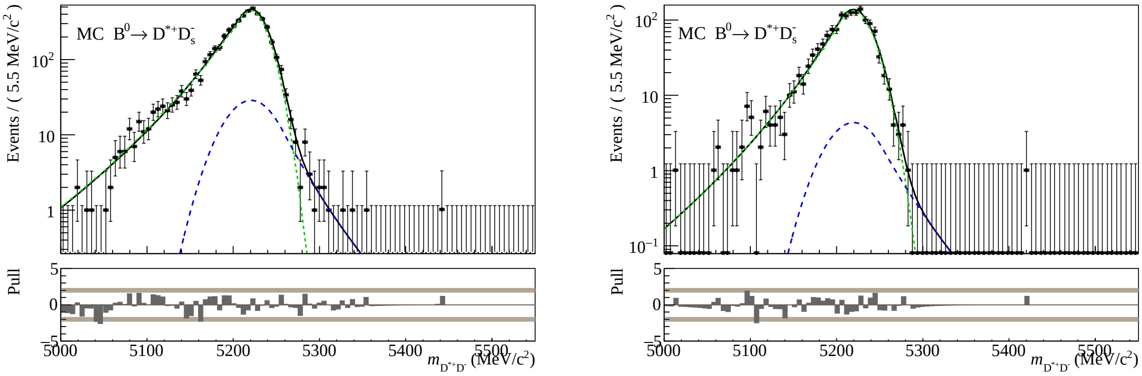


Figure 7.17: Mass distribution of the $B^0 \to D^{*+} D_s^-$ $(D^0 \to K^-\pi^-\pi^+\pi^+)$ MC sample reconstructed as $B^0 \to D^{*\pm} D^{\mp}$ overlaid with the projection of the two crystal-ball PDFs, for Run 1 (left) and Run 2 (right). The y-axis has a logarithmic scale.

values found are $N_{B^0}^{D^0 \to K^-\pi^+} = 823 \pm 30$ for Run 1 and $N_{B^0}^{D^0 \to K^-\pi^+} = 1111 \pm 35$ for Run2 respectively. In Fig. 7.18 and Fig. 7.19 the data samples are plotted overlaid with the PDF projections and their components for $D^0 \to K^-\pi^-\pi^+\pi^+$ and $D^0 \to K^-\pi^+$ respectively.

Table 7.12: Fixed mass parameters for $D^0 \to K^-\pi^-\pi^+\pi^+$.

Sample	Parameter	Fixed Value
Run 1	$\sigma_{1,B^0\to D^{*\pm}D^{\mp}}$	6.86
	$\sigma_{2,B^0\to D^{*\pm}D^{\mp}}$	10.90
	$f_{B^0\to D^{*\pm}D^{\mp}}$	0.50
	$\alpha_{1,B^0\to D^{*\pm}D^{\mp}}$	1.54
	$\alpha_{2,B^0\to D^{*\pm}D^{\mp}}$	-2.36
	$n_{B^0\to D^{*\pm}D^{\mp}}$	2.16
Run 2	$\sigma_{1,B^0\to D^{*\pm}D^{\mp}}$	8.39
	$\sigma_{2,B^0\to D^{*\pm}D^{\mp}}$	18.18
	$f_{B^0\to D^{*\pm}D^{\mp}}$	0.85
	$\alpha_{1,B^0\to D^{*\pm}D^{\mp}}$	2.15
	$\alpha_{2,B^0\to D^{*\pm}D^{\mp}}$	-2.51
	$n_{B^0\to D^{*\pm}D^{\mp}}$	1.74
Run 1	$\sigma_{1,B^0\to D^{*+}D_s^-}$	16.88
	$\sigma_{2,B^0\to D^{*+}D_s^-}$	26.96
	$f_{B^0\to D^{*+}D_s^-}$	0.92
	$\alpha_{1,B^0\to D^{*+}D_s^-}$	0.63
	$\alpha_{2,B^0\to D^{*+}D_s^-}$	-1.36
	$n_{B^0\to D^{*+}D_s^-}$	10.
Run 2	$\sigma_{1,B^0\to D^{*+}D_s^-}$	18.63
	$\sigma_{2,B^0\to D^{*+}D_s^-}$	26.85
	$f_{B^0\to D^{*+}D_s^-}$	0.96
	$\alpha_{1,B^0\to D^{*+}D_s^-}$	0.81
	$\alpha_{2,B^0\to D^{*+}D_s^-}$	-1.26
	$n_{B^0\to D^{*+}D_s^-}$	10.

Table 7.13: Results of the floating shape parameters in the wide mass fit for $D^0 \to K^-\pi^-\pi^+\pi^+$.

Sample	Parameter	Value
Run 1	μ_{B^0}	5278.15 ± 0.78
	$\mu_{B^0 \to D^{*+}D_s^-}$	5222.62 ± 3.40
	$\beta_{K^-\pi^-\pi^+\pi^+K^+\pi^-\pi^-}$	-0.00409 ± 0.00021
	$f_{B^0 \to D^{*+}D^{*-}}$	0.604 ± 0.085
	$\mu_{1,B^0 \to D^{*+}D^{*-}}$	5119.99 ± 16.19
	$\mu_{2,B^0 \to D^{*+}D^{*-}}$	5072.43 ± 7.05
	$\sigma_{B^0 \to D^{*+}D^{*-}}$	20.52 ± 4.65
Run 2	μ_{B^0}	5278.89 ± 0.70
	$\mu_{B^0 \to D^{*+}D_s^-}$	5217.38 ± 4.78
	$\beta_{K^-\pi^-\pi^+\pi^+K^+\pi^-\pi^-}$	-0.00351 ± 0.00014
	$f_{B^0 \to D^{*+}D^{*-}}$	0.506 ± 0.071
	$\mu_{1,B^0 \to D^{*+}D^{*-}}$	5114.68 ± 3.59
	$\mu_{2,B^0 \to D^{*+}D^{*-}}$	5075.226 ± 4.078
	$\sigma_{B^0 \to D^{*+}D^{*-}}$	14.35 ± 2.54

Table 7.14: Results for the yields of the background components from the wide mass fit for $D^0 \to K^-\pi^-\pi^+\pi^+$ Run1 and Run2.

Sample	Parameter	Integral	Fraction
Run 1	$N_{B_s^0 \to D^{*+}D^-}$	26	0.02
	N_{comb}	1030	0.8
	$N_{B^0 \to D^{*+}D_s^-}$	242	0.94
	$N_{B^0 \to D^{*+}D^{*-}}$	15	0.06
Run 2	$N_{B_s^0 \to D^{*+}D^-}$	1	0.0005
	N_{comb}	2086	0.89
	$N_{B^0 \to D^{*+}D_s^-}$	250	0.99
	$N_{B^0 \to D^{*+}D^{*-}}$	1	0.005

Table 7.15: Results for the yields in the narrow mass fit.

