Measuring the decay $B^0 \to \pi^- D_s^+$ with sensitivity for the $b \to u$ quark transition

Master Project
Physics and Astronomy
60 EC
research conducted between September 1, 2014 and June 26, 2015

Abstract

In this thesis, the measurement of $\mathcal{B}(B^0 \to \pi^- D_s^+)$ at LHCb is discussed. $B^0 \to \pi^- D_s^+$ is a non-leptonic tree decay, which only occurs via the $b \to u$ quark transition. As a result, it probes the CKM matrix element $|V_{ub}|$. As $B^0 \to \pi^- D_s^+$ is purely non-leptonic, non-factorisable contributions $|a_{NF}|$ from the strong interaction have to be taken into account. They are estimated using $B \to D\bar{D}$ decays while taking non-tree (penguin, exchange and penguin annihilation)contributions into account. It is found that $\mathcal{B}(B^0 \to \pi^- D_s^+) = (26.7 \pm 2.0(\text{stat.}) \pm 1.5(\text{syst.}) \pm 1.6(\mathcal{B})) \times 10^{-6}$, which results in $|V_{ub}||a_{NF}| = (3.40 \pm 0.14(\mathcal{B}) \pm 0.29(\text{ext.})) \times 10^{-3}$, where the first uncertainty comes from the measurement of $\mathcal{B}(B^0 \to \pi^- D_s^+)$ and the second uncertainty from external inputs.

Author: Mick Mulder (2092514)

PhD supervisor: Lennaert Bel

First supervisor: Niels Tuning

Second supervisor: Rick Bethlem

Contents

1 Introduction

Dirac postulated in 1928 the existence of matter and anti-matter [1]. Furthermore, a particle and its anti-particle will have the same mass, lifetime and equal but opposite charge. This is known as the CPT theorem, which is closely related to Lorentz invariance, implying that the laws of physics are the same in every inertial fram. The symmetry between particles and anti-particles would seem to imply that both matter and anti-matter are present in equal quantities in our universe. However, there is no experimental evidence for significant amounts of anti-matter in the universe. This would suggest one of two things: either our universe started with an asymmetry between the amount of matter and anti-matter, or some process produces more matter than anti-matter. Such a process is called baryogenesis.

In 1967, Sakharov proposed a set of three necessary conditions for baryogenesis [2]:

- Baryon number violation is obviously necessary for a process which needs to produce more baryons than anti-baryons.
- C and CP violation are also necessary, as otherwise equal amounts of left-handed matter and right-handed anti-matter would be produced. In this case, no asymmetry between the total amount of matter and anti-matter would appear.
- Interactions out of thermal equilibrium, as otherwise CPT symmetry would still lead to compensation of these processes.

In the current model of particle physics, called the Standard Model (SM), it is possible to accommodate CP violation, necessary for baryogenesis. However, the amount of observed CP violation seems to be too small. Therefore, searching for sources of CP violation might lead to the discovery of New Physics (NP), beyond the SM.

To explain the second condition for baryogenesis, the discrete symmetries C, P, and T first are introduced. C is the operation which exchanges a particle for its anti-particle, P inverts the spatial coordinates, and T reverses time. Up until 1956, it was thought that all of these operations were good symmetries of nature, when Wu [3] and Lederman [4] found this was not the case: P symmetry, also called parity, was discovered to be maximally violated in Co^{60} decays and $\pi^+ \to \mu^+ \bar{\nu}_\mu$ decays, both of which occur through the weak interaction. It turned out the weak interaction only couples to left-handed particles and right-handed anti-particles (where the handedness is defined by the spin orientation). Lederman found that CP symmetry was still conserved by also considering $\pi^- \to \mu^- \nu_\mu$, the charge-conjugated process, and discovering that the combination CP leads to the same experimental outcome.

However, in 1964 Cronin and Fitch investigated kaon decays and discovered CP violation [5]. Since then, CP violation has also been observed in B meson decays. Kaons and B mesons decay through the weak interaction, *i.e.* through the exchange of a W boson. The coupling strengths of the quark-to-quark transitions in the weak interaction are given by the Cabibbo-Kobayashi-Maskawa (CKM) matrix, as shown in Equation (1) [6, 7]. An example of one of these transitions is shown in Figure 1.

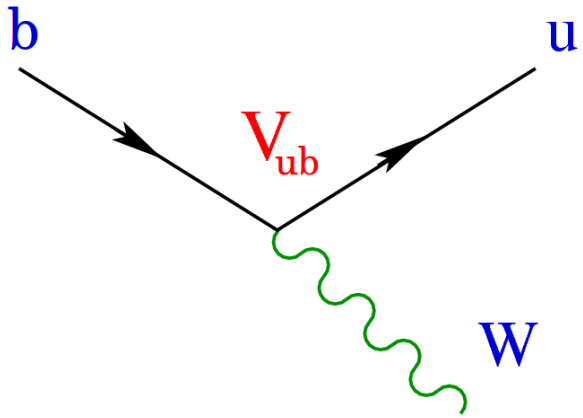

Figure 1: A diagram showing the $b \to u$ quark transition, as modified using Ref. [8].

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1}$$

The CKM matrix is a 3x3 unitary matrix which describes the mixing between the mass and weak interaction eigenstates of the three generations of down quarks, as well as quantifying the coupling strength of the mass eigenstates to the weak interaction eigenstates. Because of its unitarity, it contains four independent parameters: three real angles and a complex phase. This phase causes CP violation in the weak interaction.

A common parametrization of the CKM matrix is the so-called Wolfenstein parametrization [9]. It describes the CKM matrix using A, λ, ρ and η which are all real parameters. Using this parametrisation and the unitarity conditions, one can construct a triangle in the complex plane with an apex at (ρ, η), fixing A and λ to their well measured values. The specific relation used is:

$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0 \tag{2}$$

The resulting unitary triangle is obtained by taking the side with length $V_{cb}^* V_{cd}$ as a base and normalising this side to be of length 1. Many measurements of the magnitudes and relative phases of the CKM elements can be combined and tested for consistency. The current experimental situation is shown in Figure 2, as obtained by the CKMfitter group [10].

The CKM element V_{ub} describes the transition from bottom to up quarks and has the largest current experimental uncertainty. The main reason for this large uncertainty is an inconsistency in the presently available measurements. The work in this thesis might help to solve this long standing discrepancy on the magnitude of V_{ub}.

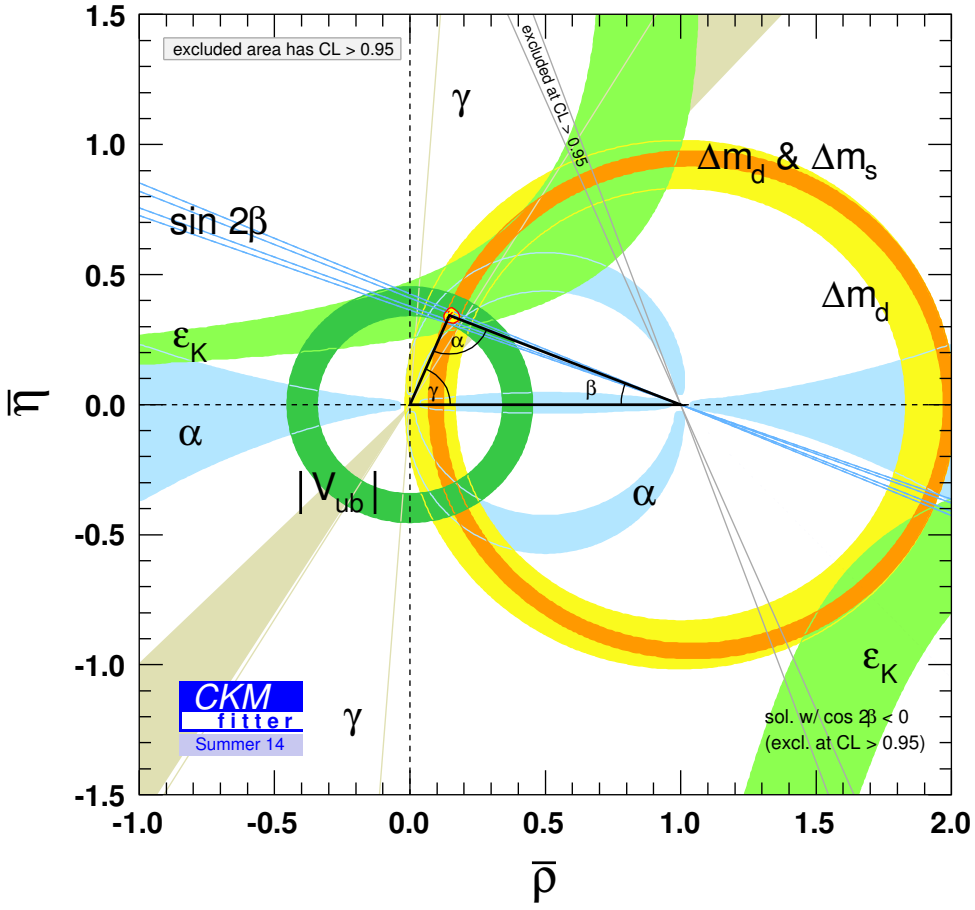

Figure 2: A plot showing the current CKM fit, as obtained from Ref. [10]. The red hashed region of the global combination corresponds to the 68% confidence level.

First, a possible method to determine $|V_{ub}|$ using the decay $B^0 \to \pi^- D_s^+$ is discussed in Section 2. Subsequently, the LHC and the LHCb detector are introduced in Section 3. Next, Sections 4 and 5 describe the analysis performed and systematic uncertainties, respectively. Subsequently, the branching fraction of $B^0 \to \pi^- D_s^+$ is determined in Section 6, relative to the well known decay $B^0 \to D^- \pi^+$. In addition, the implications for $|V_{ub}|$ are discussed. Finally, this thesis is concluded in Section 7.

2 Determining $|V_{ub}|$

2.1 Semileptonic decays

Currently, the magnitude of V_{ub} has been measured in two different ways: inclusively and exclusively, as discussed by the Particle Data Group (PDG) in Ref. [11]. Inclusive measurements are obtained with $\overline{B} \to X_u \ell \overline{\nu}_\ell$ decays, where X_u contains any number of hadrons resulting from the $b \to u$ transition. They are limited in a number of ways. First, it is difficult to measure the total decay rate due to the large $\overline{B} \to X_c \ell \overline{\nu}_\ell$ background, where X_c is any final state from the $b \to c$ transition. Secondly, some theoretical models need to be used to convert the partial decay rate into a measurement of V_{ub}. As a result, the PDG determines

$$|V_{ub}| = (4.41 \pm 0.15_{\text{exp}} \, {}^{+0.15}_{-0.17\text{theo}}) \ \times \ 10^{-3} (\text{inclusive}), \tag{3}$$

i.e. with a total uncertainty of 5%.

Exclusive measurements, using $B \to \pi \ell \overline{\nu}_\ell$, are a complementary means to determine $|V_{ub}|$. Experimentally, the specific final state ensures better background rejection. However, the smaller branching fraction ensures lower yields compared to inclusive decays. Theoretically, the calculation of the form factor $F(B \to \pi)$, which describes the effects of the strong interaction in the $B \to \pi$ transition, is the limiting factor. The main example of an exclusive decay is $B^0 \to \pi^- \ell^+ \overline{\nu}_\ell$, which is used by the PDG to obtain

$$|V_{ub}| = (3.28 \pm 0.29) \ \times \ 10^{-3} (\text{exclusive}), \tag{4}$$

which results in 9% uncertainty on $|V_{ub}|$. This determination of $|V_{ub}|$ is restricted by theoretical uncertainties. Clearly, there is a large disparity between the central values determined using both methods. Combining these two estimates, the PDG arrives at an average

$$|V_{ub}| = (4.13 \pm 0.49) \ \times \ 10^{-3} (\text{average}), \tag{5}$$

where the uncertainty has been scaled by a factor of 2.7 to reflect the discrepancy between the two values of $|V_{ub}|$. As a result, the average value $|V_{ub}|$ has a 12% uncertainty. All current measurements are shown in Figure 3.

To attempt to improve this unsatisfactory situation, this thesis will investigate the possible use of $\mathcal{B}(B^0 \to \pi D_s^+)$ to determine $|V_{ub}|$ exclusively.

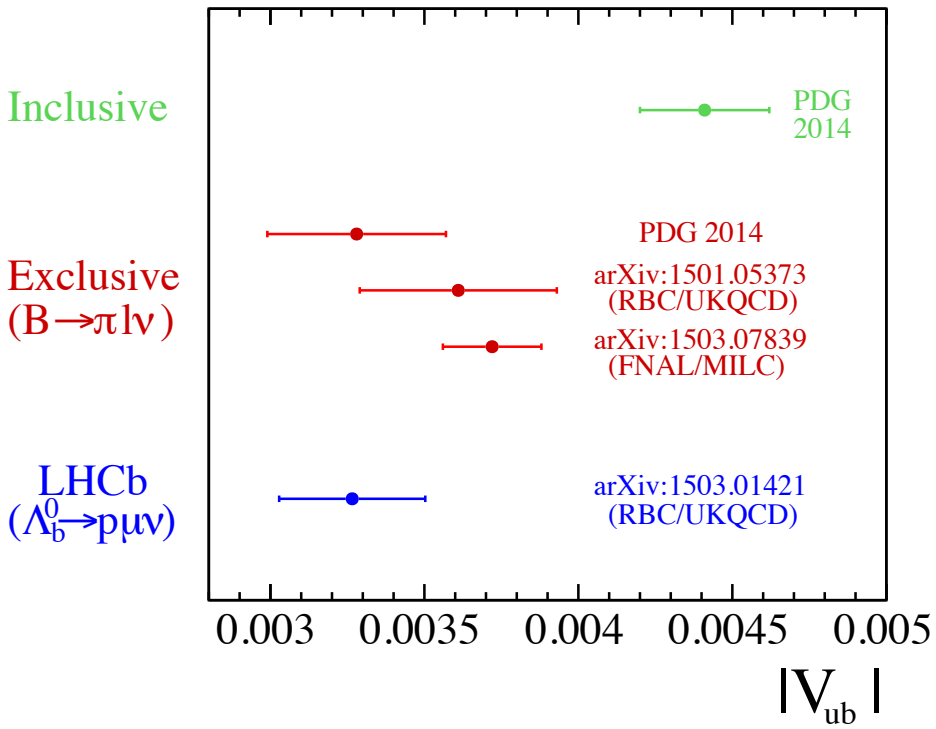

Figure 3: Recent measurements, along with the PDG 2014 averages, of $|V_{ub}|$. The different exclusive measurements (shown in red) use different assumptions for the $B \to \pi$ form factor. Recently, LHCb has published value for $|V_{ub}|$ using the exclusive baryonic decay $\Lambda_b^0 \to p\mu\nu$.

2.2 Properties of $B^0 \to \pi^- D_s^+$

The decay $B^0 \to \pi^- D_s^+$ can only occur through the simplest of transitions, a tree diagram, which is shown in Figure 4. The fact that no other diagrams have to be taken into account makes a measurement of the branching ratio of $B^0 \to \pi^- D_s^+$ easier to interpret in terms of physics parameters.

However, $B^0 \to \pi^- D_s^+$ is a non-leptonic B decay, as the final state contains two hadrons. This means that it is more complicated to investigate from a theoretical point of view. To calculate the amplitudes of the associated diagrams, a low energy effective Hamiltonian can be used. Therefore, assumptions are needed to determine a decay rate. In literature, the four-fermion matrix element, which describes the $\bar{b}$ quark decaying to c, $\bar{s}$ and $\bar{u}$ quarks, is often approximated as a multiplication of two currents: the $\bar{b} \to \bar{u}$ transition, and the creation of a D_s^+ meson from the vacuum. In other words, the possible exchange of gluons between the two final state hadrons is neglected. This approximation is called factorisation. Using this approximation, a general expression for the decay rate is obtained:

$$\Gamma(B \to LH) \sim G_F^2 |V_{BL}||V_H| F(B^0 \to L)^2|_{q^2=M_H^2} f_{Hs}^2 |a_{NF}^T(LH)|^2 \tag{6}$$

where L and H are the light and heavy hadron respectively, G_F is the Fermi constant, $|V_{BL}|$ and $|V_H|$ are the appropriate CKM elements, $F(B \to L)$ is the so-called form factor that describes the transition in the lower half of the diagram, f_H is the decay constant which quantifies the contribution from the 'ejected' meson, and $|a_{NF}^T(LH)|$ is a quantity which describes the deviation from naive factorisation for a tree diagram. For $B^0 \to \pi^- D_s^+$, this is equal to the total amount of deviation from naive factorisation $|a_{NF}(\pi^- D_s^+)|$.