Sample		Parameter	Fitted Value
$D^0 \to K^-\pi^-\pi^+\pi^+$	Run 1	N_{B^0}	347 ± 23
		N_{Bkg}	1343 ± 39
	Run2	N_{B^0}	441 ± 28
		N_{Bkg}	2356 ± 52
$D^0 \to K^-\pi^+$	Run1	N_{B^0}	823 ± 30
		N_{Bkg}	845 ± 29
	Run 2	N_{B^0}	1111 ± 35
		N_{Bkg}	1663 ± 41

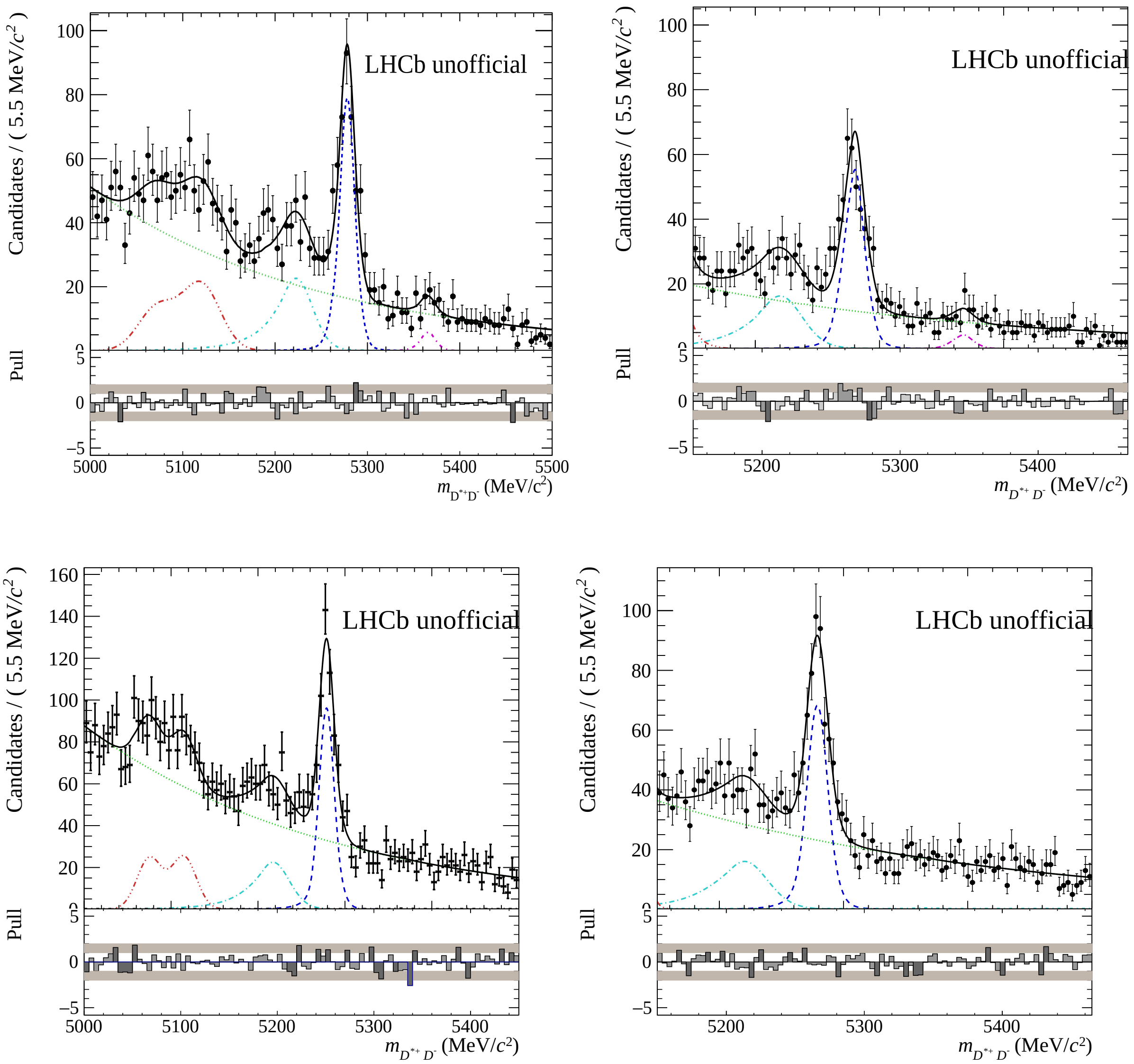


Figure 7.18: Plot of the reconstructed mass of the $B^0 \to D^{*\pm}D^{\mp}$ ($D^0 \to K^-\pi^-\pi^+\pi^+$) data sample with the projected PDF and pull distributions for Run1 (upper plot) and Run2 (bottom plot). Left plots are the result of the wide mass fit while plots on the right are obtained from the narrow mass fit. Besides the data points and the full PDF (solid black) the projections of the B^0 signal (dashed blue), the $B^0_s \to D^{*+}D^-$ background (short-dash-dotted magenta), the $B^0 \to D^{*+}D^-_s$ background (short-dash-dotted turquoise), the $B^0 \to D^{*+}D^{*-}$ background (dash-dotted red) and the combinatorial background (dotted green) are shown.

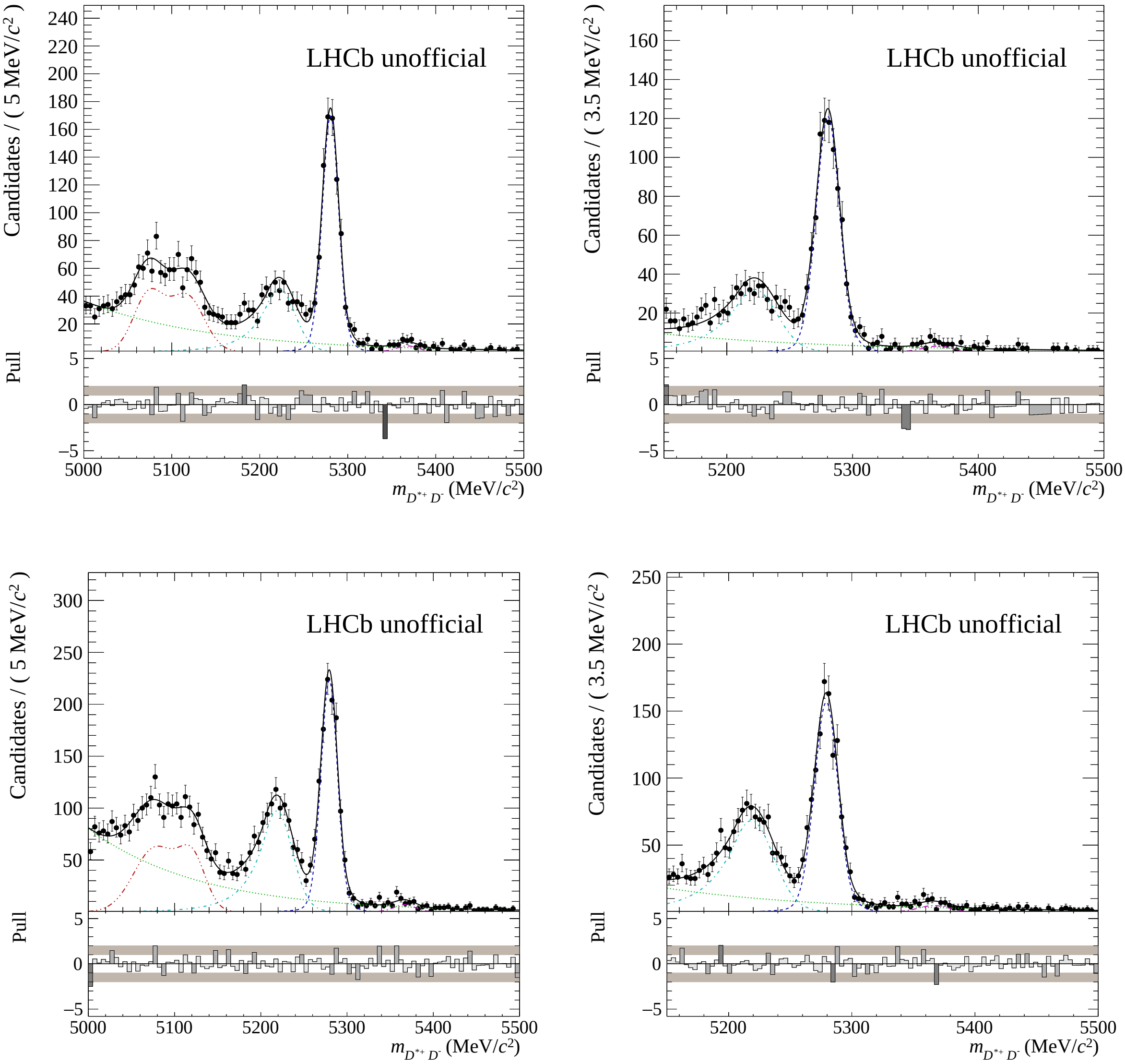


Figure 7.19: Plot of the reconstructed mass of the $B^0 \to D^{*\pm}D^{\mp}$ ($D^0 \to K^-\pi^+$) data sample with the projected PDF and pull distributions for Run1 (upper plots) and Run2 (bottom plots). Left plots are the result of the wide fit while plots on the right are obtained from the narrow fit. Besides the data points and the full PDF (solid black) the projections of the B^0 signal (dashed blue), the $B_s^0 \to D^{*+}D^-$ background (short-dash-dotted magenta), the $B^0 \to D^{*+}D_s^-$ background (short-dash-dotted turquoise), the $B^0 \to D^{*+}D^{*-}$ background (dash-dotted red) and the combinatorial background (dotted green) are shown.

7.3 Decay Time Fit

In order to measure the CP coefficients S_{D^*D}, ΔS_{D^*D}, C_{D^*D} and ΔC_{D^*D} on data, the ${}_s\mathcal{F}it$ technique is used. This method consists in weighting the data sample on a per-candidate basis using the ${}_s\mathcal{W}eights$ extracted from the mass fit, as described in Sec. 7.2. Once the data sample is weighted, its composition is described by a single PDF associated to signal $B^0 \to D^{*\pm}D^{\mp}$ candidates, whereas all background components are subtracted in the reweighting. The PDF describing the measured B^0 decay time distribution and tag decisions $\vec{d} = (d_{\rm OS}, d_{\rm SS})$, given the mistag probability estimates $\vec{\eta} = (\eta_{\rm OS}, \eta_{\rm SS})$ can be expressed as follows:

$$P\left(t, f, \vec{d} \,|\, \vec{\eta}\right) = \epsilon(t)\Big(\mathcal{P}(t', f, \vec{d} \,|\, \vec{\eta}) \otimes \mathcal{R}(t-t')\Big), \tag{7.12}$$

where $\mathcal{P}(t', f, \vec{d} \,|\, \vec{\eta})$ is the PDF describing the distribution of the true decay times t' , $\mathcal{R}(t-t')$ is the decay time resolution function described in Sec. 7.3.1, while $\epsilon(t)$ describes the decay time acceptance, discussed in Sec. 7.3.2. The PDF describing the B^0 decay time distribution can be written in the most general case as follows:

$$\mathcal{P}(t|\vec{d},\vec{\eta}) = Ne^{-\frac{t}{\tau}}\Big[C^{eff}_{\cosh}\cosh\Big(\frac{\Delta\Gamma t}{2}\Big) + C^{eff}_{\sinh}\sinh\Big(\frac{\Delta\Gamma t}{2}\Big) + C^{eff}_{\cos}\cos\Big(\Delta m t\Big) + C^{eff}_{\sin}\Big(\Delta m t\Big)\Big]. \tag{7.13}$$

The four effective CP coefficients inside Eq. (7.13) depend on the final state (f=$D^{*+}D^-$ or f=$D^{*-}D^+$), the tagging decision ($d_i = \pm 1$, $d_i = 0$, for i=OS,SS), the mistag and the tagging efficiency of the OS and SS taggers, and on the asymmetries (production, detection and tagging efficiency). In the theoretical case of no asymmetries, zero mistag and 100% tagging efficiency, these effective coefficients become the physical coefficients as described in Sec. 3.5. The OS and SS taggers are combined on the fly during the time fit. All the steps to build the final PDF are described in details in the following. The tagging efficiency ϵ^i_{tag} of the ith tagger is corrected for the tagging efficiency asymmetry $\Delta\epsilon^i_{tag}$ as follows:

$$\begin{aligned} \epsilon^i_{B^0} &= \epsilon^i_{tag} - \frac{1}{2}\Delta\epsilon_{tag}, \\ \epsilon^i_{\overline{B}^0} &= \epsilon^i_{tag} + \frac{1}{2}\Delta\epsilon_{tag}, \end{aligned} \tag{7.14}$$

where $\epsilon^i_{B^0}$ and $\epsilon^i_{\overline{B}^0}$ are the two different tagging efficiencies for B^0 and $\overline{B}^0$. Moreover, the OS and SS mistag are calibrated taking into account asymmetries between B^0 and $\overline{B}^0$ according to Eq. (5.5). They will be referred in the following as $\omega^i_{B^0}$ and $\omega^i_{\overline{B}^0}$. It is also convenient to define the quantity $\Delta^{\pm}(\vec{d})$, according to the OS and SS tagging decisions. If $d_{OS} = d_{SS} = 0$, the B^0 candidate is untagged and the $\Delta^{\pm}(\vec{d})$ term becomes

$$\Delta^{\pm} = (1 - \epsilon^{OS}_{B^0} - \epsilon^{SS}_{B^0} + \epsilon^{OS}_{B^0}\epsilon^{SS}_{B^0}) \pm (1 - \epsilon^{OS}_{\overline{B}^0} - \epsilon^{SS}_{\overline{B}^0} + \epsilon^{OS}_{\overline{B}^0}\epsilon^{SS}_{\overline{B}^0}). \tag{7.15}$$

If $d_i = 0$, $d_j \neq 0$ and $i \neq j$, it means that the B^0 candidate is tagged by one tagger only, so that

$$\begin{aligned}\Delta^{\pm} = &\frac{1}{2}\Big[1-\epsilon^{j}_{B^0} + d_j\Big(1-\epsilon^{i}_{B^0} - 2\omega^{j}_{B^0} + 2\omega^{j}_{B^0}\epsilon^{i}_{B^0}\Big)\Big] \\ \pm &\frac{1}{2}\Big[1-\epsilon^{j}_{\overline{B}^0} + d_j\Big(1-\epsilon^{i}_{\overline{B}^0} - 2\omega^{j}_{\overline{B}^0} + 2\omega^{j}_{\overline{B}^0}\epsilon^{i}_{\overline{B}^0}\Big)\Big].\end{aligned} \tag{7.16}$$

Finally if $d_{OS} = d_{SS} = 1$ it means that the B^0 candidate is tagged by both taggers and $\Delta^{\pm}$ becomes:

$$\begin{aligned}\Delta^{\pm} = &\frac{1}{4}\epsilon^{SS}_{B^0}\epsilon^{OS}_{B^0}\Big[1 + d_{SS}\Big(1-2\omega^{SS}_{B^0}\Big) + d_{OS}\Big(1-2\omega^{OS}_{B^0}\Big) \\ &+ d_{\mathrm{OS}}d_{\mathrm{SS}}\Big(1-2\omega^{SS}_{B^0} - 2\omega^{OS}_{B^0} + 4\omega^{SS}_{B^0}\omega^{OS}_{B^0}\Big)\Big] \\ \pm &\frac{1}{4}\epsilon^{SS}_{\overline{B}^0}\epsilon^{OS}_{\overline{B}^0}\Big[1 + d_{SS}\Big(1-2\omega^{SS}_{\overline{B}^0}\Big) + d_{OS}\Big(1-2\omega^{OS}_{\overline{B}^0}\Big) \\ &+ d_{\mathrm{OS}}d_{\mathrm{SS}}\Big(1-2\omega^{SS}_{\overline{B}^0} - 2\omega^{OS}_{\overline{B}^0} + 4\omega^{SS}_{\overline{B}^0}\omega^{OS}_{\overline{B}^0}\Big)\Big].\end{aligned} \tag{7.17}$$

The final effective *CP* coefficients to be used in the time PDF to data can be be then written as:

$$C^{eff}_{\sin,\cos} = \begin{cases}(1+A_{det})C^{phys}_{\sin,\cos}(\Delta^- - A_{prod}\Delta^+), & f = D^{*-}D^+ \\ -(1-A_{det})C^{phys}_{\sin,\cos}(\Delta^- - A_{prod}\Delta^+), & f = D^{*+}D^-\end{cases} \tag{7.18}$$

$$C^{eff}_{\sinh,\cosh} = \begin{cases}(1+A_{det})C^{phys}_{\sinh}(\Delta^+ - A_{prod}\Delta^-), & f = D^{*-}D^+ \\ (1-A_{det})C^{phys}_{\sinh}(\Delta^+ - A_{prod}\Delta^-), & f = D^{*+}D^-\end{cases} \tag{7.19}$$