Non-factorisable effects can occur, for example, due to gluon exchange between the two hadrons in the final state. If $|a_{NF}^T(LH)|$ would be equal to or close to one, factorisation would be a good approximation. For most B decays, where the ejected meson has a low mass (as it is a pion or kaon), this is expected to be the case [12].

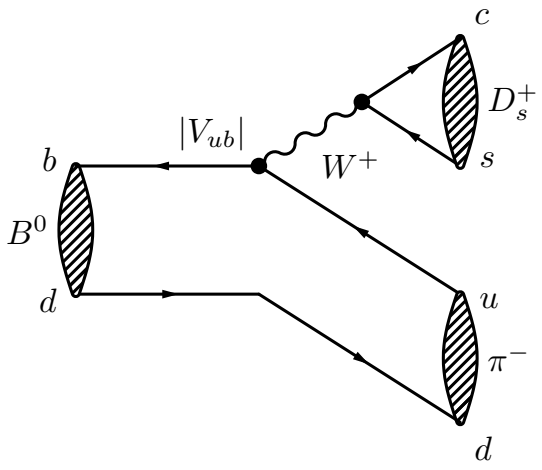

Figure 4: A tree diagram, which is the only diagram contributing to $B^0 \to \pi^- D_s^+$.

However, $B^0 \to \pi^- D_s^+$ is a decay where the low mass particle (pion) contains the spectator quark, while the high mass particle (D meson) is formed from the W. Thus, it is not possible to estimate $|a_{NF}^T(LH)|$ from decays such as $B^0 \to D^- K^+$ and apply it to this situation.

The measurement of the branching fraction of $B^0 \to \pi^- D_s^+$ is sensitive to the product of $|V_{ub}|$ and $|a_{NF}|$. By comparing the decay width of $B^0 \to \pi^- D_s^+$ with the decay rate of $B^0 \to \pi^- \ell^+ \nu_\ell$, the non-factorisable effects $|a_{NF}|$ can be quantified, as discussed in Ref. [12], using the following expression:

$$R_{D_s} = \frac{\Gamma(B^0 \to \pi^- D_s^+)}{d\Gamma(B^0 \to \pi^- \ell^+ \nu_\ell)/dq^2|_{q^2=m_{D_s}^2}} = 6\pi^2 |V_{cs}|^2 f_{D_s}^2 |a_{NF}(\pi D_s)|^2 X_{B^0\pi}^{D_s} \tag{7}$$

where

$$X_{B^0\pi}^{D_s} = \frac{(m_{B^0}^2 - m_\pi^2)^2}{[m_{B^0}^2 - (m_\pi^2 + m_{D_s})^2][m_{B^0}^2 - (m_\pi^2 - m_{D_s})^2]} \left[\frac{F_0^{B^0\pi}(m_{D_s}^2)}{F_1^{B^0\pi}(m_{D_s}^2)}\right]^2 \tag{8}$$

is a phase space parameter, with the ratio of form factors

$$\left[\frac{F_0^{B^0\pi}(m_{D_s}^2)}{F_1^{B^0\pi}(m_{D_s}^2)}\right] = 1 - \frac{m_{D_s}^2}{(m_{B^0} + m_\pi)^2} \tag{9}$$

However, if this formula is used to determine $|a_{NF}|$, the decay $B^0 \to \pi^- D_s^+$ cannot be used to also determine $|V_{ub}|$. This is clear when combining Eq. (6) with Eq. (7), which results in a formula for the semileptonic decay rate:

$$d\Gamma(B^0 \to \pi^- \ell^+ \nu_\ell)/dq^2|_{q^2=m_{D_s}^2} \sim \frac{G_F^2 |V_{ub}|^2 F(B^0 \to \pi)^2|_{q^2=M_{D_s}^2}}{6\pi^2 X_{B^0\pi}^{D_s}} \tag{10}$$

In this case, the use of hadronic information has proven superfluous, as this formula only contains information from the semi-leptonic decay, and can be used to determine $|V_{ub}|$. Using the current world average for $|V_{ub}|$ and $\mathcal{B}(B^0 \to \pi^- D_s^+)$, as given by PDG,

$$\mathcal{B}(B^0 \to \pi^- D_s^+) = (21.6 \pm 2.6) \times 10^{-6} \tag{11}$$

we can determine $|a_{NF}|$. This is done by calculating $d\Gamma(B^0 \to \pi^- \ell^+ \nu_\ell)/dq^2|_{q^2=m_{D_s}^2}$ using the prescription in Ref. [13]. It is found that

$$|a_{NF}(\pi D_s)| = 0.79 \pm 0.12 \tag{12}$$

This result is consistent with naive factorisation.

In order to determine $|V_{ub}|$, external input for $|a_{NF}|$ is needed. To determine $|a_{NF}^T(\pi D_s)|$ without using $\mathcal{B}(B^0 \to \pi^- D_s^+)$, $B \to D\overline{D}$ decays are examined, as they also contain a heavy ejected meson. Therefore, factorisation is investigated in these decays to obtain an estimate which might be used in the decay $B^0 \to \pi^- D_s^+$.

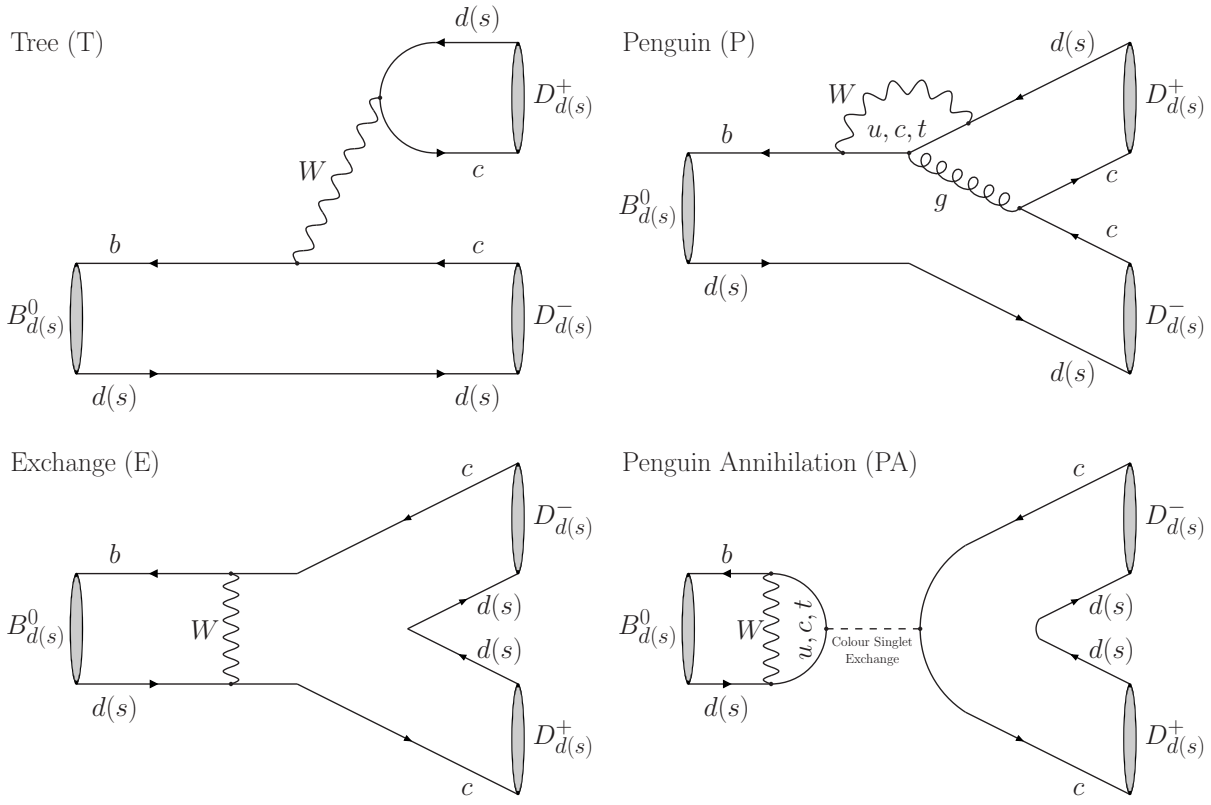

Figure 5: Illustration of topologies contributing to $B^0 \to D\bar{D}$ decays.

2.3 Factorisation in $B \to D\overline{D}$ decays

2.3.1 $B \to D\overline{D}$ decay structure

The first decay to compare to $B^0 \to \pi^- D_s^+$ is $B^0 \to D^- D_s^+$, as they proceed through a similar tree diagram, with the only difference being a $\bar{b} \to \bar{c}$ transition instead of a $\bar{b} \to \bar{u}$ transition in the lower part of the diagram. However, $B^0 \to D^- D_s^+$ also contains a contribution from so-called penguin diagrams, a loop diagram, of which an example is shown in Figure 5.

Depending on the quark being exchanged in the penguin loop, three combinations of CKM elements play a role: $V_{us}V_{ub}{}^*, V_{cs}V_{cb}{}^*$, and $V_{ts}V_{tb}{}^*$. In the Wolfenstein parametrisation [9], these combinations are given by $[\lambda][A\lambda^3(\rho + i\eta)], [1 - \frac{\lambda^2}{2}][A\lambda^2]$, and $[-A\lambda^2][1 - \frac{\lambda^2}{2}]$ respectively. In the following, we will reparametrise $(\rho - i\eta) = R_b e^{i\gamma}$. Using the unitarity condition, shown in Eq. (2), one of these combinations can be eliminated. The combination which occurs under exchange of top quarks is chosen. As a result, the following amplitudes can be written down for the $B^0 \to D^- D_s^+$ decay:

$$A(B^0 \to D^- D_s^+) = \left(1 - \frac{\lambda^2}{2}\right) \tilde{\mathcal{A}}[1 + \varepsilon \tilde{a} e^{i\tilde{\theta}} e^{i\gamma}], \tag{13}$$

where γ is the CP-violating weak phase which is one of the angles of the unitarity triangle, while $\mathcal{A}$ quantifies the amplitudes from the tree diagram and the penguin diagrams with the same weak phase,

$$\mathcal{A} = \lambda^2 A[\tilde{T} + \tilde{P}^{(ct)}], \tag{14}$$

with

$$\tilde{P}^{(q_1 q_2)} \equiv \tilde{P}^{(q_1)} - \tilde{P}^{(q_2)}. \tag{15}$$

The factor $\tilde{a}e^{i\tilde{\theta}}$ quantifies the magnitudes and the relative phase of the remaining penguin contributions,

$$\tilde{a}e^{i\tilde{\theta}} = R_b \left[\frac{\tilde{P}^{(ut)}}{\tilde{T} + \tilde{P}^{(ct)}} \right] \tag{16}$$

and ε is the CKM suppression factor,

$$\varepsilon \equiv \frac{\lambda^2}{1-\lambda^2} = 0.0536 \pm 0.0003. \tag{17}$$

Another decay which is interesting to investigate is $B^0 \to D^- D^+$, as it has the same kinematic strucure as $B^0 \to D^- D_s^+$. In addition to tree and penguin diagrams, $B^0 \to D^- D^+$ also contains exchange and penguin annihilation diagrams, which are also shown in Figure 5. The size of these diagrams can be determined using the decay $B^0 \to D_s^- D_s^+$, which only decays via exchange and penguin annihilation diagrams. An overview of the contributing decays is given in Table 1. Depending on the kind of up-quark being exchanged in the penguin or penguin annihilation diagrams for both of these decays, three combinations of CKM elements play a role: $V_{ud}V_{ub}{}^*$, $V_{cd}V_{cb}{}^*$, and $V_{td}V_{tb}{}^*$. In the Wolfenstein parametrisation, these combinations are given by $[1-\frac{\lambda^2}{2}][A\lambda^3(\rho+i\eta)]$, $[-\lambda][A\lambda^2]$, and $[A\lambda^3(1-\rho-i\eta)][1-\frac{\lambda^2}{2}]$ respectively. Again, the top quark contribution is eliminated using Eq. (2). The amplitudes of both decays can be written as

$$A(B^0 \to D^- D^+) = -\lambda \mathcal{A}[1 - a e^{i\theta} e^{i\gamma}], \tag{18}$$

where

$$\mathcal{A} = \lambda^2 A[T + E + P^{(ct)} + PA^{(ct)}], \tag{19}$$

$$a e^{i\theta} = R_b \left[\frac{P^{(ut)} + PA^{(ut)}}{T + E + P^{(ct)} + PA^{(ct)}} \right], \tag{20}$$

and

$$A(B^0 \to D_s^- D_s^+) = -\lambda \mathcal{A}_{EPA}[1 - a_{EPA} e^{i\theta_{EPA}} e^{i\gamma}], \tag{21}$$

where

$$\mathcal{A}_{EPA} = \lambda^2 A[\hat{E} + \hat{PA}^{(ct)}], \tag{22}$$

$$a_{EPA} e^{i\theta_{EPA}} = R_b \left[\frac{\hat{PA}^{(ut)}}{\hat{E} + \hat{PA}^{(ct)}} \right]. \tag{23}$$

Decay	$\mathcal{A}$	Topologies				Used for:
		T	P	E	PA	
$B^0 \to D^- D^+$	$\mathcal{A}$	x	x	x	x	determination of a and θ (and ϕ_d)
$B^0 \to D^- D_s^+$	$\tilde{\mathcal{A}}$	x	x			non-factorisable effect $\tilde{a}'_{\rm NF}$
$B^0 \to D_s^- D_s^+$	$\mathcal{A}_{EPA}$			x	x	quantify $E + PA$ contribution $\tilde{x}$
$B_s^0 \to D_s^- D_s^+$	$\mathcal{A}'$	x	x	x	x	physics goal ϕ_s
$B_s^0 \to D_s^- D^+$	$\tilde{\mathcal{A}}'$	x	x			SU(3) breaking non-fact. $\tilde{a}_{\rm NF}/\tilde{a}'_{\rm NF}$
$B_s^0 \to D^- D^+$	$\mathcal{A}'_{EPA}$			x	x	quantify $E + PA$ contribution $\tilde{x}'$

Table 1: Overview of the various topologies contributing to the $B \to D\bar{D}$ decays. The naming convention is indicated in the second column.

The difficulty in determining the exchange and penguin annihilation contributions comes from the fact that $\mathcal{B}(B^0 \to D_s^- D_s^+)$ has not been measured yet. Additionally, it is not possible to relate the topological amplitudes with each other directly, due to the different orders of λ with which the various penguin terms $ae^{i\theta}$ appear in the decay amplitudes.

Therefore, the same range of B_s^0 (instead of B^0) decays was used to determine the non-factorisable and the $E+PA$ contributions. This is possible because of SU(3)-symmetry, also known as U-spin flavour symmetry: to good approximation, the exchange of one of the triplet of (u, d, s)-quarks for another does not change the topological amplitude.

First, the following decays were compared: $B_s^0 \to D_s^- D_s^+$ with $B^0 \to D^- D_s^+$, and $B_s^0 \to D^- D^+$ with $B^0 \to D^- D_s^+$. The advantage of examining these pairs of decays is that they have the same kind of $P^{(ut)}$ structure. Taking the ratio will allow us to estimate the contributions from exchange and penguin annihilation contributions (E+PA).