7.3.1 Decay time resolution

As already described in Sec. 6.3, despite the small effect of the decay time resolution on CP observables, it is worthwhile to include this effect in the decay time fit. To determine the decay time resolution model, an unbinned maximum likelihood fit is performed to the simulated events where the difference between true and reconstructed decay time is below 0.6 ps. To account for different sources introducing the decay time resolution an effective model consisting of three Gaussians is used. The complete parametrization of the resolution model is given in Eq. (7.20).

$$\begin{aligned}\mathcal{R}(t-t_{\mathrm{true}}) = &f_3\frac{1}{\sqrt{2\pi}\sigma_3}\exp\left(-\frac{(t-t_{\mathrm{true}}-\mu_t)^2}{2\sigma_3^2}\right) + f_2\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(t-t_{\mathrm{true}}-\mu_t)^2}{2\sigma_2^2}\right) \\ &+ (1-f_3-f_2)\frac{1}{\sqrt{2\pi}\sigma_3}\exp\left(-\frac{(t-t_{\mathrm{true}}-\mu_t)^2}{2\sigma_3^2}\right).\end{aligned} \tag{7.20}$$

The results of the fit are shown in Table 7.16 for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ and $K^-\pi^+$ respectively, while the corresponding plots are presented in Fig. 7.20.

Parameter	Value		Error
μ	-0.00068	±	0.00031 ps
f_3	0.683	±	0.041
f_2	0.0075	±	0.0027
σ_3	0.03895	±	0.00097 ps
σ_2	0.203	±	0.026 ps
σ_1	0.0720	±	0.0029 ps

Parameter	Value		Error
μ	0.00075	±	0.00016 ps
f_3	0.055	±	0.009
f_2	0.51	±	0.04
σ_3	0.124	±	0.005 ps
σ_2	0.0368	±	0.0008 ps
σ_1	0.0632	±	0.0020 ps

Table 7.16: Fit parameters of the decay time resolution function for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ (left) and $K^-\pi^+$ (right).

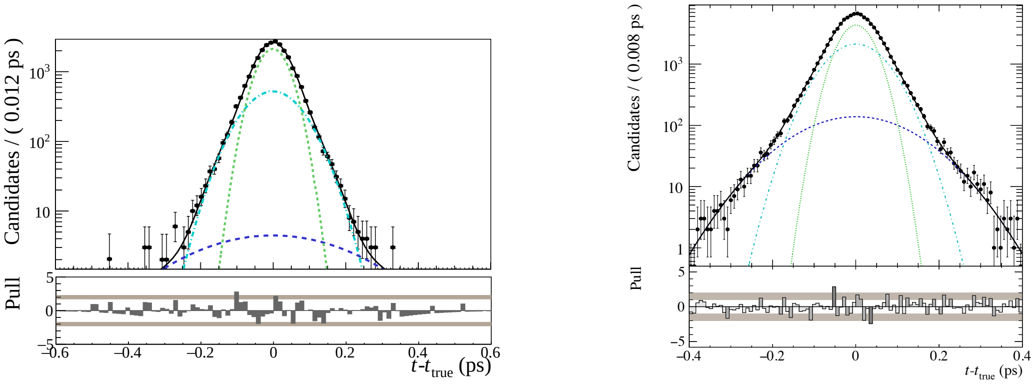


Figure 7.20: Fits of resolution model to the difference between true and reconstructed decay time in signal MC for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ (left) and $K^-\pi^+$ (right). The black solid line is the projection of the full PDF. The blue dashed, the green dotted and the turquoise dashed-dotted lines represent the three components.

7.3.2 Decay time acceptance

As described in Sec. 6.3, in order to correctly describe effects related to the different selection steps, in the fit to the decay time, a decay time acceptance model has to be taken into account. A data-driven approach is employed. In the fit to data the B^0 lifetime is constrained to its PDG value of $\tau = (1.519 \pm 0.005)\,\mathrm{ps}$ [28]. Any deviation from the pure exponential behavior is described using a cubic spline model. The knot positions for the splines are tuned on MC sample, they are chosen to be at $\{0.3, 0.5, 2.7, 6.3, 10.3\,\mathrm{ps}\}$, and the second to last spline parameter is fixed to 1.0 for normalization.

Exploiting the truth information, available in MC sample, the shape of the decay time acceptance can be separated from the pure exponential decay. This shape can be used to make a comparison with the spline configuration described previously, showing that it is the one having the minimum number of knots necessary to provide a good fit both on MC and data. The plots obtained from MC samples, for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ and $K^-\pi^+$ are shown in Fig. 7.21.

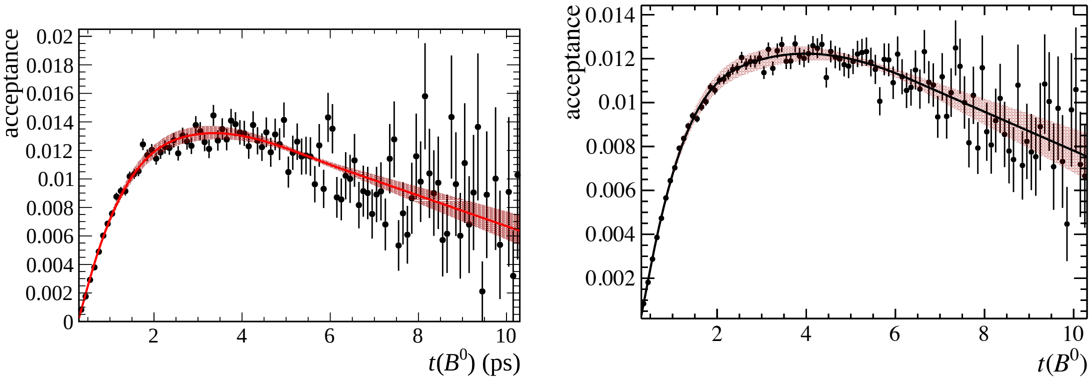


Figure 7.21: Decay time acceptance of truth-matched signal MC for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ (left) and the $K^-\pi^+$ final state (right). The black data points show the true decay time acceptance calculated dividing the reconstructed by the true decay time distribution. The red line is the spline acceptance function with five knots, while the red stripes indicate the $1\,\sigma$ error band taking into account the statistical uncertainties.

The studies performed on MC samples showed that there is a very small correlation between spline coefficients and other parameters. Moreover in spite of the usage of two different BDT in the selection for D^0 final states, the same acceptance configuration is valid for both. For this reason in the decay time fit to data the same parametrization will be employed.

7.4 Results

7.4.1 Fitter Validation

In order to validate the fit model, a fit on signal MC is performed and the results are compared to the generation values of $S_{D^*D}^{\mathrm{Gen}} = 0.73$, $\Delta S_{D^*D}^{\mathrm{Gen}} = 0$, $C_{D^*D}^{\mathrm{Gen}} = 0$, $\Delta C_{D^*D}^{\mathrm{Gen}} = 0$.

Fig. 7.22 shows the results for the decay time fit, performed on D^0 decaying into $K^-\pi^-\pi^+\pi^+$ and $K^-\pi^+$ signal MC respectively.

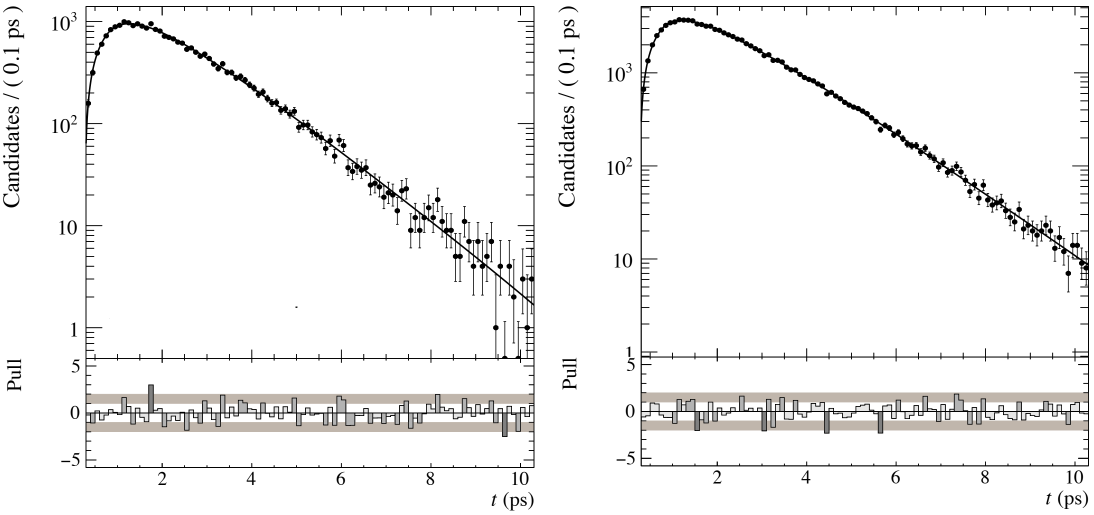


Figure 7.22: Plot of the decay time distributions of MC sample with the projection of the PDF and the pull distribution for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ (left) and $K^-\pi^+$ (right).

In the fit to MC sample the production and detection asymmetries are set to zero, while in the nominal fit to data they are included according to what described in Sec. 7.4.3 and truth information on B^0 flavour is exploited. In Table 7.17 the results are listed, where it can be seen that *CP* parameters are compatible with values used in generation.

	Generation	Fit
S_{D^*D}	0.73	0.725± 0.007
ΔS_{D^*D}	0.	0.002± 0.007
C_{D^*D}	0.	-0.010± 0.010
ΔC_{D^*D}	0.	-0.003± 0.010

	Generation	Fit
S_{D^*D}	0.73	0.737±0.004
ΔS_{D^*D}	0.	-0.002±0.004
C_{D^*D}	0.	-0.010± 0.005
ΔC_{D^*D}	0.	-0.001± 0.005

Table 7.17: Results of the decay time fit on MC samples for D^0 decaying into $K^-\pi^-\pi^+\pi^+$ (left) and $K^-\pi^+$ (right).

7.4.2 Blinding

Before performing the fit, a blinding transformation is applied on the *CP* parameters S_f, $S_{\bar{f}}$, C_f and $C_{\bar{f}}$, where f=$D^{*+}D^-$ and $\bar{f}$=$D^{*-}D^+$. The blinded offset is drawn from a uniform distribution with a range $\pm s$, using a random number generator with known seed, called *blinding string*. In Table 7.18 the blinding strings and s parameters used in the fit are listed.

Table 7.18: Blinding strings and scales s for RooUnblindUniform.

Parameter	Blinding string	s
S_f	SB02DstD5fb	2.0
$S_{\bar{f}}$	SbarB02DstD5fb	2.0
C_f	CB02DstD5fb	2.0
$C_{\bar{f}}$	CbarB02DstD5fb	2.0

7.4.3 Input Parameters

The decay time resolution parameters (given in Table 7.16), the flavour tagging parameters $\langle\eta^{OS}\rangle = 0.3675$ and $\langle\eta^{SS}\rangle = 0.4301$ (which are averaged from the $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*+} D_s^-$ calibration channels) and the B^0 lifetime difference $\Delta\Gamma = 0$ ps^{-1}, are fixed in the nominal decay time fit.

Other parameters instead are introduced in the fit through a Gaussian constraint, in order to account for their associated uncertainties, as described by Eq. (6.18). The constrained parameters with their uncertainties are summarized in Table 7.19. The value of the production asymmetry is obtained from the LHCb measurement [91], applying a weighting procedure based on transverse momentum and pseudorapidity, in order to match the kinematics of the nominal data set. As an example the p_T dependent values of production asymmetry used to recompute the integrated values on $B^0 \to D^{*\pm} D^{\mp}$ MC are shown in Fig. 7.23, from them the weighted averaged is calculated. The tagging efficiency asymmetries are also used as constrained parameters in the fit, their values are obtained, for each tagger, by fitting the decay time distribution of control channels. According to what described in Sec. 7.3, in order to extract through the decay time fit, the value for

Table 7.19: Parameters that are constrained in the fit to $B \to D^{*+} D^-$ data.

Parameter	Value and uncertainty
A_{prod} [%]	-0.14±0.68
$\Delta\epsilon_{\text{tag}}^{\text{OS}}$ [%]	1.50±1.36
$\Delta\epsilon_{\text{tag}}^{\text{SS}}$ [%]	-0.46±0.71
p_0^{OS}	0.3679±0.0071
p_0^{SS}	0.4288±0.0056
p_1^{OS}	1.107±0.068
p_1^{SS}	0.927±0.092
Δp_0^{OS}	0.013±0.010
Δp_0^{SS}	-0.0037±0.0079
Δp_1^{OS}	0.0226±0.0993
Δp_1^{SS}	-0.0042±0.1288

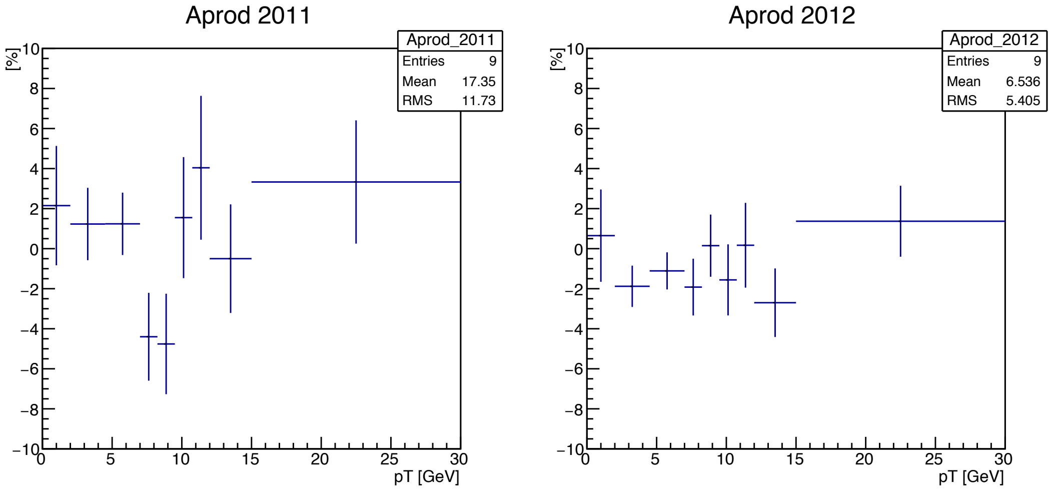


Figure 7.23: Production asymmetry as a function of transverse momentum for $B^0 \rightarrow D^{*\pm}D^{\mp}$. The weighted average of the integrated values obtained is used as input in the nominal decay time fit.

Table 7.20: Study on the influence of the asymmetries using MC sample of D^0 decaying into $K^-\pi^-\pi^+\pi^+$.