The decay structures for the different $B_s^0 \to D_{(s)} D_{(s)}$ are given by:

$$A(B_s^0 \to D_s^- D_s^+) = \left(1 - \frac{\lambda^2}{2}\right) \mathcal{A}'[1 + \varepsilon a' e^{i\theta'} e^{i\gamma}], \tag{24}$$

$$A(B_s^0 \to D_s^- D^+) = \lambda \tilde{\mathcal{A}}'[1 - \tilde{a}' e^{i\tilde{\theta}'} e^{i\gamma}], \tag{25}$$

$$A(B_s^0 \to D^- D^+) = \left(1 - \frac{\lambda^2}{2}\right) \mathcal{A}'_{EPA}[1 + \varepsilon a'_{EPA} e^{i\theta'_{EPA}} e^{i\gamma}], \tag{26}$$

where both $\mathcal{A}'$ and $a' e^{i\theta'}$ are defined analogously to the "unprimed" expressions. U-spin flavour symmetry implies that these "primed" observables are equal to their "unprimed" counterparts:

$$\mathcal{A} = \mathcal{A}', \qquad a e^{i\theta} = a' e^{i\theta'} \tag{27}$$

2.3.2 Determining the size of $E + PA$ contributions

By taking the ratio of amplitudes for $B_s^0 \to D_s^- D_s^+$ and $B^0 \to D^- D_s^+$, we rewrite

$$\frac{A(B_s^0 \to D_s^- D_s^+)}{A(B^0 \to D^- D_s^+)} = \left(\frac{\mathcal{A}'}{\tilde{\mathcal{A}}}\right)\left[\frac{1+\varepsilon a' e^{i\theta'} e^{i\gamma}}{1+\varepsilon \tilde{a} e^{i\tilde{\theta}} e^{i\gamma}}\right] = \left[\frac{T'+P^{(ct)'}}{\tilde{T}+\tilde{P}^{(ct)}} + \tilde{x}'\right]\left[\frac{1+\epsilon a' e^{i\theta'} e^{i\gamma}}{1+\epsilon \tilde{a} e^{i\tilde{\theta}} e^{i\gamma}}\right], \quad (28)$$

where

$$\tilde{x}' \equiv |\tilde{x}'| e^{i\tilde{\sigma}'} \equiv \frac{E'+PA^{(ct)'}}{\tilde{T}+\tilde{P}^{(ct)}} \quad (29)$$

quantifies the amount of $E + PA$ contributions for the decay $B^0 \to D^- D^+$. This is a complex quantity, which means that the size and phase have both to be determined. As the penguin parameters enter with a factor ε, they are neglected in the ratio. The definition

$$\varrho' \equiv |\varrho'| e^{i\omega'} \equiv \frac{T'+P^{(ct)'}}{\tilde{T}+\tilde{P}^{(ct)}} = \left[\frac{T'}{\tilde{T}}\right]\left[\frac{1+P^{(ct)'}/T'}{1+\tilde{P}^{(ct)}/\tilde{T}}\right], \quad (30)$$

is introduced, after which

$$\frac{A(B_s^0 \to D_s^- D_s^+)}{A(B^0 \to D^- D_s^+)} = \varrho' + \tilde{x}'\,. \quad (31)$$

is obtained. Rewriting this to use the experimental information directly, it is found that

$$\Xi(B_s^0 \to D_s^- D_s^+, B^0 \to D^- D_s^+) = |\varrho'|^2 + 2|\varrho'||\tilde{x}'|\cos(\omega' - \tilde{\sigma}') + |\tilde{x}'|^2 \quad (32)$$

which fixes a circle for $|\tilde{x}'|$ in the complex plane, as shown in Figure 6. The experimental input is given by

$$\begin{aligned}&\Xi(B_s^0 \to D_s^- D_s^+, B^0 \to D^- D_s^+)\\ &\quad \equiv \left[\frac{m_{B_s^0}}{m_{B^0}} \frac{\Phi(m_D/m_{B^0}, m_{D_s^+}/m_{B^0})}{\Phi(m_{D_s^+}/m_{B_s^0}, m_{D_s^+}/m_{B_s^0})} \frac{\tau_{B^0}}{\tau_{B_s^0}}\right]\left[\frac{\mathcal{B}(B_s^0 \to D_s^- D_s^+)_{\text{theo}}}{\mathcal{B}(B^0 \to D^- D_s^+)}\right] = 0.647 \pm 0.049\end{aligned} \quad (33)$$

Here Φ describes a kinematic factor given by

$$\Phi(x,y) = \sqrt{[1-(x+y)^2]\,[1-(x-y)^2]}, \quad (34)$$

$\tau_{B^0{}_q}$ is the lifetime of a B_q meson, and the branching ratio of $B_s^0 \to D_s^- D_s^+$ is scaled to account for a difference between the theoretical and experimental definition.

To determine $\tilde{x}'$, another ratio of branching ratios is needed. Using $B_s^0 \to D^- D^+$ and $B^0 \to D^- D_s^+$ means that the ratio of $T + P$ and $E + PA$ contributions can be estimated directly. This ratio can be written as

$$\frac{A(B_s^0 \to D^- D^+)}{A(B^0 \to D^- D_s^+)} = \left(\frac{\mathcal{A}'_{EPA}}{\tilde{\mathcal{A}}'}\right)\left[\frac{1+\epsilon a'_{EPA} e^{i\theta'_{EPA}} e^{i\gamma}}{1+\epsilon \tilde{a} e^{i\tilde{\theta}} e^{i\gamma}}\right] = \varsigma' \tilde{x}'\,, \quad (35)$$

wherethe terms proportional to the tiny ϵ factor have been neglected and the parameter

$$\varsigma' \equiv \frac{\hat{E}' + \hat{PA}'^{(ct)}}{E' + PA'^{(ct)}} \approx \left(\frac{f_D}{m_D}\frac{m_{D_s}}{f_{D_s}}\right)^2 = 0.700 \pm 0.042\,. \quad (36)$$

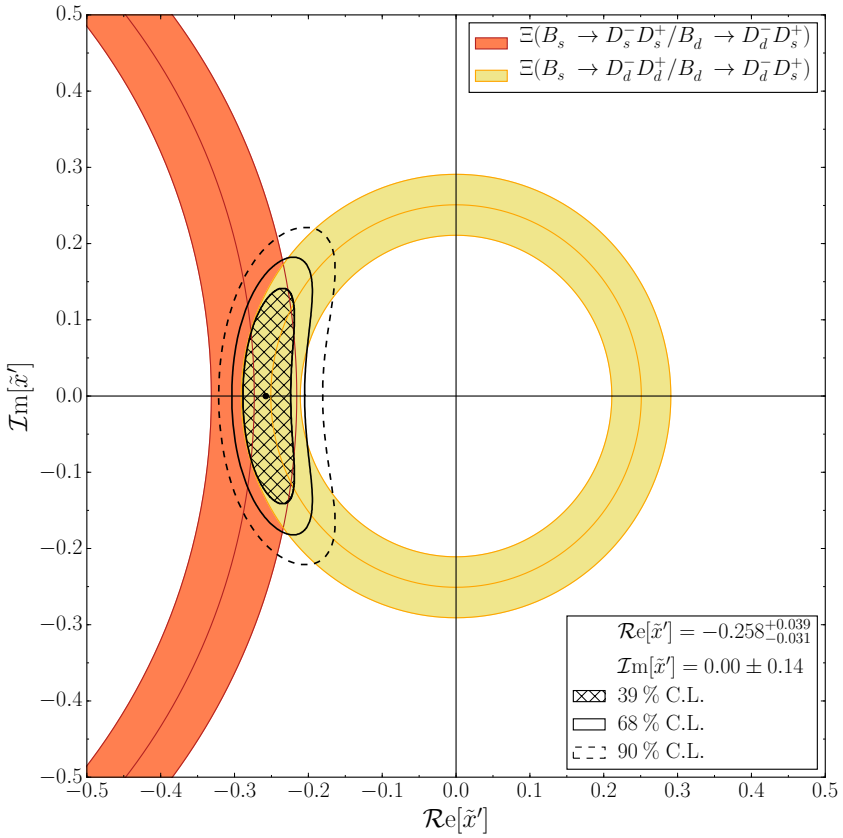

Figure 6: The fit for the amount of E + PA contributions.

has been introduced. Consequently, the ratio

$$\Xi(B_s^0 \to D^- D^+, B^0 \to D^- D_s^+) \equiv \left[\frac{m_{B_s^0}}{m_{B^0}} \frac{\Phi(m_D/m_{B^0}, m_{D_s^+}/m_{B^0})}{\Phi(m_D/m_{B_s^0}, m_D/m_{B_s^0})} \frac{\tau_{B^0}}{\tau_{B_s^0}}\right] \left[\frac{\mathcal{B}(B_s^0 \to D^- D^+)_{\text{theo}}}{\mathcal{B}(B^0 \to D^- D_s^+)}\right] = 0.031 \pm 0.009 \quad (37)$$

takes the simple form

$$\Xi(B_s^0 \to D^- D^+, B^0 \to D^- D_s^+) = |\varsigma' \tilde{x}'|^2 \,, \quad (38)$$

which fixes a circle with radius $|\varsigma' \tilde{x}'|$ around the origin in the complex plane, as shown in Figure 6.

To combine Eq. 33 with Eq. 37 and determine $\tilde{x}'$, another piece of information: the size of ρ' needs to be included. This parameter describes the SU(3)-breaking effects due to the spectator quark difference of $B^0 \to D^- D_s^+$ and $B_s^0 \to D_s^- D_s^+$, and is determined to be

$$\rho' = 1.078 \pm 0.051 \quad (39)$$

At this point, $\tilde{x}'$ can be determined, assuming that $\omega = 0$. It is found that

$$|\tilde{x}'| = 0.258^{+0.031}_{-0.039} \,, \qquad \tilde{\sigma}' = (180 \pm 34)^\circ \,, \quad (40)$$

Finally, to convert this to x, we use the SU(3)-relation

$$x \approx \left(\frac{m_{B_s^0} m_{D_s}}{m_{B^0} m_D}\right)\left(\frac{f_{B_s^0} f_{D_s}}{f_{B^0} f_D}\right)\tilde{x}' = -0.182^{+0.028}_{-0.022}\,, \tag{41}$$

We now know the relative contribution of exchange and penguin amplitudes in $B \to D\bar{D}$ decays is around 20%. The next step is to determine the non-factorisable effects in the tree diagram $|a_{NF}^T|$.

2.3.3 Using semileptonic $B \to D$ decays

Now that the structure of $B \to D\bar{D}$ decays has been investigated, $|a_{NF}|$ can be determined. As for $B^0 \to \pi^- D_s^+$, the hadronic decay width is compared with a semileptonic decay rate:

$$R_L \equiv \frac{\Gamma(B \to DD_q)}{d\Gamma(B \to D\ell\nu)/dq^2|_{q^2=m^2_{D_q}}} \tag{42}$$

$$= 6\pi^2 |V_{cq}|^2 f^2_{D_q} |a_{NF}(DD_q)|^2 X_L [1 - 2b_q \cos p_q \cos\gamma + b_q^2] \tag{43}$$

with $q \in \{d, s\}$, $b_q e^{ip_q} = ae^{i\theta}, \varepsilon\tilde{a}e^{i\tilde{\theta}}$ for $B^0 \to D^- D^+$ and $B^0 \to D^- D_s^+$ respectively. As a result, this has to be taken into account for determining $|a_{NF}(DD)|$, while it can be ignored when determining $|a_{NF}(DD_s)|$. $|a_{NF}|$ is a quantity which describes the effect of non-factorisable contributions in the amplitudes defined using Eq. (14) and Eq. (19), which can be written as

$$\mathcal{A} = \lambda^2 A T_{\text{fact}} a_{NF} \tag{44}$$

where

$$T_{\text{fact}} = -i\frac{G_F}{\sqrt{2}} f_{D_q} (m^2_{B^0} - m^2_{D_q}) [F(B^0 \to D)]^2_{q^2=M^2_{D_q}} a_1 \tag{45}$$

is the amplitude of the tree diagram in factorisation, G_F is the Fermi constant and a_1 is the appropriate combination of Wilson coefficient functions of the current-current operators. Note that a_{NF} contains contributions due to penguin diagrams and due to non-factorisable effects in the tree diagram $|a_{NF}^T|$.

For $B^0 \to D^- D_s^+$, the penguin contributions from $P^{(ut)}$ are suppressed by a factor ε, and it is possible to neglect them when determining $|a_{NF}(DD_s)|$. However, the contribution from $P^{(ct)}$ is not suppressed, and

$$a_{NF}(DD_s) = a_{NF}^T(DD_s)(1 + \tilde{r_P}) \tag{46}$$

is obtained, where $\tilde{r_P} = \tilde{P}^{(ct)}/\tilde{T}$.

For $B^0 \to D^- D^+$, there is an extra non-factorisable contribution from $E + PA$ diagrams, and the penguin contributions are not suppressed. This means that both the determination of $|a_{NF}(DD)|$ and the conversion to $|a_{NF}(DD)^T|$ are different. First, the determination of $|a_{NF}(DD)|$ will now include the penguin parameters directly. Secondly, the formula to relate $|a_{NF}(DD)|$ to $|a_{NF}(DD)^T|$ is now given by

$$a_{NF}(DD) = a_{NF}^T(DD)(1 + r_P)(1 + x) \tag{47}$$

where $r_P = P^{(ct)}/T$, $x = (E + PA^{(ut)})/(T + P^{(ct)})$.

To determine $|a_{NF}|$ using Eq.(43), the form factor parametrisation given by Ref. [14] is used to calculate the semileptonic decay rates:

$$\frac{d\Gamma(B \to D\ell\nu)}{dw} = \frac{G_F^2 |V_{cb}|^2}{48\pi^3}(m_B + m_D)^2 m_D^3 (w^2-1)^{3/2} |V_1(w)|^2 \tag{48}$$

where

$$w = \frac{m_{B^0}^2 + m_D^2 - m_{D_q}^2}{2m_B m_D}. \tag{49}$$

As a result, the differential rates are related by

$$\frac{d\Gamma}{dw} = \frac{1}{2m_B m_D}\frac{d\Gamma}{dq^2}. \tag{50}$$

Finally, the kinematic dependence is modeled by

$$\frac{V_1(w)}{V_1(1)} \approx 1 - 8\rho_1^2 z + (51\rho_1^2 - 10)z^2 - (252\rho_1^2 - 84)z^3 \tag{51}$$

where $z = (\sqrt{w+1} - \sqrt{2})/(\sqrt{w+1} + \sqrt{2})$. HFAG [15] has determined $|V_{cb}|V_1(1) = 42.65 \pm 0.72 \pm 1.35$ and $\rho_1^2 = 1.185 \pm 0.035 \pm 0.041$.

The non-factorisable effects for the heavy-light $B \to Dh$ decays are expected to be smaller than for $B^0 \to D\bar{D}$ decays, and do not suffer from penguin effects. The value of the non-factorisable effects for $B \to Dh$ decays are included in Table 2 for completeness.

Table 2: The non-factorisable contributions $|a_{NF}|$ determined with Eq (43). For all decays except $B^0 \to D^- D^+$, penguin contributions are neglected while determining $|a_{NF}|$.

Decay	$\lvert a_{NF}\rvert$	$\lvert a_{NF}^T\rvert$
$B^0 \to D^- D_s^+$	0.75 ± 0.06	1.36 ± 0.36
$B^0 \to D^- D^+$	0.50 ± 0.08	1.11 ± 0.34
$B^0 \to D^- K^+$	0.87 ± 0.07	0.87 ± 0.07
$B^0 \to D^- \pi^+$	0.89 ± 0.06	0.89 ± 0.06

2.3.4 Factorisation results for $B \to DX$ decays

Using these expressions, $|a_{NF}(DD_q)|$ is calculated. The resulting factorisation parameters are shown in Table 2. As discussed earlier, $|a_{NF}|$ only corresponds to the amount of non-factorisability in the tree diagram $|a_{NF}^T|$ if other contributions are small, which is the case for $B^0 \to D^- K^+$ and $B^0 \to D^- \pi^+$. However, for $B^0 \to D^- D_s^+$ and $B^0 \to D^- D^+$, additional contributions have to be taken into account. A method to obtain a and θ was discussed in Ref. [16]. The same method to estimate non-factorisable and $E + PA$ effects was used. In addition, $B^0 \to D^- D^+$ was compared with $B_s^0 \to D_s^- D_s^+$ to determine a and θ. The following values were obtained:

$$a = 0.35^{+0.19}_{-0.20}\,, \qquad \theta = \left(215^{+51}_{-17}\right)^\circ\,. \tag{52}$$

As this parameter describes the relative size of the $P^{(ut)} + PA^{(ut)}$ amplitudes compared to the $T + P^{(ct)} + PA^{(ct)}$, it is used in the determination of $|a_{NF}(DD)|$.

To determine the amount of non-factorisability in the tree diagram, the relative size of the penguin contributions in both $B^0 \to D^- D^+$ and $B^0 \to D^- D_s^+$ need to be determined. To do so, some assumptions are made. First, it is assumed that $P^{(ct)} = P^{(ut)}$, such that the approximation $ae^{i\theta} \approx R_b r_P (1 + r_P) \approx R_b \tilde{r}'_P (1 + \tilde{r}'_P)$ can be made. In this way, it is obtained that

$$|1 + r_P| \approx |1 + \tilde{r}'_P| = 0.55 \pm 0.14, \tag{53}$$

where the accompanying angle in the complex plane is $\theta_{1+r_P} = -16^\circ$ with unknown uncertainty. To good approximation, this angle can be ignored as the projection on the real axis is given by $\cos(16^\circ) = 0.96$.

As x has been determined in Eq. (41) ($x \approx -0.182^{+0.028}_{-0.022}$), it is now possible to calculate $|a_{NF}^T|$ for both $B^0 \to D^- D^+$ and $B^0 \to D^- D_s^+$. It should be emphasized that the uncertainty on this calculation might be larger, due to additional uncertainties from the $P^{(ct)} = P^{(ut)}$ assumption.

The resulting factorisation parameters are shown in Table 2. Comparing all final results, it is clear that they are quite consistent with factorisation. However, it also shows that the value of $|a_{NF}^T|$ from $B \to D\bar{D}$ decays is still poorly known, thereby preventing the possible use of this result to determine $|a_{NF}^T|$ for $B^0 \to \pi^- D_s^+$ at this stage. On the other hand, future theoretical and experimental developments might change the situation.

3 LHC and the LHCb detector

3.1 LHC

The Large Hadron Collider (LHC) is the largest and most powerful particle collider ever built [17]. It was designed to accelerate protons (lead ions) to an energy of 7(2.8) TeV/nucleon, resulting in a centre of mass energy of 14(574) TeV. The main goal of the LHC is to investigate the laws of nature at these higher energies, and to search for evidence of physics beyond the Standard Model. As evidence might be found in multiple ways, four different experiments are using the LHC:

- ATLAS [18] and CMS [19] are the two general purpose detectors at the LHC. They have observed the Higgs boson, one of their main goals. In addition, they will continue to search for directly produced new, heavy particles, for example predicted by supersymmetry.
- ALICE [20] has been built to investigate heavy-ion collisions. The extreme energy densities produced by heavy-ion collisions at the LHC should result in the creation of a new state of matter, called quark-gluon plasma. Studying such a state would help us to understand QCD.
- LHCb [21] is a detector which aims to detect rare decays and CP violation, mainly by studying mesons containing b quarks, so-called B mesons. New, heavy particles can affect many observables in B decays. The rest of this chapter will discuss the LHCb detector and its components in detail.

3.2 The LHCb Detector

The LHCb detector has been designed as a single-arm forward spectrometer. This choice is justified by the fact that b and $\bar{b}$-hadrons are mainly produced in the forward and backward cones around the beam axis.

The coordinate system used by LHCb can be expressed in the three coordinates: z, ϕ and η. First, the cartesian coordinates x, y, z have to be defined. Pointing along the beam axis is z, with y pointing upwards and x pointing away from the LHC ring. The resulting polar coordinates use z, the angle from the x to the y axis as ϕ, and the angle from the z to the y axis as θ. The pseudorapidity η is given by

$$\eta = -\log \tan \theta/2 \tag{54}$$

and is often quoted to indicate the geometric acceptance of a detector. LHCb has a range of $2 < \eta < 5$, equivalent to $0.77° < \theta < 15°$.

LHCb consists of a number of subdetectors, which all play a part in identifying particles and measuring their properties. These are the VELO, RICH, Transition Tracker, Inner and Outer Tracker, RICH2, SPD/PS, ECAL, HCAL and Muon stations.

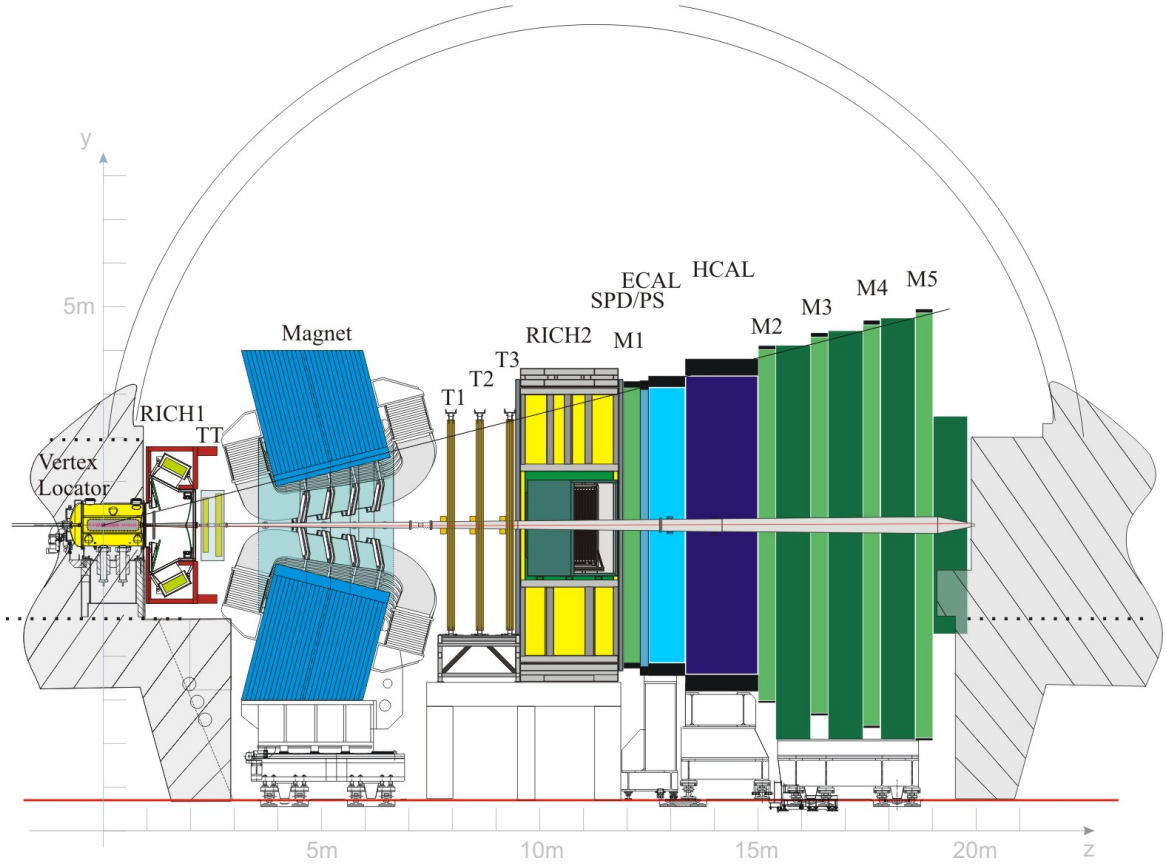

Figure 7: A cross section of the LHCb detector in the x-plane. From [21].

The Vertex Locator (VELO) is the subdetector closest to the beam pipe and beam collision point. This allows it to locate the vertices in particle decays with very high precision. For are a number of reasons, this is essential for LHCb. Most importantly, B mesons typically decay within ~ 1 cm, and measuring this distance is vital to determine the lifetime. Subsequently, time-dependent measurements, which form an important part of measuring CP violation, will all depend on the ability to determine the vertex locations with great precision.

Ring Imaging CHerenkov (RICH) detectors are needed for particle identification. Identifying the light decay products of B meson decays, pions or kaons, is essential for identifying the exact decay type. By using Cherenkov radiation, the RICH detectors can measure the velocity of the different charged particles, and thus distinguish pions from kaons and protons.

The different tracking detectors are used to measure the momenta of charged particles passing through the detector. Because the magnet in between the different tracking detectors bends charged particles, the curvature can be converted into a momentum measurement. In addition, the particle's charge can be determined from the direction of curvature. To achieve this, the LHCb magnet has a bending power of 4 Tm. The magnet

polarity can be reversed, thus cancelling possible detector asymmetries.

The Scintllating Pad Detector (SPD) helps to distinguish electrons from neutral pions and photons, because charged particles will deposit energy in a scintillator while neutral ones will not. The Pre-Shower (PS) detector can separate charged pions from more energetic electrons by using the energy deposits of particles individually.

Both ECAL and HCAL are calorimeters, which measure the electromagnetic and hadronic energy deposition, respectively.

Finally, the muon stations consist of five tracking stations, which are used to identify muons. As the rest of the detector absorbs almost all non-muons before they reach the Muon stations, muons are succesfully identified in 95% of the cases.

3.3 Data flow

To understand the analysis, the various stages of the selection are summarised below. The selection process already starts during data-taking, as a trigger is needed to reduce the data rate to a managable level. This trigger is composed of a hardware level (L0) trigger and a software trigger (HLT1 and HLT2). From a collision rate of 20 MHz, the hardware trigger records 1 MHz, which is sent to the software trigger. The software trigger can use the full event information, and further reduces the rate to 3 kHz of raw data. All data are saved and can be accessed at a later stage.

To keep data volume at a managable level, another selection step is needed. This step is called *Stripping*. After Stripping, the data can be used in an analysis. Stripping can be repeated at a later stage, which means that even the results obtained from the same raw data might improve. For example, new algorithms might enable more efficient track reconstruction, thus reducing combinatorial backgrounds which consist of random tracks and improving the amount of detected signal. A more detailed description of the event selection is described in the next chapter.

4 Measuring $\mathcal{B}(B^0 \to \pi^- D_s^+)$ at LHCb

4.1 Analysis strategy

In this section, we will discuss the various aspects involved in measuring the branching fraction of the decay $B^0 \to \pi^- D_s^+$ at LHCb, where the D_s^+ meson decays to $K^+K^-\pi^+$, such that four charged particles need to be reconstructed. The distribution of the invariant mass of the four final state particles will show a peak around the B^0 mass. The experimental analysis includes the selection of suitable signal events for analysis and the modeling backgrounds: combinatorial background and partially reconstructed(PR) and/or misidentified(misID) backgrounds.

As well as performing the event selection and modeling the backgrounds, it is necessary to normalise the result to a known branching fraction, as it is unknown how many B mesons are produced at LHCb to start with. At the B-factories, however, absolute branching fractions *have* been measured with sufficient precision. As a result, it is possible to measure a ratio of branching fractions at LHCb and obtain the interesting branching fraction by using the previously obtained result.

For this analysis, the $B^0 \to D^-\pi^+$ decay is used as normalisation channel. Of all topologically similar decays ($B^0 \to Dh$), this channel has the smallest uncertainty on the measured branching fraction: $\mathcal{B}(B^0 \to D^-\pi^+) = (2.68 \pm 0.13) \times 10^{-4}$. This induces a 4.9% uncertainty on the absolute branching fraction measurement.

The expression used to determine $\mathcal{B}(B^0 \to \pi^- D_s^+)$ is

$$\mathcal{B}(B^0 \to \pi^- D_s^+) = \mathcal{B}(B^0 \to D^-\pi^+) \frac{\varepsilon_{B^0 \to D^-\pi^+}}{\varepsilon_{B^0 \to \pi^- D_s^+}} \frac{N_{B^0 \to \pi^- D_s^+}}{N_{B^0 \to D^-\pi^+}} \frac{\mathcal{B}(D^- \to K^+\pi^-\pi^-)}{\mathcal{B}(D_s^+ \to K^-K^+\pi^+)} \tag{55}$$

where N_X is the measured yield and ε_X the efficiency of decay X. In this calculation, the following branching ratios are used from Ref. [11]:

$$\mathcal{B}(B^0 \to D^-\pi^+) = (2.68 \pm 0.13) \times 10^{-3}, \tag{56}$$
$$\mathcal{B}(D^- \to K^+\pi^-\pi^-) = (9.13 \pm 0.19) \times 10^{-2}, \tag{57}$$
$$\mathcal{B}(D_s^+ \to K^+K^-\pi^+) = (5.42 \pm 0.14) \times 10^{-2}. \tag{58}$$

Combined, this will contribute with a 5.9% uncertainty to the measurement. The goal of this analysis is to measure of the yields and efficiencies for $B^0 \to D^-\pi^+$ and $B^0 \to \pi^- D_s^+$ as precisely as possible, in order to determine $\mathcal{B}(B^0 \to \pi^- D_s^+)$.

4.2 Selection

To make an analysis possible, interesting events need to be selected and separated from background. This analysis is performed using data from pp collisions at the LHCb detector, of which 1 fb^{-1} is obtained at $\sqrt{s} = 7$ TeV and 2 fb^{-1} at $\sqrt{s} = 8$ TeV, for a total of 3

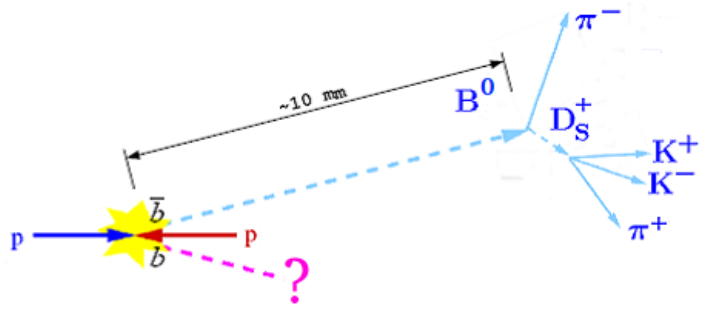

Figure 8: A schematic view of the $B^0 \to \pi^- D_s^+$ decay at LHCb, where the D_s^+ meson decays into $K^+K^-\pi^+$. This is an edited version of a diagram obtained from [22].

fb^{-1}. Because not every collision can be recorded and not every collision is interesting, a so-called trigger is used. The trigger applies a first selection, which is meant to satisfy two guidelines. First, as much interesting signal as possible should remain after selection. Second, a selection bias should be avoided.

4.2.1 Trigger selection

The trigger contains two stages: a hardware and a software stage. At the hardware level, events containing a hadron with transverse energy greater than 3.6 GeV are selected. The subsequential software trigger is slightly more complicated. It requires a two-, three- or four-track secondary vertex where the tracks have large transverse momenta(p_T), as well as a significant displacement from the primary vertices(PVs). At least one track should have $p_T > 1.7\,\text{GeV}$ and $\chi^2_{\text{IP}} > 16$ with respect to any PV. Here χ^2_{IP} is defined as the difference between the χ^2 of a PV reconstructed with and without the considered candidate track. After the triggering stage, all data are saved.

4.2.2 Stripping selection

The next step in the event selection is a pre-selection called Stripping. This step is mainly based on kinematical and geometrical variables. The same stripping lines as used in Ref. [23] were used. [1]

First, a $D_s^\pm$ candidate is constructed by requiring a combination of three light hadrons (pions or kaons), each of which has a good quality track, and $p_T > 100\,\text{MeV}/c$ and $p > 1000\,\text{MeV}/c$, which are typical cuts for decay products of charmed mesons. Secondly, final state particles which originate from a $B^0_{(s)}$ decay have to be detached from the primary vertex. This is implemented by requiring the primary vertex fit to reduce in quality if their

[1] These are *B02DKD2HHH* and *B02DPiD2HHHBeauty2Charmline*. Stripping versions 20 and 20r1 were used to analyse 2012 and 2011 data, respectively. There are small differences between these versions: the ghost probability of a track can be at most 0.4 or 0.3, respectively. In addition v20 requires a specific software trigger "HLT2Topo" or "HLT2IncPhi" to fire.

tracks are added to the primary vertex fit. Thirdly, one of the $D^{\pm}{}_{(s)}$ candidate daughters is required to have $p_T > 500\,\mathrm{MeV}/c$ and $p > 5000\,\mathrm{MeV}/c$. The invariant mass of this candidate should be between 1769.62 and 2068.49 MeV/c^2, equivalent to a mass window from 100 MeV/c^2 below the $D^{\pm}$ mass to 100 MeV/c^2 above the $D_s^{\pm}$ mass. The $D^{\pm}{}_{(s)}$ candidate should have a p_T greater than $1800\,\mathrm{MeV}/c$, and a distance of closest approach to the primary vertex of at least 0.5 mm. Finally, this $D^{\pm}{}_{(s)}$ candidate is combined with a bachelor pion. This bachelor pion should have a good track quality, be displaced from the primary vertex, and have $p_T > 500\,\mathrm{MeV}/c$ and $p > 5000\,\mathrm{MeV}/c$. The resulting $B^0_{(s)}$ candidate is required to have an invariant mass between 4750 and 6000 MeV/c^2 and a minimum lifetime of 2 ps. Of all of the final state particles, at least one should have $p_T > 1700\,\mathrm{MeV}/c$ and $p > 10\,\mathrm{GeV}/c$.