Parameter	A_{det}, $A_{prod}+\Delta\epsilon_{tag}$ fixed	$A_{det} free$	A_{det}, $A_{prod}+\Delta\epsilon_{tag}$ free
S_{D^*D}	0.726±0.007	0.726±0.007	0.715±0.013
ΔS_{D^*D}	-0.003±0.007	-0.003±0.007	-0.003±0.007
C_{D^*D}	-0.012±0.010	-0.012±0.010	-0.006±0.011
ΔC_{D^*D}	0.001±0.010	0.001±0.010	0.001±0.010
A_{det} [%]	0	0.25±0.63	0.26±0.63
$A_{prod}+\Delta\epsilon_{tag}$ [%]	0	0	1.8±1.6

A_{D^*D} along with those of *CP* observables, a constraint on the detection asymmetry should be considered. This value is computed from signal MC and used to extract A_{D^*D}, from the term (A_{D^*D} + A_{det}) that appears in the nominal PDF. From studies performed on MC sample (see Table 7.20) it has been observed that the tagging asymmetry along with the production asymmetry strongly limit the precision of the fitted *CP* observables, while the detection asymmetry term appear to be uncorrelated to any other parameter, for this reason in the final fit configuration the latter is left free while the others are constrained.

7.4.4 Result perspectives

The fit model is currently under validation using pseudo-experiments. A preliminary estimate of the sensitivity that can be achieved through this measurement can be made starting from the result obtained in a previous LHCb analysis of $B^0 \to D^{*\pm}D^{\mp}$ considering only D^0 decaying into $K^-\pi^+$ and performed on Run1 sample [92]. The sensitivity found for the parameters S_{D^*D} and ΔS_{D^*D} is $\sigma_{S_{D^*D}} = 0.27$ and $\sigma_{\Delta S_{D^*D}} = 0.27$ respectively. This result was obtained from a CP fit on tagged events, where only OS standard combination was considered, and corresponding to a tagging power of 3.85%. In order to increase the sensitivity on CP parameters, the analysis presented in this work plans to include along with the $D^0 \to K^-\pi^+$ decay also the $D^0 \to K^-\pi^-\pi^+\pi^+$ one, this leads to an increase of about 50 % in statistics with respect to the previous $B^0 \to D^{*\pm}D^{\mp}$ analysis. For both the D^0 decays, the analysis aims to add on top of the decay $D^+ \to K^-\pi^+\pi^+$ also the $D^+ \to K^-K^+\pi^+$ with an expected gain in statistic around 10%. Moreover the full Run1 and Run2 data samples are exploited, the addition of the Run2 data sample, bring to a gain of about 85% in statistics. Finally, adding to the OS standard combination used in the previous analysis, the new SS taggers developed by the LHCb collaboration could allow to reach a similar tagging power to what obtained in $B^0 \to D^+D^-$ analysis (Sec. 5.6.2), doubling the value reached in the previous $B^0 \to D^{*\pm}D^{\mp}$ analysis, and corresponding to an increase in statistics of about 100%. The total gain coming from these improvements, corresponds to a factor of about 9 in statistics, the expected sensitivity for CP parameters results to be

$$\sigma_{S_{D^*D}} = 0.09, \qquad \sigma_{\Delta S_{D^*D}} = 0.09.$$

7.5 Systematic uncertainties

Different sources of systematics will be taken into account in $B^0 \to D^{*\pm}D^{\mp}$ analysis. At first in order to estimate the size of systematic effects the full sample will be divided into categories, according to the year of data taking, the magnet polarity, the D^0 final state and the tagging algorithms. Individual studies will be performed using pseudo-experiment to estimate the effect of an inaccurate determination of the resolution parameters, and of fixing the parameters as the production asymmetry, the decay width difference and the mass difference in the decay time fit. Pseudo-experiments will be used also to evaluate the bias due to the likelihood fit. The systematic uncertainties associated to the mass fit will be studied using a different parametrization for the mass components.

Chapter 8

Conclusion

8.1 Result Discussion for $B^0 \to D^+D^-$

A measurement of the CP observables S and C in the decay channel $B^0 \to D^+D^-$ has been performed. The full Run 1 data set which corresponds to a total integrated luminosity of 3 fb^{-1} of pp collisions has been used. To separate signal from combinatorial background a multivariate classifier has been trained. This has been optimized with respect to the statistical sensitivity of the CP observables. The sPlot technique has been applied to calculate signal sWeights based on a fit to the invariant mass distribution. The total signal yield has been determined to be 1610 ± 49. Using the OS and for the first time the combination of the SSπ (BDT) and SSp taggers, an effective tagging efficiency of (8.1±0.6)% is achieved. The decay time distribution of the flavour-specific decay channel $B^0 \to D_s^+D^-$ is fitted in order to estimate the values of the flavour tagging calibration parameters. These are taken as input into the signal PDF, which is used to perform the likelihood fit of the tagged decay time distribution of the $B^0 \to D^+D^-$ candidates, the obtained result for the CP observables is

$$S_{D^+D^-} = -0.54^{+0.17}_{-0.16}(stat) \pm 0.05(syst),$$

$$C_{D^+D^-} = 0.26^{+0.18}_{-0.17}(stat) \pm 0.02(syst),$$

with a statistical correlation coefficient of $\rho(S_{D^+D^-}, C_{D^+D^-}) = 0.48$. The largest systematic uncertainty comes from neglecting partially charmless contributions in the mass fit. This result excludes the conservation of CP symmetry by 4.0 standard deviations. In Fig. 8.1 the results from BaBar, Belle and LHCb, along with their combination, are presented in the two-dimensional plane ($C_{D^+D^-}$, $S_{D^+D^-}$). When comparing the uncertainty ellipses it is clear that the precision of LHCb matches the one of Belle, while it is significantly better than the one of BaBar. Moreover the orientation of the ellipses shows that in the measurements of the B-factories the two CP observables result to be almost uncorrelated. In order to obtain a comparison between the central values it is useful to take into account the condition $S_{D^+D^-}{}^2+C_{D^+D^-}{}^2=1$, which defines an arc of circumference in the plane that delimits the physically allowed region. The result by Belle [5], of $S_{D^+D^-}$ =-1.06 and

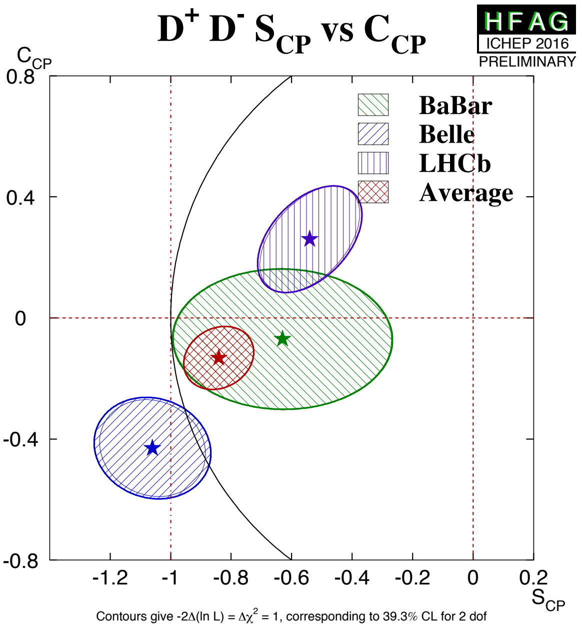


Figure 8.1: Comparison of CP observables from $B^0 \to D^+D^-$ decay in ($S_{D^+D^-}$,$C_{D^+D^-}$) plane. The black arc is defined by the condition $S_{D^+D^-}\ {}^2$+$C_{D^+D^-}\ {}^2$=1.