4.2.3 Analysis selection

After these general pre-selection criteria, more specific requirements are imposed. These can be divided in two types of cuts: kinematic and particle identification (PID) cuts. All of these cuts are discussed below. First, mass cuts are implemented. The $B^0_{(s)}$ candidate mass is required to be between 5000 and 5800 MeV/c^2, while the $D^{\pm}(D^{\pm}{}_s)$ candidate is required to have a mass between 1844 and 1890 MeV/c^2 (1944 and 1990 MeV/c^2). Secondly, so-called fiducial cuts are added, which limit the sample to the range in which the detector operates. The p_T of the $B^0_{(s)}$ candidate is limited to be from 1500 to 40000 MeV/c, and the pseudorapidity η to [2,5]. Thirdly, to make sure the final state particles originate from a $D^{\pm}{}_{(s)}$ particle, prompt decays are removed. This removes decays such as $B^0 \to K^+K^-\pi^+\pi^-$, where no intermediate D_s^+ state is part of the process. Finally, a cut on the particle identification is applied, for example to distinguish $D^+ \to K^-\pi^+\pi^+$ from $D_s^+ \to K^-K^+\pi^+$ decays. These are cuts on the difference of the log likelihood of the two particle hypotheses, called the $\mathrm{DLL}_{K\pi}$ and $\mathrm{DLL}_{p\pi}$ variables. The cuts, for each different decay channel, are outlined in Table 3. For the $B^0 \to \pi^- D_s^+$ analysis, an extra PID requirement is imposed: a veto on $\Lambda_b^0 \to \Lambda_c^+\pi^-$ ($\Lambda_c^+ \to pK^-\pi^+$) events, where the proton is misidentified as a kaon. This is done by reconstructing the D_s^+ mass under the p mass hypothesis and calculating the resulting invariant mass. If this reconstructed mass is within $21\,\mathrm{MeV}/c^2$ of the Λ_c^+ mass $2286.46\,\mathrm{MeV}/c^2$ and $\mathrm{DLL}_{p\pi}$- $\mathrm{DLL}_{K\pi}$ for an event is greater than 0, the event is rejected. This Λ_c^+ veto ensures almost a fully efficient rejection of this background.

A multivariate discrimination variable, called a Boosted Decision Tree (BDT), is used to separate combinatorial background from signal. It combines several kinematic and geometrical variables, and outputs a single value in the interval $[-1, 1]$ for each event. The lower the value, the more likely it is that the event is background; the higher the value, the more likely it is that the event is signal. The BDT has been calibrated using $B_s^0 \to D_s^-\pi^+$ events in the relevant mass window, weighted depending on its similarity to actual signal events. Background events have been sampled from the upper mass sideband with an invariant mass $> 5500\,\mathrm{MeV}$.

Table 3: PID cuts used for $B^0 \to \pi^- D_s^+$ and $B^0 \to D^- \pi^+$.

Cut	$B^0 \to \pi^- D_s^+$	$B^0 \to D^- \pi^+$
Bachelor π	$\mathrm{DLL}_{K\pi} < 0$	$\mathrm{DLL}_{K\pi} < 0$
$\pi^\pm$ from $D_{(s)}^\pm$	$\mathrm{DLL}_{K\pi} < 5$	$\mathrm{DLL}_{K\pi} < 5$
	$\mathrm{DLL}_{p\pi} < 15$	
K^+ from $D_{(s)}^-$	$\mathrm{DLL}_{K\pi} > 0$	$\mathrm{DLL}_{K\pi} > 0$
K^- from D_s^+	-	$\mathrm{DLL}_{K\pi} > 5$

4.2.4 Additional selection

Compared to the previous analysis of events with $D_s^+\pi^-$ candidates [23], we include extra cuts. These are needed to eliminate backgrounds and thus decrease the systematic uncertainty. The background events can be divided in two categories: combinatorial and physics backgrounds. The second category consists of partially reconstructed and/or misID decays, such as $B_s^0 \to D_s^{*-}\pi^+$ or $B^0 \to \pi(D^+ \to D_s^+)$ decays (see Subsection 4.4).

Two additional cuts are included. First, a muon veto is included: if the bachelor π is identified as a muon, the event is rejected. As the identification of muons at LHCb is very accurate, this leads to the elimination of 97% of this background while excluding only 0.75% of signal.

Secondly, different BDT cuts are investigated. The previous analysis (Ref. [23]) used BDT > 0.3 as a cut. In this analysis, two BDT regions are used for the final result: one with $0.85 >$ BDT > 0.0 and one with BDT > 0.85, called region 1 and 2 respectively. These regions are chosen for two reasons: a BDT > 0.0 cut will reject as little as 1.3% of signal while removing 55% of combinatorial background, and the total number of events in both regions is about the same. The analysis in two BDT bins allows for a consistency check, and aims to show that the result will not depend on a specific BDT requirement. Additionally, various BDT cuts are investigated to study the shape of combinatorial background.

As discussed earlier, the relative efficiencies of this selection procedure need to be determined to calculate the branching ratio. As $B^0 \to D^-\pi^+$ and $B^0 \to \pi^- D_s^+$ have similar decay topologies, these should be similar, but there might still be a relative difference due to kinematics or PID differences.

Apart from the PID efficiency, the selection efficiency is determined from MC 2011 and 2012 simulated samples, as the kinematic distributions are well modeled. These efficiencies are shown for $B^0 \to \pi^- D_s^+$ and $B^0 \to D^-\pi^+$ in Tables 4 and 5. The assumption is made that generator, reconstruction and stripping efficiencies are equal for $B^0 \to \pi^- D_s^+$ and $B_s^0 \to D_s^-\pi^+$.

However, the PID performance is not modelled accurately enough in simulation. Thus, a calibration sample from data is used to reweight the simulation samples. This sample consists of $D^{*+} \to (D^0 \to K^-\pi^+)\pi^+$ decays. For this decay, it is possible to determine

Table 4: Efficiencies of the kinematic cuts on $B^0 \to \pi^- D_s^+$. BDT region 2 is the high purity region with BDT > 0.85, and is an important region for the analysis.

	ε_{rel} (%)	ε_{cum} (%)
Generator level efficiency	17.54 ± 0.19	17.54 ± 0.19
Reconstruction and stripping	14.37 ± 0.03	2.52 ± 0.05
Trigger cuts	95.15 ± 0.27	2.40 ± 0.03
$B^0_{(s)}$ mass window cuts	98.54 ± 0.15	2.36 ± 0.03
$D^{\pm}_{(s)}$ mass window cuts	98.31 ± 0.16	2.33 ± 0.03
Fiducial cuts	99.41 ± 0.10	2.31 ± 0.03
Flight distance cuts	88.48 ± 0.41	2.06 ± 0.03
Λ_c^+ veto	99.41 ± 0.10	2.05 ± 0.03
μ veto	99.25 ± 0.11	2.03 ± 0.03
BDT > 0.0	98.70 ± 0.14	2.02 ± 0.03
BDT region 1(2)	34.27 ± 0.61 (64.43 ± 0.61)	$0.64 \pm 0.02(1.37 \pm 0.02)$

Table 5: Efficiencies of the kinematic cuts on $B^0 \to D^- \pi^+$. BDT region 2 is the high purity region with BDT > 0.85, and is an important region for the analysis.

	ε_{rel} (%)	ε_{cum} (%)
Generator level efficiency	17.54 ± 0.19	17.54 ± 0.19
Reconstruction and stripping	14.37 ± 0.03	2.52 ± 0.05
Trigger cuts	96.70 ± 0.07	2.18 ± 0.01
$B^0_{(s)}$ mass window cuts	99.32 ± 0.03	2.17 ± 0.01
$D^{\pm}_{(s)}$ mass window cuts	95.65 ± 0.18	2.08 ± 0.01
Fiducial cuts	99.53 ± 0.03	2.07 ± 0.01
Flight distance cuts	94.24 ± 0.09	1.95 ± 0.01
μ veto	98.94 ± 0.04	1.93 ± 0.01
BDT > 0.0	98.57 ± 0.04	1.90 ± 0.01
BDT region 1(2)	$37.90 \pm 0.18(60.67 \pm 0.19)$	$0.69 \pm 0.00(1.22 \pm 0.01)$

PID efficiencies without requiring information from the RICH detector, thus yielding an accurate estimate of the pion-kaon efficiency. This estimate is made for different track momenta, track pseudorapidity η, and number of tracks in the events. The MC samples are reweighted to the PID performance histograms which are obtained this way, such that an estimate of the PID efficiency for $B^0 \to \pi^- D_s^+$ and $B^0 \to D^- \pi^+$ is obtained.

As an extra handle on the internal consistency of the analysis, the data is split into two samples, recorded with opposite magnet polarity.

4.3 Modeling of combinatorial background

Random combinations of D_s candidates and pions form a background for genuine $B_{(s)}$ candidates. To model this combinatorial background, both the shape and size of this background need to be estimated. Preferably, one or both of these are obtained from an independent dataset to avoid model dependence.

Random combinations of D_s with a same-sign pion can not originate from B decays, and are referred to as wrong-sign(WS) decays. As this will not contain any physical signal, it is expected to find the same background as in data: random combinations of $D_s^+\pi^+$ which are accidentally combined. An exponential function turns out to describe WS data. In addition, a constant is included in the fitting function. Thus, the fitting function is parametrised as:

$$f[m(D_s^+\pi^-)] = p_0 + e^{-p_1 m(D_s^+\pi^-)} \tag{59}$$

An example of wrong-sign data is shown in Figure 9. The resulting fit parameters were found to be $p_0 = (0.00 \pm 0.35) \times 10^{-2}$, $p_1 = (-5.36 \pm 0.19) \times 10^{-3}$ for magnet down data, and $p_0 = (0.00 \pm 0.17) \times 10^{-2}$, $p_1 = (-6.52 \pm 0.20) \times 10^{-3}$ for magnet up data. Thus, the WS data contains an exponential distribution without an offset p_0.

However, a close comparison of the shape of the combinatorial background for right sign and wrong sign candidates reveals a different behaviour. For various BDT requirements, WS data is fitted and compared with right-sign data. The results are shown

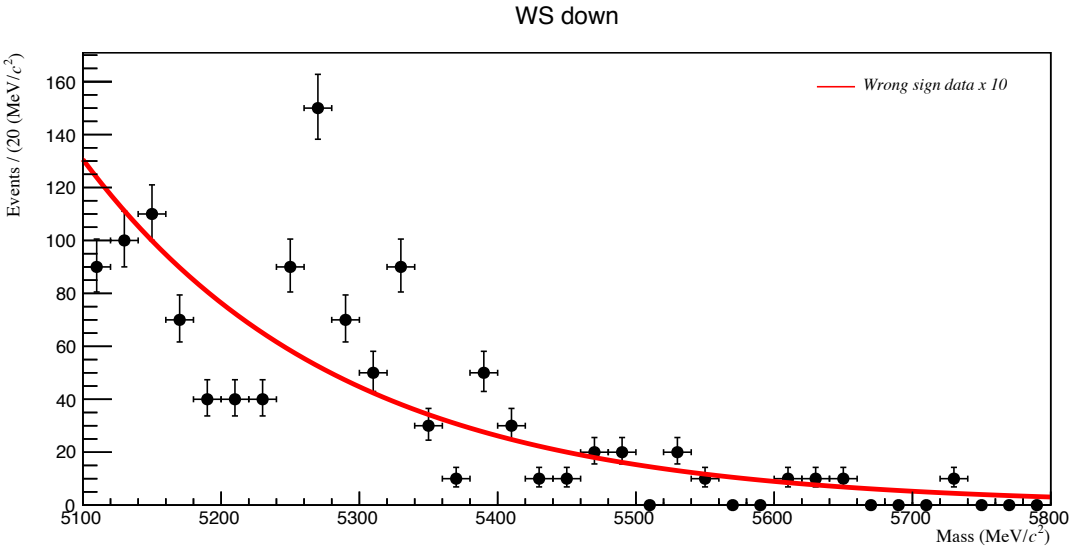

Figure 9: An example of wrong sign data, in this case for BDT region 2. The amount of events per bin has been multiplied by 10, to compensate for the prescale of a factor 10 in the stripping process. The resulting fit parameters were obtained to be $p_0 = (0.00 \pm 0.35) \times 10^{-2}$, $p_1 = (-5.36 \pm 0.19) \times 10^{-3}$ for magnet down data, and $p_0 = (0.00 \pm 0.17) \times 10^{-2}$, $p_1 = (-6.52 \pm 0.20) \times 10^{-3}$ for magnet up data.

in Figure 10. Interestingly, right sign candidates require a constant in addition to the exponential component. A possible explanation is that fragmentation of quarks is not charge independent, leading to differences in the shape of combinatorial background. As a result, the combinatorial background cannot be determined from an independent sample, and has to be determined from the signal sample itself. To limit any systematic uncertainty from this, the combinatorial background is rejected as much as possible.

The amount of combinatorial rejection in the final selection is shown in Figure 11. This plot compares the amount of $B_s^0 \to D_s^- \pi^+$ events with the amount of combinatorial background. Compared to no cut at all, the requirement of BDT > 0.0 ensures that 55% of combinatorial is rejected, while 98.7% of $B^0 \to \pi^- D_s^+$ candidates is still accepted, as was shown in Table 4.

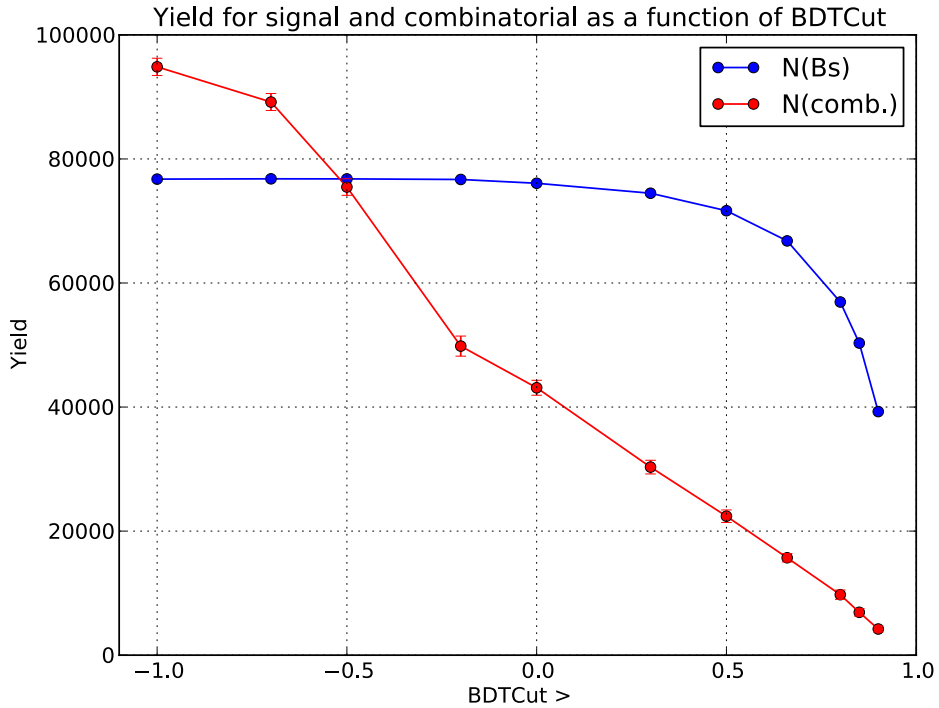

Figure 10: The combinatorial parameters p_0 and p_1 for different BDT cuts without muon veto. A clear difference between right-sign and wrong-sign data can be observed, while magnet up and down results are clearly consistent with each other. The smooth behaviour is seen, indicating that the fits are stable.