$C_{D^+D^-}$ =-0.43 lies outside of this region, while the results of BaBar [4], $S_{D^+D^-}$ =-0.63 and $C_{D^+D^-}$ =-0.07, and the one obtained from LHCb, $S_{D^+D^-}$ =-0.54 and $C_{D^+D^-}$ =0.26 are inside. LHCb measurement is compatible with the one obtained by BaBar, while a proper evaluation of the compatibility with the result from the Belle experiment can not be performed due to its non-Gaussian uncertainties [80]. The CP observables S and C are related to the B^0 mixing angle ϕ_d and a phase shift $\Delta\phi_d$ from the decay amplitude via Eq. (8.1)

$$\sin(\phi_d + \Delta\phi_d) = -\frac{S}{\sqrt{1-C^2}}. \tag{8.1}$$

In the SM $\phi_d = 2\beta$, and if the $B^0 \to D^+D^-$ decay amplitude can be described by a dominant tree-level $b \to c\bar{c}d$ transition, the phase shift $\Delta\phi$ vanishes and the CP observables are given by C=0 and S=-sin ϕ_d. Instead, if contribution of higher-order Standard Model corrections exists, the phase shift $\Delta\phi_d$ can be obtained, complicating the determination of β. The fit results of $S_{D^+D^-}$ and $C_{D^+D^-}$ corresponds to

$$\sin(\phi_d + \Delta\phi_d) = -\frac{S_{D^+D^-}}{\sqrt{1-C^2_{D^+D^-}}} = 0.56^{+0.16}_{-0.17}, \tag{8.2}$$

here the statistical uncertainty is estimated by generating three million sets of $S_{D^+D^-}$ and $C_{D^+D^-}$ using a two-dimensional Gaussian distribution including their correlation, calculating $\sin(\phi_d + \Delta\phi_d)$ for each of them, and then taking the two-sided 68% confidence intervals. In order to extract the phase shift $\Delta\phi_d$, the measurement of $\sin(2\beta) = 0.691 \pm 0.017$ coming from [7] is used as input, resulting in

$$\Delta\phi_d = -0.16^{+0.19}_{-0.21} rad,$$

which represents the world's most precise determination of this quantity. With this measurement the room for possible higher order corrections got smaller by a factor three with respect to what was allowed with previous measurements [35].

8.2 Result Discussion for $B^0 \to D^{*\pm}D^{\mp}$

An analysis for the measurement of the CP observables S_{D^*D}, ΔS_{D^*D}, C_{D^*D} and ΔC_{D^*D} in the decay channel $B^0 \to D^{*\pm}D^{\mp}$ has been presented. The full Run 1 and Run2 data sets, corresponding to a total integrated luminosity of 5 fb^{-1} of pp collisions have been used. In order to separate signal from combinatorial background a multivariate classifier has been employed. This has been optimized making use of a figure of merit, that takes into account the tagging power, the signal significance, the decay time resolution and includes also a term sensitive to $\sin 2\phi_d^{eff}$. The sPlot technique has been applied to calculate signal sWeights based on a fit to the invariant mass distribution. The total signal yield was found to be 2722 ± 59. The analysis aims to use the OS along with the combination of the SSπ (BDT) and SSp taggers. The fit model to the decay time has been tested on MC samples and further validation is ongoing using pseudo-experiment. The decay channels $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*+}D_s^-$ have been used in order to estimate the values of the flavour tagging calibration parameters. The weighted average of these parameters coming from the two decays will be used as input into the signal PDF, which is used to perform the likelihood fit of the decay time distribution to the $B^0 \to D^{*\pm}D^{\mp}$ candidates. An initial estimate of the sensitivity that can be achieved through this measurement can be made starting from the result obtained in a previous LHCb analysis of $B^0 \to D^{*\pm}D^{\mp}$, done using only the decay $D^0 \to K^-\pi^+$ on Run 1 sample, considering the gain in statistics coming from adding another D^0 decay i.e. $D^0 \to K^-\pi^-\pi^+\pi^+$, along with the two final states $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K^-K^+\pi^+$, including the Run 2 statistics, and exploiting the full combination of the available taggers.The sensitivity on CP observables is expected to be

$$\sigma_{S_{D^*D}} = 0.09, \qquad \sigma_{\Delta S_{D^*D}} = 0.09.$$

Bibliography

[1] Y. Fukuda et al. Evidence for Oscillation of Atmospheric Neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998.

[2] A. Sakharov. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *Soviet Physics Uspekhi*, 34(5):392, 1991.

[3] O. Brüning et al. *LHC Design Report*, volume v.1 : the LHC Main Ring of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2004.

[4] B. Aubert et al. Measurements of time-dependent CP asymmetries in $B^0 \to D^{(*)+}D^{(*)-}$ decays. *Phys. Rev. D*, 79:032002, Feb 2009.

[5] M. Röhrken et al. Measurements of branching fractions and time-dependent CP violating asymmetries in $B^0 \to D^{(*)\pm}D^{\mp}$ decays. *Phys. Rev. D*, 85:091106, May 2012.

[6] Y. Amhis et al. Heavy Flavor Averaging Group. `http://www.slac.stanford.edu/xorg/hfag/`.

[7] Heavy Flavor Averaging Group (HFAG). Results on Time-Dependent CP Violation, and Measurements Related to the Angles of the Unitarity Triangle: Winter 2016 (Moriond 2016, Italy, etc.). http://www.slac.stanford.edu/xorg/hflav/triangle/moriond2016/.

[8] D. Griffiths. *Introduction to Elementary Particles.* Wiley, 2008.

[9] D. Perkins. *Introduction to High Energy Physics.* Cambridge University Press, 1982.

[10] W. Kenneth. Confinement of Quarks. *Phys. Rev. D*, 10:2445–2459, Oct 1974.

[11] R. Aaij et al. Observation of $J/\psi\phi$ Structures Consistent with Exotic States from Amplitude Analysis of $B^+ \to J/\psi\phi K^+$ Decays. *Phys. Rev. Lett.*, 118:022003, Jan 2017.

[12] R. Aaij et al. Amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays. *Phys. Rev. D*, 95:012002, Jan 2017.

[13] R. Aaij et al. Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays. *Phys. Rev. Lett.*, 115:072001, Aug 2015.

[14] R. Aaij et al. Evidence for exotic hadron contributions to $\Lambda_b^0 \rightarrow J/\psi p \pi^-$ decays. *Phys. Rev. Lett.*, 117(8):082003, 2016. [Addendum: Phys. Rev. Lett.118,119901(2017)].

[15] P. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[16] K. Zuber. *Neutrino Physics, Second Edition.* Series in High Energy Physics, Cosmology and Gravitation. Taylor & Francis, 2011.

[17] A. Bettini. *Introduction to Elementary Particle Physics.* Cambridge University Press, 2008.

[18] H. Georgi et al. Unity of All Elementary-Particle Forces. *Phys. Rev. Lett.*, 32:438–441, Feb 1974.

[19] A. Georges et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015.

[20] J. Gervais et al. Field theory interpretation of supergauges in dual models. *Nuclear Physics B*, 34(2):632 – 639, 1971.

[21] Y. Golfand et al. Extension of the Algebra of Poincare Group Generators and Violation of P invariance. *JETP Lett. 13*, 323(13), 1971.

[22] D. Volkov et al. Is the neutrino a goldstone particle? *Physics Letters B*, 46(1):109 – 110, 1973.

[23] D. Volkov et al. Possible universal neutrino interaction. *JETP Lett.*, 16(11), 1972.

[24] Q. Ahmad et al. Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301, Jul 2001.

[25] Q. Ahmad et al. Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, Jun 2002.

[26] M. Kobayashi et al. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, 1973.

[27] C. Jarlskog. A basis independent formulation of the connection between quark mass matrices, CP violation and experiment. *Zeitschrift für Physik C Particles and Fields*, 29(3):491–497, Sep 1985.

[28] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[29] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945, 1983.

[30] The CKM Fitter Group. `http://ckmfitter.in2p3.fr`.

[31] G. Eigen et al. Global CKM Fits with the Scan Method. *Phys. Rev.*, D89(3):033004, 2014.

[32] The UT Fitter Group. `http://utfit.org/UTfit/WebHome`.

[33] M. Ciuchini et al. Unitarity triangle analysis in the standard model and sensitivity to new physics. *eConf*, C0304052:WG306, 2003.

[34] D. Boutigny et al. The BABAR physics book: Physics at an asymmetric B factory. In *Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting) Pasadena, California, September 22-24, 1997*, 1998.