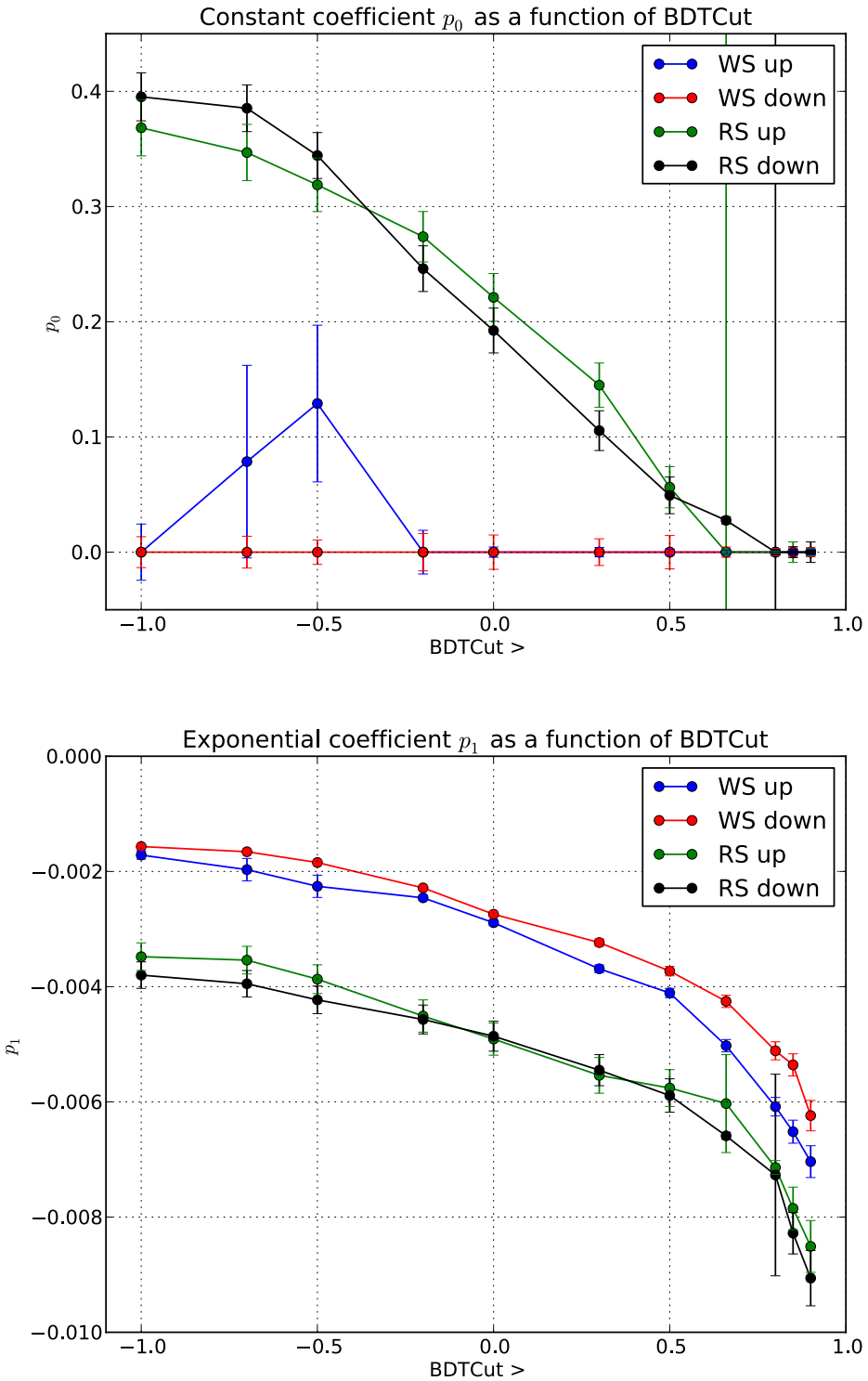

Figure 11: The combinatorial parameters p_0 and p_1 for different BDT cuts without muon veto. A clear difference between right-sign and wrong-sign data can be observed, while magnet up and down results are clearly consistent with each other. A smooth behaviour is seen, indicating that the fits are stable.

4.4 Modeling of partially reconstructed and misidentified backgrounds

For an analysis investigating the $D_s^+\pi^-$ candidates, backgrounds from other B decays can occur for two reasons:

Partial reconstruction Excited mesons in the final state, such as $B^0 \to \pi^- D_s^{*+}$ or $B_s^0 \to D_s^- \rho^+$, can be reconstructed without the photon or $\pi^0 \to \gamma\gamma$, from the decay $D_s^{*+} \to D_s^+\gamma$ or $\rho \to \pi\pi^0$, respectively, thereby missing momentum, leading to a lower reconstructed mass.

Misidentification If one of the four final states particle was misidentified, decays such as $B^0 \to (D^+ \to K^-\pi^+\pi^+)\pi^-$ can end up as backgrounds in our mass window of $B^0{}_{(s)} \to \pi^-(D_s^+ \to K^-K^+\pi^+)$ candidates.

In some cases, a combination of these two effects might occur. For example, $B^0 \to D^+\rho^-$ can be misidentified and partially reconstructed and therefore end up as a background in this analysis.

To incorporate these backgrounds, we need to determine two things: how much of the background events pass the selection (yield), and what distribution the remaining events will have (shape). Both of these are obtained from Monte Carlo(MC) simulations. In this analysis, both 2011 and 2012 MC is used. These are weighted, such that 1/3(2/3) of these backgrounds is obtained from 2011(2012) MC. All signal and background decays are simulated by using Pythia to generate pp collisions with a specific LHCb configuration. Subsequent decays of hadronic particles are described by EvtGen, in which final state radiation is included using Photos. Finally, Geant4 is used to simulate the interactions of these particles with the detector. At the end of these simulations, a simulated dataset is obtained which is very similar to real events recorded by the LHCb detector.

The main difference is that MC data includes so-called truth information: as MC is simulated, it is known which events are caused by which decay process. Thus, a selection can be made which keeps only events from a specific physics background.

We apply the same selection steps on these MC samples as were applied on data (see Subsection 4.2). Subsequently, we obtain both the selection efficiency and a shape, which is parametrised as a so-called template. A non-parametric template, called RooKeys was used. The sharpness, denoted as ρ, of these templates was varied, depending on the background. An example template is shown in Figure 12.

The selection efficiency has to be combined with a number of other variables to estimate the yields of different backgrounds. These include the measured branching ratios for both the B and D meson decays, $\mathcal{B}(D^* \to D\pi^0)$, the production difference between B_s^0 and B^0 quarks f_s/f_d and particle identification efficiencies for both the D meson and bachelor pion. Table 6 shows the resulting relative efficiencies of the different backgrounds.

It is possible that two backgrounds cannot be distinguished very well if they have a similar shape and appear in the same mass region, which occurs for example between $B_s^0 \to D_s^{+*}\pi^-$ and $B_s^0 \to D_s^+\rho^-$. To reduce this correlation, we implement so-called gaussian

Background	Relative efficiency	Yield (gaussian constrained, BDT2)
$B^0 \to (D^+ \to D_s^+)\pi^-$	4.2%	2118
$B^0 \to (D^+ \to D_s^+)\rho^-$	1.5%	760
$B_s^0 \to D_s^+(K^- \to \pi^-)$	0.17%	84
$B_s^0 \to D_s^+\rho^-$	29.7%	15139
$B_s^0 \to D_s^{*+}\pi^-$	45.9%	23364
$B_s^0 \to D_s^{*+}\rho^-$	0.5%	236
$\Lambda_b^0 \to (\Lambda_c^+ \to D_s^+)\pi^-$	0.8%	430

Table 6: A table containing the relative efficiencies for all backgrounds. The presence of ($D^+ \to D_s^+$) indicates that a D^+ has to be misidentified as D_s^+ to become a background.

constraints, which add physics information for every background yield, which is then taken into account while fitting. The uncertainty on these constraints is set to their branching ratio uncertainty as given by the PDG, which is about 20% for the main backgrounds that are included in the final fit.

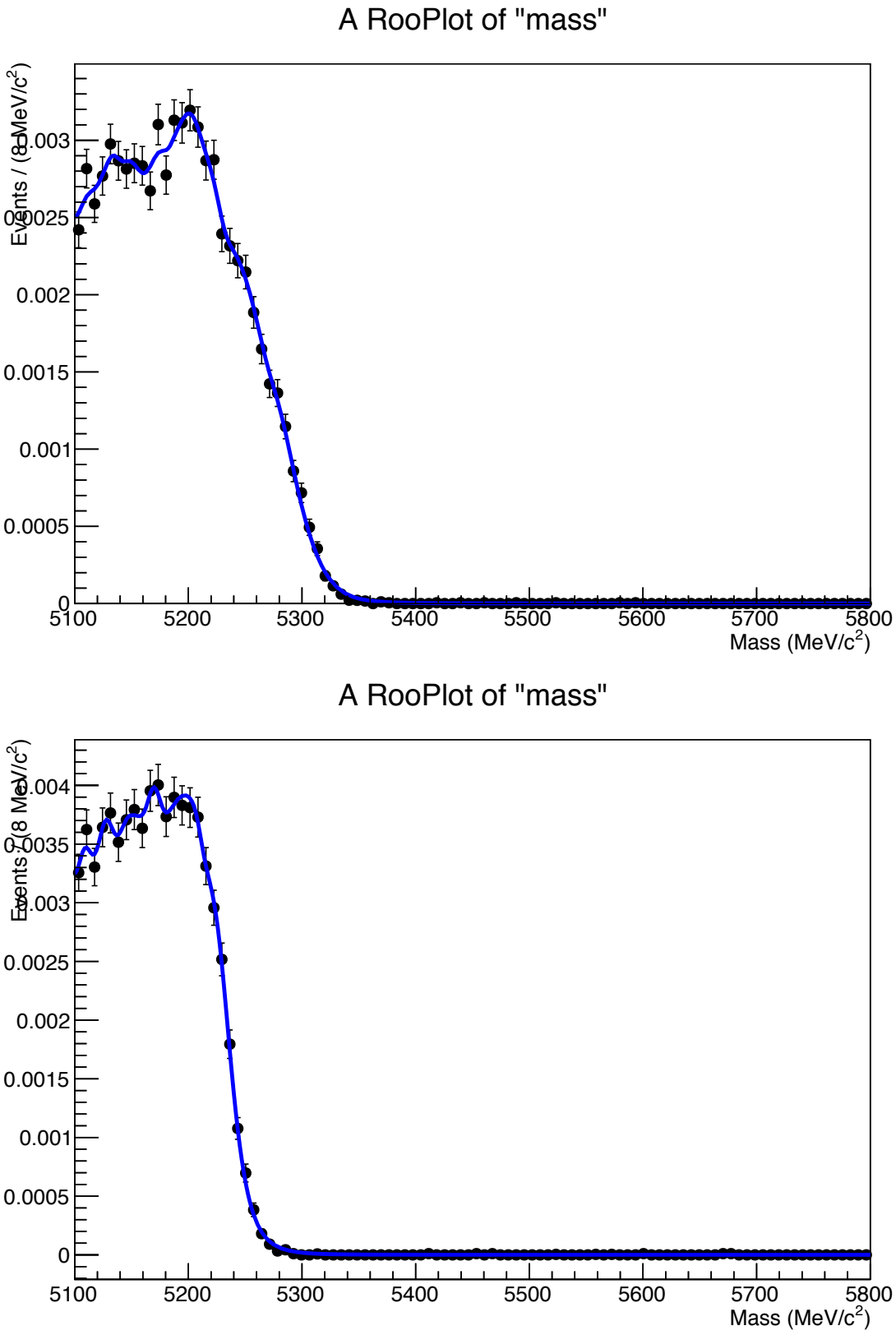

Figure 12: Two examples of background templates for $B_s^0 \to D_s^{*+}\pi^-$ and $B_s^0 \to D_s^+\rho^-$ respectively.

4.5 Signal shape

Apart from modeling background, a shape for the mass distribution of signal events also needs to be chosen. This shape has been determined in the previous analysis [23] to be of the form

$$f(x;\overline{x},\sigma,\alpha_1,\alpha_2,n_1,n_2,\varepsilon) = \\ N\begin{cases} \varepsilon\left(\frac{n_1}{|\alpha_1|}\right)^{n_1}\exp\left(-\frac{|\alpha_1|}{2}\right)\left(\frac{n_1}{|\alpha_1|}-|\alpha_1|-\frac{x-\overline{x}}{\sigma}\right)^{-n_1} & \text{for } \frac{x-\overline{x}}{\sigma}>\alpha_1 \\ (1-\varepsilon)\left(\frac{n_2}{|\alpha_2|}\right)^{n_2}\exp\left(-\frac{|\alpha_2|}{2}\right)\left(\frac{n_2}{|\alpha_2|}-|\alpha_2|-\frac{x-\overline{x}}{\sigma}\right)^{-n_2} & \text{for } \frac{x-\overline{x}}{\sigma}<\alpha_2 \\ \exp\left(-\frac{(x-\overline{x})^2}{2\sigma^2}\right) & \text{otherwise.} \end{cases} \quad (60)$$

This is a so-called Double Crystal Ball function. This consists of a central Gaussian with two exponential tails. This function has seven parameters: the mean $\overline{x}$ and standard deviation σ of the Gaussian, the exponential constants n_1, n_2 of the two tails, and the distance from the mean at which the tails start α_1, α_2 in units of σ. For $n = 1$, the tail has the shape of a radiative photon tail. For large n, a tail behaves exponentially. Each of these parameters is left free in the fit for $B^0 \to D^-\pi^+$, while for $B_s^0 \to D_s^-\pi^+$ and $B^0 \to \pi^- D_s^+$ the tail parameters are fixed to the MC values. Finally, the parameter ε is fixed to 0.5 in the fit. It represents the relative size of the two Crystal Ball functions. The normalisation factor N is implicitly included in the final value for the fitted signal yield. The mean value $\bar{x}$ for the $B^0 \to \pi^- D_s^+$ signal is taken from the fitted value of the $B_s^0 \to D_s^-\pi^+$ candidates, shifted with the known $B_s^0 - B^0$ mass difference.

Parameter	MC $B_s^0 \to D_s^-\pi^+$	MC $B^0 \to D^-\pi^+$	Data $B^0 \to D^-\pi^+$
$\text{Fract}_{\text{Sig}-\text{CBs}}$	0.500000	0.500000	0.500000
α_1	1.60 ± 0.05	1.49 ± 0.05	1.04 ± 0.02
α_2	-1.59 ± 0.07	-1.87 ± 0.08	-1.23 ± 0.01
$\overline{\text{x}}$	5371.53 ± 0.12	5285.03 ± 0.15	5284.24 ± 0.05
n_1	1.70 ± 0.10	1.88 ± 0.13	7.11 ± 0.85
n_2	6.98 ± 1.16	3.45 ± 0.40	139.91 ± 79.16
σ	15.86 ± 0.12	16.69 ± 0.15	17.48 ± 0.05

Table 7: A table showing the resulting fit parameters for signal shapes from MC (for $B_s^0 \to D_s^-\pi^+$ and $B^0 \to D^-\pi^+$) and from data (for $B^0 \to D^-\pi^+$). For the fit to $B^0 \to D^-\pi^+$ in data, n_2 was limited at 140 to prevent numerical issues from occuring. As the size of n does not matter in the limit of large n, this will not affect the fit result.

4.6 Event yield of $B^0 \to \pi^- D_s^+$

The resulting plots for both BDT regions are shown in Figure 13. As is clear from these plots, the low purity BDT region 1 contains only about half the amount of signal and six times more combinatorial background than BDT region 2.

The final fit amounts to 2021 ± 150 signal $B^0 \to \pi^- D_s^+$ decays. Note that the $B^0 \to \pi^- D_s^+$ signal around $m_{B^0} = 5280$ MeV is much smaller compared to the abundant $B_s^0 \to D_s^- \pi^+$ decays around $m_{B_s^0} = 5366$ MeV. Detailed results of this fit are shown in Table 8.

4.7 Event yield of normalisation channel $B^0 \to D^- \pi^+$

To determine $N_{B^0 \to D^- \pi^+}$ and $\varepsilon_{B^0 \to D^- \pi^+}$, we perform a similar fit as for the $B^0 \to \pi^- D_s^+$ candidates. The same selection procedure is followed, backgrounds are modeled similarly, and the same signal parametrisation is used. The main difference is the mass window, which runs from 5000 to 5800 MeV/c^2 for this fit. Additionally, the tail parameters of the signal shape are left free in the fit, while they are fixed for the $B^0 \to D_s^+ \pi^-$ fit.

The statistical and systematic uncertainties are smaller for the $B^0 \to D^- \pi^+$ candidates. The main reason for this is that $\frac{\mathcal{B}(B^0 \to \pi^- D_s^+)}{\mathcal{B}(B^0 \to D^- \pi^+)} \sim 10^{-2}$, and thus a lot more events will be available in the $B^0 \to D^- \pi^+$ fit. The fit results are shown in Figure 14 and Table 9.

Table 8: Fit results for $B_s^0 \to D_s^- \pi^+$ in BDT region 2.