[35] L. Bel et al. Anatomy of $B \to D\overline{D}$ decays. *JHEP*, 07:108, 2015.

[36] O. Brüning et al. *LHC Design Report*, volume v.2 : the LHC Infrastructure and General Services of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2004.

[37] M. Benedikt et al. *LHC Design Report*, volume v.3 : the LHC Injector Chain of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2004.

[38] S. Myers. *The LEP Collider, from design to approval and commissioning*. John Adams' Lecture. CERN, Geneva, 1991. Delivered at CERN, 26 Nov 1990.

[39] Symmetry dimensions of particle physics. `http://www.symmetrymagazine.org`.

[40] F. Marcastel. CERN's Accelerator Complex. La chane des acclrateurs du CERN, Oct 2013. General Photo.

[41] International Science Grid. `http://archive.isgtw.org/?pid=1000416`.

[42] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *JINST*, 3(08):S08002, 2008.

[43] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3(08):S08003, 2008.

[44] The CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3(08):S08004, 2008.

[45] The TOTEM Collaboration. The TOTEM Experiment at the CERN Large Hadron Collider. *JINST*, 3(08):S08007, 2008.

[46] The LHCf Collaboration. The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3(08):S08006, 2008.

[47] B. Acharya et al. The physics programme of the MoEDAL experiment at the LHC. *International Journal of Modern Physics A*, 29(23):1430050, 2014.

[48] The LHCb Collaboration. LHCb detector performance. *International Journal of Modern Physics A*, 30(07):1530022, 2015.

[49] The LHCb Collaboration. LHCb Operations Plots Webpage. `http://lhcb-operationsplots.web.cern.ch/lhcb-operationsplots/index.htm`.

[50] The LHCb Collaboration. The LHCb Detector at the LHC. *JINST*, 3(08):S08005, 2008.

[51] J. Harrison. *Radiation damage studies in the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays.* PhD thesis, Manchester University, 2014. Presented 16 05 2014.

[52] M. Adinolfi et al. Performance of the LHCb RICH detector at the LHC. *EPJ C*, 73(5):2431, 2013.

[53] R. Aaij et al. The LHCb trigger and its performance in 2011. *JINST*, 8(04):P04022, 2013.

[54] R. Aaij et al. Measurement of the track reconstruction efficiency at LHCb. *JINST*, 10(02):P02007, 2015.

[55] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262(2):444 – 450, 1987.

[56] R. Aaij et al. The DaVinci project. `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/`, 2017. [Online].

[57] T. Sjöstrand et al. PYTHIA 6.4 physics and manual. *JHEP*, 2006(05):026, 2006.

[58] T. Sjöstrand et al. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852 – 867, 2008.

[59] D. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462(1):152 – 155, 2001. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[60] P. Golonka et al. PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays. *EPJ C - Particles and Fields*, 45(1):97–107, Jan 2006.

[61] S. Agostinelli et al. Geant4a simulation toolkit. *Nucl. Instrum. and Meth.*, A506(3):250 – 303, 2003.

[62] J. Allison et al. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1):270–278, Feb 2006.

[63] R. Aaij et al. The Boole project. `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/`, 2017. [Online].

[64] R. Aaij et al. The Moore project. `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/`, 2017. [Online].

[65] R. Aaij et al. The Brunel project. `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/`, 2017. [Online].

[66] R. Aaij et al. New algorithms for identifying the flavour of B^0 mesons using pions and protons. *EPJ C*, 77:238, 2017.

[67] R. Aaij et al. Opposite-side flavour tagging of B mesons at the LHCb experiment. *EPJ C*, 72(6):2022, Jun 2012.

[68] The LHCb Collaboration. Optimization and calibration of the LHCb flavour tagging performance using 2010 data, Mar 2011. LHCb-ANA-2011-003.

[69] M. Grabalosa. *Flavour Tagging developments within the LHCb experiment.* PhD thesis, Barcelona University, 2012-03-28.

[70] J. Wimberley et al. Espresso Performance Monitor. `https://gitlab.cern.ch/lhcb-ft/EspressoPerformanceMonitor`, 2017. [Online].

[71] P. Koppenburg. Dealing with multiple candidates - an update. Technical Report LHCb-INT-2016-045. CERN-LHCb-INT-2016-045, CERN, Geneva, Nov 2016.

[72] M. Pivk et al. SPlot: A Statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, A555:356–369, 2005.

[73] W. Verkerke et al. *RooFit.* `http://roofit.sourceforge.net`.

[74] T. Skwarnicki. *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances.* PhD thesis, Institute of Nuclear Physics, Krakow, 1986. DESY-F31-86-02.

[75] R. Aaij et al. Selections and lifetime measurements for exclusive $b \to J/\psi X$ decays with $J/\psi \to \mu^+\mu^-$ with 2010 data, Apr 2011. LHCb-CONF-2011-001.

[76] R. Aaij et al. Measurement of the $\overline{B}^0$–B^0 and $\overline{B}^0_s$–B^0_s production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B739:218, 2014.

[77] C. Bozzi et al. Measurement of the semileptonic CP violating asymmetry a^d_{sl}, 2013. LHCb-ANA-2013-050.

[78] P. Seyfert et al. Flavour Tagging Combinations: Definitions and Best Practices. Technical Report LHCb-INT-2016-008. CERN-LHCb-INT-2016-008, CERN, Geneva, Feb 2016.

[79] F. Meier et al. Measurement of *CP* violation in $B^0 \to D^{*\pm}D^{\mp}$ decays, May 2017. LHCb-ANA-2017-031.

[80] R. Aaij et al. Measurement of *CP* Violation in $B^0 \to D^+D^-$ Decays. *Phys. Rev. Lett.*, 117:261801, Dec 2016.

[81] W. Hulsbergen. Decay chain fitting with a Kalman filter. *Nucl. Instrum. Meth.*, A552(3):566 – 575, 2005.

[82] C. Jones. Charged Particle ID. `https://indico.cern.ch/getFile.py/access?contribId=4&sessionId=0&resId=`. [Online].

[83] L. Breiman et al. *Classification and Regression Trees.* Wadsworth and Brooks, Monterey, CA, 1984. new edition.

[84] B. Roe et al. Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nucl. Instrum. Meth.*, A543(2):577 – 584, 2005.

[85] Y. Freund et al. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1):119 – 139, 1997.

[86] O Callot. FastVelo, a fast and efficient pattern recognition package for the Velo. Technical Report LHCb-PUB-2011-001. CERN-LHCb-PUB-2011-001, CERN, Geneva, Jan 2011. LHCb.

[87] R. Aaij et al. Measurement of the semileptonic *CP* asymmetry in $B^0 - \overline{B}^0$ mixing. *Phys. Rev. Lett.*, 114:041601, 2015.

[88] A. Poluektov. Correction of simulated particle identification response in LHCb using transformation of variables. Technical Report LHCb-INT-2017-007. CERN-LHCb-INT-2017-007, CERN, Geneva, Apr 2017.

[89] P. Mackowiak et al. Measurement of *CP* violation in $B^0 \to J/\psi K^0_S$ and $B^0 \to \psi(2S)K^0_S$ decays , Oct 2014. LHCb-ANA-2014-085.

[90] Y. Xie. Principles to optimize event selections for measurements of CP asymmetries. Technical Report LHCb-INT-2009-012. CERN-LHCb-INT-2009-012, CERN, Geneva, Jul 2009.

[91] R. Aaij et al. Measurement of B^0, B^0_s, B^+ and Λ^0_b production asymmetries in 7 and 8 TeV proton-proton collisions. *Phys. Lett.*, B774:139–158, 2017.

[92] M. Schellenberg. Messung der zerfallszeitabhängigen *CP*-Asymmetrie im Zerfall $B^0 \to D^{*\pm}D^{\mp}$ mit dem LHCb Experiment. Master's thesis, TU Dortmund University, Sep 2015.