	Magnet Up	Magnet Down
Parameters	Fit Results	Fit Results
$N_{B_s^0 \to D_s^- \pi^+}$	$23\,880 \pm 179$	$26\,436 \pm 190$
$N_{B^0 \to \pi^- D_s^+}$	$1\,034 \pm 103$	988 ± 109
$N_{B_s^0 \to D_s^- \rho^+}$	$8\,317 \pm 522$	$8\,694 \pm 560$
$N_{B_s^0 \to D_s^{*-} \pi^+}$	$12\,994 \pm 535$	$15\,496 \pm 571$
$N_{B_s^0 \to D_s^{*-} \rho^+}$	104 ± 23	111 ± 25
$N_{B^0 \to \pi^- D_s^{*+}}$	148 ± 28	162 ± 31
$N_{B^0 \to D^- \pi^+}$	$1\,210 \pm 83$	$1\,477 \pm 91$
$N_{B^0 \to D^- \rho^+}$	366 ± 36	401 ± 40
$N_{B_s^0 \to D_s^{\mp} K^{\pm}}$	41 ± 4	45 ± 4
$N_{\Lambda_b^0 \to \Lambda_c^+ \pi^-}$	197 ± 19	234 ± 20
p_0	0.00000 ± 0.00893	0.00000 ± 0.00473
p_1	-0.00785 ± 0.00037	-0.00828 ± 0.00036
$N_{\text{combinatorial}}$	$3\,363 \pm 417$	$3\,544 \pm 429$
Common Parameters		
$\overline{x}$	$5\,371.25 \pm 0.10$	
σ	17.64 ± 0.09	
$\text{mean}_{B^0 \to \pi^- D_s^+}$	$5\,284.2 \pm 0.0$	
σ	17.64 ± 0.09	
Fixed Parameters		
α_1	1.5966	
α_2	-1.5879	
n_1	1.70	
n_2	6.98	
$\text{Fract}_{\text{Sig}-\text{CBs}}$	0.50	
$\alpha_1(B^0 \to \pi^- D_s^+)$	1.5966	
$\alpha_1(B^0 \to \pi^- D_s^+)$	-1.5879	
$n_1(B^0 \to \pi^- D_s^+)$	1.70	
$n_2(B^0 \to \pi^- D_s^+)$	6.98	
$\text{Fract}_{\text{Sig}-\text{CBs}}(B^0 \to \pi^- D_s^+)$	0.50	

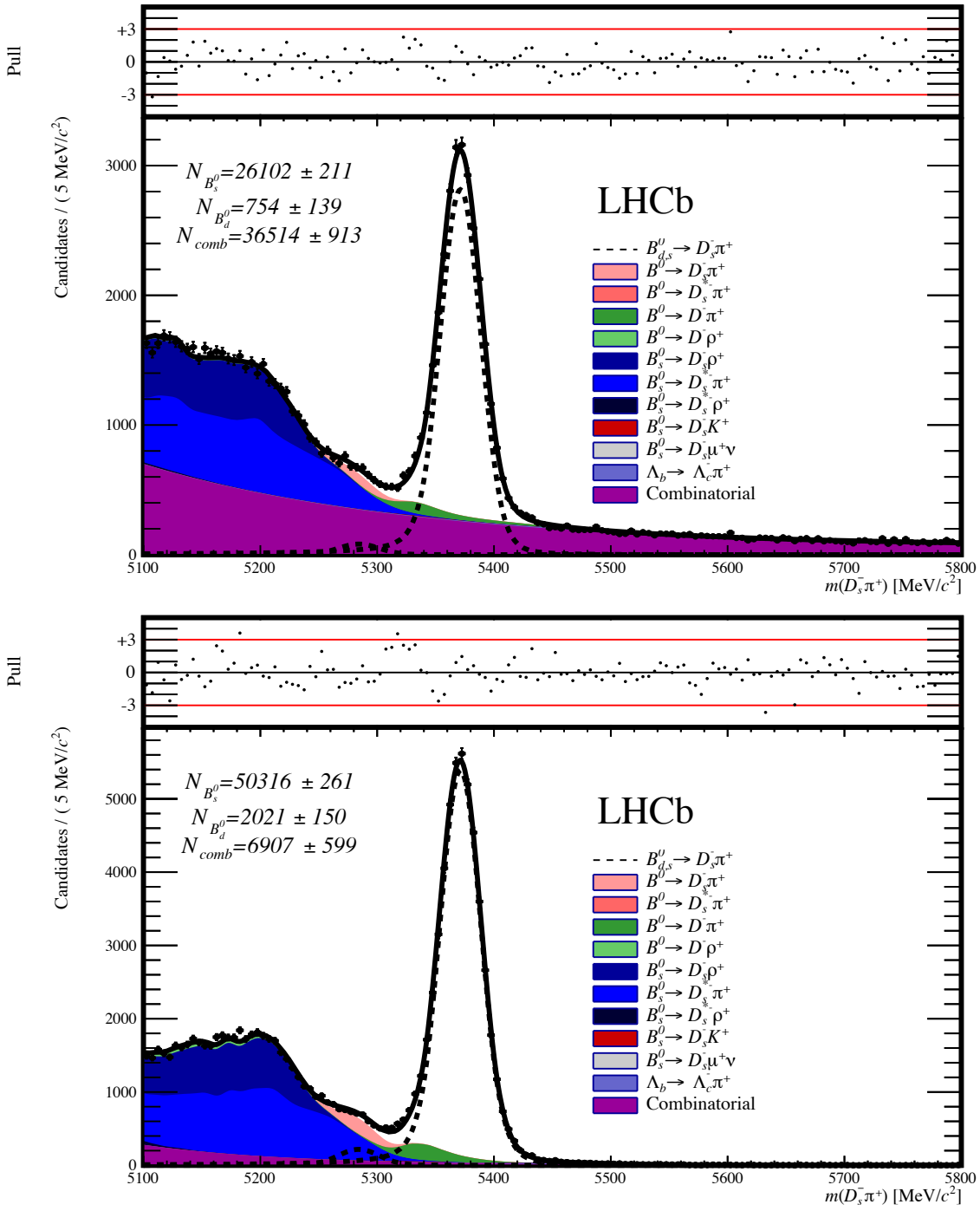

Figure 13: The fit results for BDT region 1 (top) and region 2 (bottom), for $(\pi^- D_s^+)$ candidates. The small B^0 signal is shown around 5280 MeV.

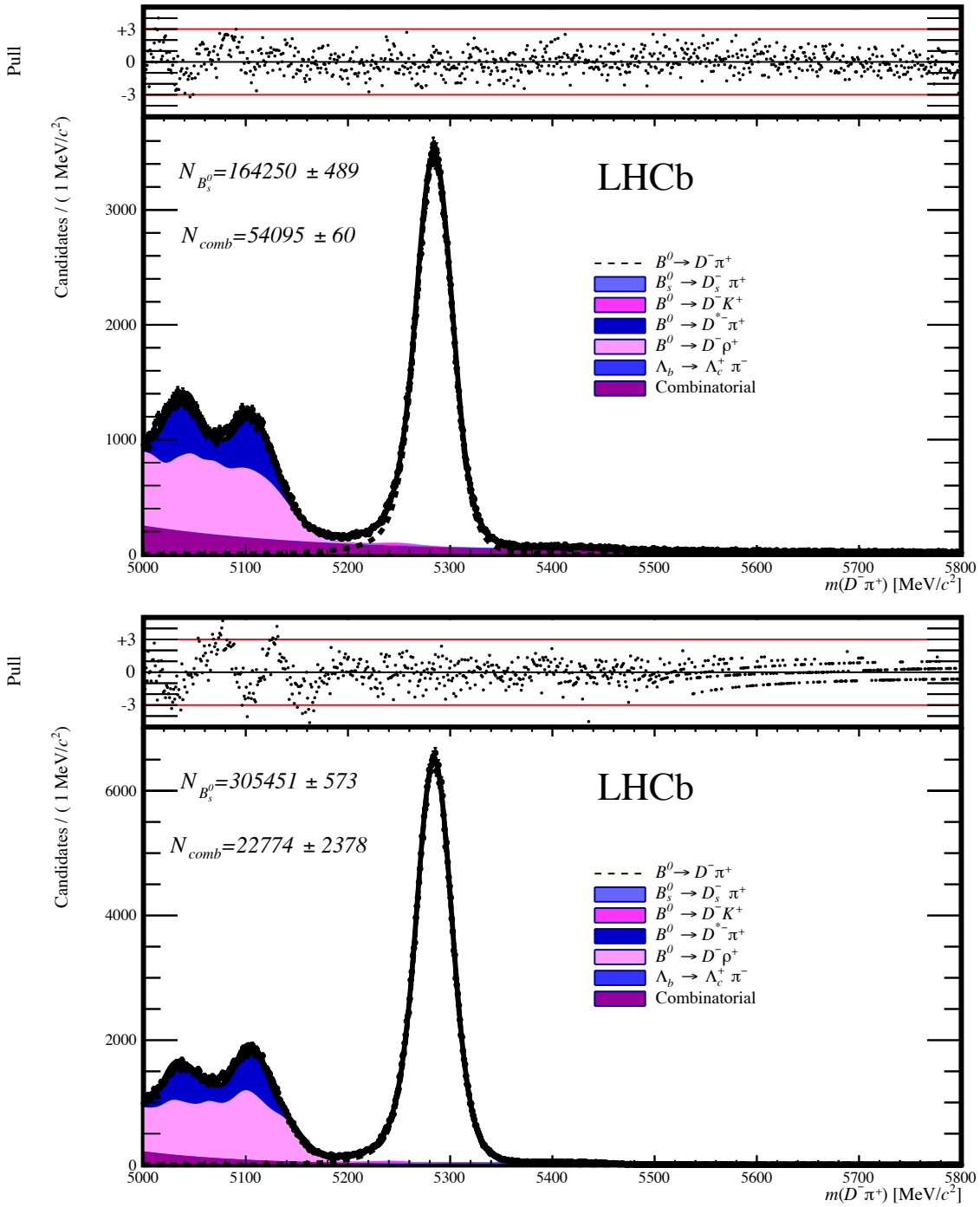

Figure 14: The fit results for BDT region 1 (top) and region 2 (bottom), for $B^0 \to D^-\pi^+$.

Table 9: Fit results for $B^0 \to D^- \pi^+$, BDT region 2

	Magnet Up	Magnet Down
Parameters	Fit Results	Fit Results
$N_{B^0 \to D^- \pi^+}$	$144\,425 \pm 435$	$160\,386 \pm 440$
$N_{B^0 \to D^- K^+}$	$1\,192 \pm 129$	$1\,533 \pm 60$
$N_{B^0 \to D^- \rho^+}$	$64\,662 \pm 121$	$71\,901 \pm 103$
$N_{B^0 \to D^{*-} \pi^+}$	$28\,491 \pm 442$	$31\,997 \pm 467$
$N_{\Lambda_b^0 \to \Lambda_c^+ \pi^-}$	$2\,513 \pm 80$	$2\,841 \pm 79$
p_0	0.01992 ± 0.00422	0.01975 ± 0.00341
p_1	-0.00959 ± 0.00031	-0.00998 ± 0.00665
$N_{combinatorial}$	$11\,215 \pm 545$	$12\,029 \pm 563$
Common Parameters		
$\overline{x}$	$5\,284.24 \pm 0.05$	
σ	17.48 ± 0.05	
α_1	1.0360 ± 0.0204	
α_2	-1.2345 ± 0.0128	
n_1	7.11 ± 0.85	
n_2	139.91 ± 79.16	
Fixed Parameters		
$Fract_{Sig-CBs}$	0.50	

5 Systematic uncertainties

After implementing the analysis strategy and obtaining the signal yield and efficiency, it is important to investigate both the selection procedure and fit model for systematic effects, which will result in an uncertainty on the efficiency and yield, respectively. The total systematic uncertainty is shown in Table 11.

5.1 Selection efficiency from hardware (L0) trigger

One of the selection efficiency uncertainties enters while taking data. As there is a limited amount of information, and only triggered data is stored, an uncertainty on the efficiency of the hardware (also known as L0) trigger is introduced. This manifests itself as a difference in the detection efficiency between pions and kaons. The hardware trigger efficiency has been measured as a function of p_T for pions and kaons by the L0 trigger group (see Ref. [24]). On average, an efficiency difference of 2% is found, which is not implemented in simulation. As the ratio of $B^0 \to D_s^+\pi^-$ and $B^0 \to D^+\pi^-$ decays is investigated, only a difference in triggering on D mesons, which occurs for about 10% of the events (see Ref. [25]), will influence the final results. Therefore, a systematic uncertainty of 0.2% is assigned to the final result.

5.2 Offline selection efficiency

The subsequent offline selection is determined from MC simulations, as detailed in Section 4.2. To estimate the uncertainty introduced by this method of determining the efficiency, two methods are used.

First, the uncertainty on the efficiency of the BDT is studied, as this is the main uncertainty during selection. This uncertainty is determined by comparing data with simulation for each of the BDT variables. For two variables, a significant difference is found: the vertex χ^2/ndf of the D_s^+ meson and the ghost probability of the tracks of D_s^+ meson daughters. The simulation is then reweighted to account for these differences and thus improve the consistency between simulation and data. After this procedure, the BDT efficiency is recalculated for the reweighted simulation samples. The difference with the previous BDT efficiency is used as a systematic uncertainty on the BDT. As reweighting with the ghost track probability leads to the largest uncertainty, the reweighting in this variable is used to determine the uncertainty. An uncertainty of 2.8% is assigned for $B^0 \to D^-\pi^+$, and an uncertainty of 0.7

5.3 PID selection efficiency

The uncertainty from the PID selection efficiency has been determined for bachelor kaons and pions as well as for $D^{\pm}_{(s)}$ daughters in Ref. [26]. The PID efficiencies from simulation are compared in two ways. First, the PID efficiency is determined just as in section 4.2, but with the $D^{*\pm}$ samples obtained from MC instead of data. This makes it possible

to determine an efficiency $\varepsilon_{\text{tot,MC}}$. Secondly, the PID information is obtained from the same samples by cutting directly on the PID variables and using the truth information to determine an efficiency $\varepsilon_{\text{dir,MC}}$. The relative difference between these two is multiplied with the PID efficiency determined from data (and used in the analysis) to obtain the systematic uncertainty from the PID selection

$$\sigma_{\text{PID}} = 2\varepsilon_{\text{tot,data}} \frac{\varepsilon_{\text{tot,MC}} - \varepsilon_{\text{dir,MC}}}{\varepsilon_{\text{tot,MC}} + \varepsilon_{\text{dir,MC}}} \tag{61}$$

As a ratio of $B^0 \to \pi^- D_s^+$ and $B^0 \to D^- \pi^+$ will be investigated for this analysis, only the pion or kaon which differentiates a D^+ from a D_s^+ decay is taken into account. An uncertainty of 1.1% is assigned on the ratio of $B^0 \to \pi^- D_s^+$ and $B^0 \to D^- \pi^+$ decays.

5.4 Fit model uncertainties

After checking systematic uncertainties on the selection procedure, the fitting model needs to be investigated for possible systematic effects. As the decay $B^0 \to D_s^+ \pi^-$ is relatively rare compared to $B_s^0 \to D_s^{+(*)} \pi^{-(*)}$ decays, systematic uncertainties from these backgrounds and combinatorial backgrounds need to be investigated. The main two systematic uncertainties originate from the shape and yield of both combinatorial and physics backgrounds. In addition, the effect of fixing the tail parameters of the signal shapes is investigated.

5.4.1 Combinatorial background

A systematic uncertainty is assigned by fitting with different combinatorial shapes. The constant+exponential function used in the fit is compared with an exponential function.It turns out that the combinatorial does not contain a constant component. Therefore, this systematic check does not change the $B^0 \to \pi^- D_s^+$ yield.

The resulting fit for the chosen selection region is shown in Figure 19 as part of Appendix A. In addition to the resulting signal yields, the yields of partially reconstructed and/or misID backgrounds are examined. These are compared with the nominal fit. The results are shown in Table 14 as part of Appendix B, and are very consistent.

No systematic uncertainty is assigned, as the yields in the fit do not change.

5.4.2 Template shape of partially reconstructed and/or misidentified backgrounds

A free parameter, called ρ, defines the coarseness of the shape of the background templates. As this parameter can be chosen freely, one systematic check is to vary this parameter and investigate the resulting fit.The bounds on this variation were chosen by examining at what value of ρ point the template poorly described the mass distribution in MC. The resulting bounds where chosen to be $1.5 * \rho_{\text{nominal}}$ and $0.5\rho_{\text{nominal}}$. These correspond to a smooth and sharp template, respectively. In Figure 15, the resulting templates are shown for $B_s^0 \to D_s^{*+} \pi^-$.

The resulting fits for the chosen selection region are shown in Figure 17 as part of Appendix A. The results are shown in Tables 15 and 16 as part of Appendix B. These results are consistent with the nominal fit. A systematic uncertainty of $^{+4.3}_{-1.2}\%$ is assigned.

5.4.3 Exclusion of small backgrounds

To include another check of the modeling of partially reconstructed and/or misidentified backgrounds, a fit was performed with all but the main backgrounds ($B_s^0 \to D_s^{*-}\pi^+$, $B_s^0 \to D_s^-\rho^+$, $B^0 \to D^-\pi^+$, and $\Lambda_b^0 \to \Lambda_c^+\pi^-$) excluded. As a result of this fit, the yield changed by 0.1%, which is negligible compared to other systematic uncertainties. Therefore, this change in yield is not taken into account for the systematics.

5.4.4 Signal tail parameters

Finally, the modeling of the signal shape might lead to a systematic uncertainty on the yield. For $B^0 \to D^-\pi^+$, this effect is investigated by fixing the shape parameters α, η in the $B^0 \to D^-\pi^+$ fit to the MC shape and comparing the resulting yield with the nominal fit, which leaves these parameters free. This difference is taken as a systematic uncertainty on the $B^0 \to D^-\pi^+$ yield of 1.6%. Assuming the same differences between data and simulation for $B^0 \to \pi^- D_s^+$ and $B_s^0 \to D_s^-\pi^+$, the tail parameters are scaled using:

$$p_{\mathrm{data},D_s\pi} = \frac{p_{\mathrm{data},D\pi}}{p_{\mathrm{MC},D\pi}} p_{\mathrm{MC},D_s\pi} \tag{62}$$

where $p \in \{\alpha, n\}$. As the yield changes by 1.5%, this is assigned as an uncertainty on the final result.

5.5 Consistency check with BDT region 1

To check the consistency of our method, it is investigated the same systematic uncertainties as discussed for BDT region 2 are investigated for BDT region 1. It turns out that fit model systematics dominate in this case, as shown in Table 12. The resulting relative yields and efficiencies are shown in Table 10.The yields and efficiencies obtained for BDT region 1 and 2 (as shown in Table 13 in Section 6) are quite consistent.

$N_{B^0\to\pi^-D_s^+}/N_{B^0\to D^-\pi^+}$	$(4.56 \pm 0.94\ (\mathrm{stat.}) \pm 1.61\ (\mathrm{syst.}))\times 10^{-3}$
$\varepsilon_{B^0\to D^-\pi^+}/\varepsilon_{B^0\to\pi^-D_s^+}$	$1.078 \pm 0.034\ (\mathrm{stat.}) \pm 0.023\ (\mathrm{syst.})$

Table 10: The statistical and systematic uncertainties for the ratio of $B^0 \to D^-\pi^+$ and $B^0 \to \pi^- D_s^+$ for signal yield and efficiency.

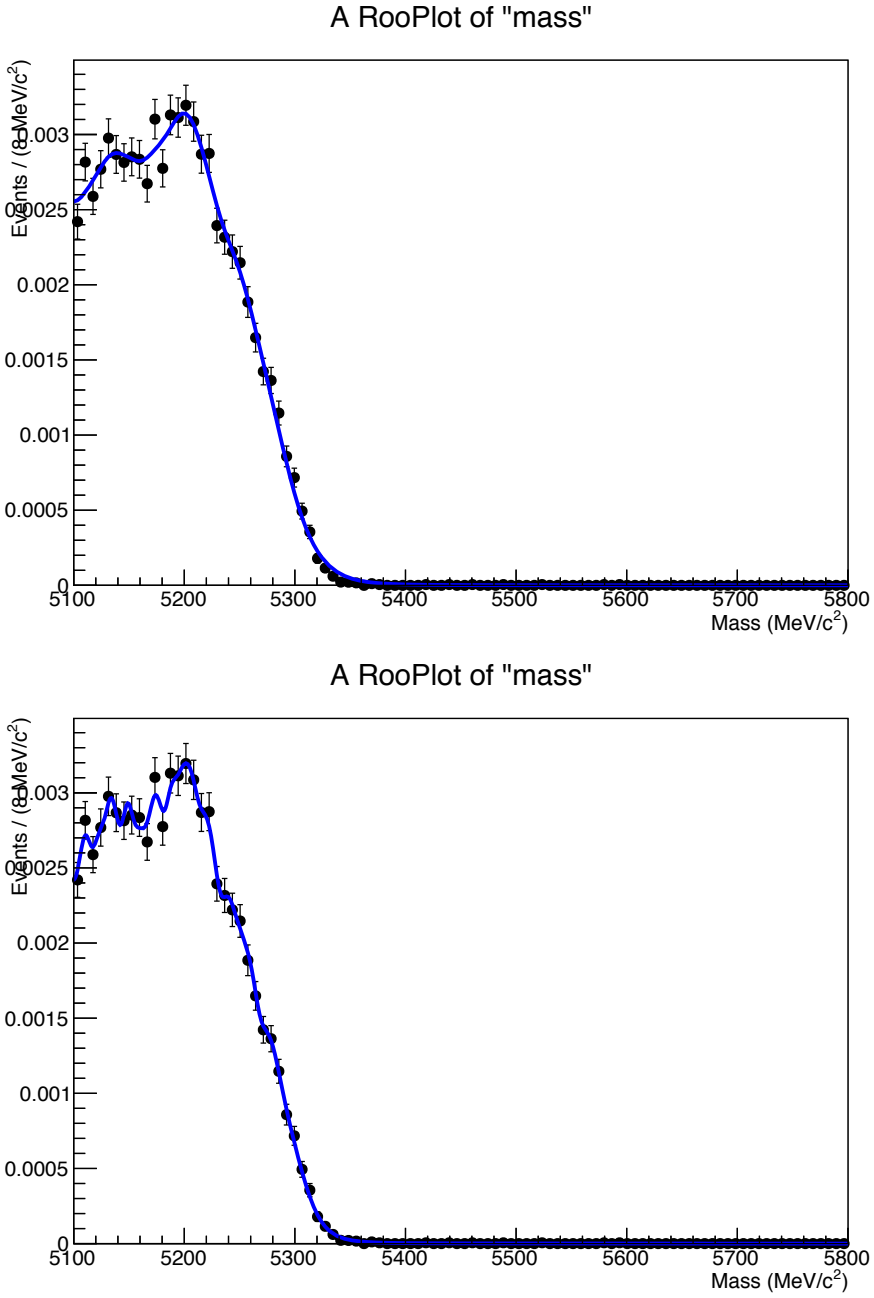

Figure 15: The two templates used to estimate the systematic uncertainty due to shape parametrisation for $B_s^0 \to D_s^{*-}\pi^+$. These were obtained with $\rho = 1.5\rho_{\text{nominal}}$ and $\rho = 0.5\rho_{\text{nominal}}$ respectively.

Systematic uncertainty	$B^0 \to D^-\pi^+$	$B^0 \to \pi^- D_s^+$	$\frac{B^0 \to \pi^- D_s^+}{B^0 \to D^-\pi^+}$
Hardware selection efficiency	1.2	1.2	0.2
BDT selection efficiency	2.8	0.7	2.1
PID selection efficiency	1.3	1.4	1.1
Total efficiency	3.3	2.0	2.4
Combinatorial shape	0.1	0	0
Background templates	0.42	$^{+4.4}_{-1.2}$	$^{+4.4}_{-1.2}$
Tail parameters	1.6	1.5	1.6
Total fit model	1.6	4.6	4.8
Total	3.7	5.0	5.4

Table 11: The systematic uncertainties for $B^0 \to D^-\pi^+$ and $B^0 \to \pi^- D_s^+$ as percentage of signal yield change for BDT region 2. The largest of the two uncertainties on background templates is taken for the calculation of the total systematic uncertainty.

Systematic uncertainty	$B^0 \to D^-\pi^+$	$B^0 \to \pi^- D_s^+$	$\frac{B^0 \to \pi^- D_s^+}{B^0 \to D^-\pi^+}$
Hardware selection efficiency	1.2	1.2	0.2
BDT selection efficiency	2.8	0.7	2.1
PID selection efficiency	1.3	1.4	1.1
Total efficiency	3.3	2.0	2.1
Combinatorial shape	1.5	27.1	27.1
Background templates	$^{+1.2}_{-0.3}$	$^{+3.7}_{-7.1}$	$^{+3.9}_{-7.1}$
Tail parameters	0.8	14.3	14.3
Total fit model	2.1	31.6	31.6
Total	3.9	31.7	31.7

Table 12: The systematic uncertainties for $B^0 \to D^-\pi^+$ and $B^0 \to \pi^- D_s^+$ as percentage of signal yield change for BDT region 1. The largest of the two uncertainties on background templates is taken for the calculation of the total systematic uncertainty.

6 Results

From the reported yields and efficiencies for $B^0 \to \pi^- D_s^+$ and $B^0 \to D^- \pi^+$, it is possible to determine both the ratio of branching fractions of these decays and the branching fraction of $B^0 \to \pi^- D_s^+$. This is the current best (individual) measurement. Additionally, this is the first reported measurement of $B^0 \to \pi^- D_s^+$ at a hadronic collider experiment.

The efficiencies which are used to calculate these branching fractions are a weighted average of the efficiencies obtained from simulations under 2011 (1/3) and 2012 (2/3) conditions, to reflect the composition of data.

In Section 4.1, it is discussed how to determine the ratio of branching fractions and the branching fraction of $B^0 \to \pi^- D_s^+$. As input, the numbers shown in Table 13 are used. It is found that

$$\frac{\mathcal{B}(B^0 \to \pi^- D_s^+)}{\mathcal{B}(B^0 \to D^- \pi^+)} = (9.95 \pm 0.73(\text{stat.}) \pm 0.51(\text{syst.}) \pm 0.33(\mathcal{B})) \times 10^{-3} \tag{63}$$

where the first uncertainty is statistical, the second systematic, and the third comes from the branching fractions uncertainties from $\mathcal{B}(D^+ \to K^- \pi^+ \pi^+)$ and $\mathcal{B}(D_s^+ \to K^+ K^- \pi^+)$. The branching fraction then becomes

$$\mathcal{B}(B^0 \to \pi^- D_s^+) = (26.7 \pm 2.0(\text{stat.}) \pm 1.5(\text{syst.}) \pm 1.6(\mathcal{B})) \times 10^{-6} \tag{64}$$

where the branching fraction uncertainties come from both D decays and $\mathcal{B}$ $(B^0 \to D^- \pi^+)$. This result is the single most precise measurement and consistent with the current world average from the PDG,

$$\mathcal{B}(B^0 \to \pi^- D_s^+) = (21.6 \pm 2.6) \times 10^{-6}. \tag{65}$$

The new world average is determined, using $1/\sigma^2$ as weights, to be

$$\mathcal{B}(B^0 \to \pi^- D_s^+) = (24.0 \pm 1.9) \times 10^{-6}. \tag{66}$$

The relative uncertainty on $\mathcal{B}(B^0 \to \pi^- D_s^+)$ is found to decrease from 12.0% to 7.9% by including our measurement.

$N_{B^0 \to \pi^- D_s^+} / N_{B^0 \to D^- \pi^+}$	$(6.63 \pm 0.48$ (stat.) ± 0.32 (syst.)$) \times 10^{-3}$
$\varepsilon_{B^0 \to D^- \pi^+} / \varepsilon_{B^0 \to \pi^- D_s^+}$	0.891 ± 0.015 (stat.) ± 0.019 (syst.)

Table 13: The statistical and systematic uncertainties for the ratio of $B^0 \to D^- \pi^+$ and $B^0 \to \pi^- D_s^+$ for signal yield and efficiency.

From these calculated branching fractions, it is possible to calculate the combined result for $|V_{ub}||a_{NF}|$, using the method discussed in Section 2. For current data, it is found that

$$|V_{ub}||a_{NF}| = (3.24 \pm 0.19(\mathcal{B}) \pm 0.27(\text{ext.})) \times 10^{-3}, \tag{67}$$

where the first uncertainty is due to $\mathcal{B}(B^0 \to \pi^- D_s^+)$, and the second uncertainty comes from $F(B^0 \to \pi)$, f_{D_s}, and $|V_{cs}|$, with the form factor dominating this uncertainty. Using the new world average as computed in Section 6, it is found that

$$|V_{ub}||a_{NF}| = (3.40 \pm 0.14(\mathcal{B}) \pm 0.29(\text{ext.})) \times 10^{-3}, \tag{68}$$

This result can be compared with the different $|V_{ub}|$ measurements discussed in Section 2. Interestingly, if $|a_{NF}| = 1$, this measurement is consistent with the exclusive measurement:

$$|V_{ub}| = (3.28 \pm 0.29) \times 10^{-3} \text{ (exclusive)}. \tag{69}$$

6.1 Interpretation of $|V_{ub}|$ or $|a_{NF}|$

A number of possible values for $|V_{ub}|$ and $|a_{NF}|$ will now be examined. First, the exclusive, inclusive and average results for $|V_{ub}|$ as given in Section 2 can be used to quantify non-factorisable effects $|a_{NF}|$. As a result, it is found that

$$|a_{NF}| = 0.77 \pm 0.03(\mathcal{B}) \pm 0.04(|V_{ub}|) \pm 0.06(\text{ext.})(\text{inclusive}), \tag{70}$$

$$|a_{NF}| = 1.03 \pm 0.04(\mathcal{B}) \pm 0.09(|V_{ub}|) \pm 0.09(\text{ext.})(\text{exclusive}), \tag{71}$$

$$|a_{NF}| = 0.82 \pm 0.03(\mathcal{B}) \pm 0.10(|V_{ub}|) \pm 0.07(\text{ext.})(\text{average}), \tag{72}$$

where the external uncertainties are the same as described for $|V_{ub}||a_{NF}|$. The form factor for $B^0 \to \pi$ remains the dominant of these uncertainties. All of these results are reasonably consistent with naive factorisation.

Secondly, the argument can be turned around, and estimates for $|a_{NF}|$ can be used to determined $|V_{ub}|$. For $|a_{NF}|$, the results for DD and DD_s as given in Section 2 are used. In addition, we examine $|V_{ub}|$ for a naive factorisation assumption $|a_{NF}| = 1.0 \pm 0.2$. It is found that:

$$|V_{ub}| = 3.40 \pm 0.74 \text{ (naive factorisation)}, \tag{73}$$

$$|V_{ub}| = 2.50 \pm 0.70 \text{ (from } B^0 \to D^- D_s^+), \tag{74}$$

$$|V_{ub}| = 3.07 \pm 0.98 \text{ (from } B^0 \to D^- D^+), \tag{75}$$

where the uncertainties are so large that sensitivity to $|V_{ub}|$ is negligible.

7 Conclusion

As a tree decay, $\mathcal{B}\ (B^0 \to \pi^- D_s^+)$ probes the $b \to u$ quark transition $|V_{ub}|$ very cleanly. However, as it is purely non-leptonic, it is also sensitive to non-factorisable effects in the final state $|a_{NF}|$. $B \to D\bar{D}$ decays are discussed as a possible probe of these non-factorisable effects, taking loop diagrams into account.

Subsequently, a new measurement of $\mathcal{B}\ (B^0 \to \pi^- D_s^+)$ at LHCb is presented. It is found that

$$\mathcal{B}(B^0 \to \pi^- D_s^+) = (26.7 \pm 2.0\ (\text{stat.}) \pm 1.5\ (\text{syst.}) \pm 1.6\ (\mathcal{B}\text{'s})) \times 10^{-6}, \tag{76}$$

which is the world's best single measurement of $\mathcal{B}\ (B^0 \to \pi^- D_s^+)$, and consistent with the world average. After calculating a new world average for $\mathcal{B}\ (B^0 \to \pi^- D_s^+)$, $|V_{ub}||a_{NF}|$ is determined to be

$$|V_{ub}||a_{NF}| = (3.40 \pm 0.14(\mathcal{B}) \pm 0.29(\text{ext.})) \times 10^{-3} \tag{77}$$

which is consistent with naive factorisation if the current exclusive measurement of $|V_{ub}|$ is used. Finally, interpretations of this result in terms of $|V_{ub}|$ and $|a_{NF}|$ are presented.

Appendices

A Systematic uncertainty plots

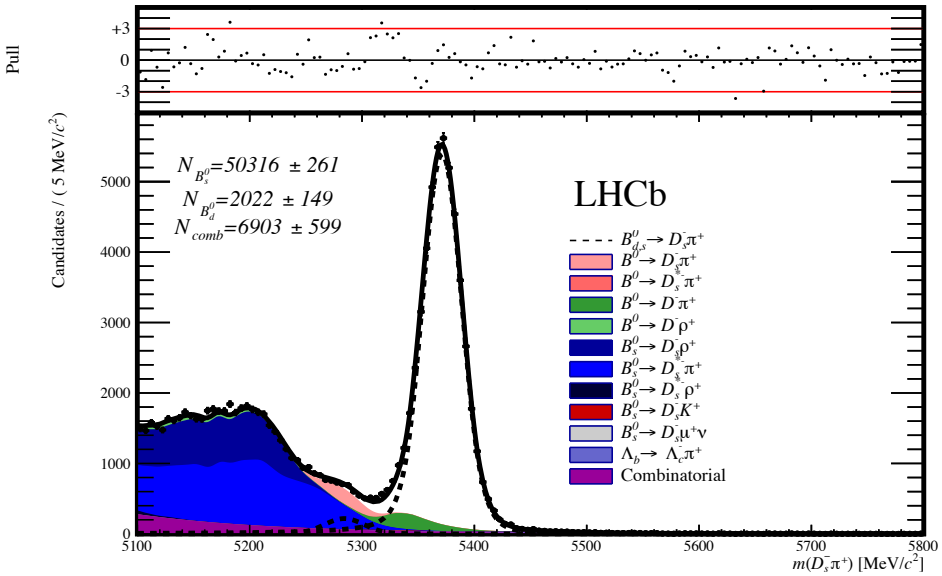

Figure 16: The fit for $B^0 \to \pi^- D_s^+$ with an exponential model for the combinatorial background.

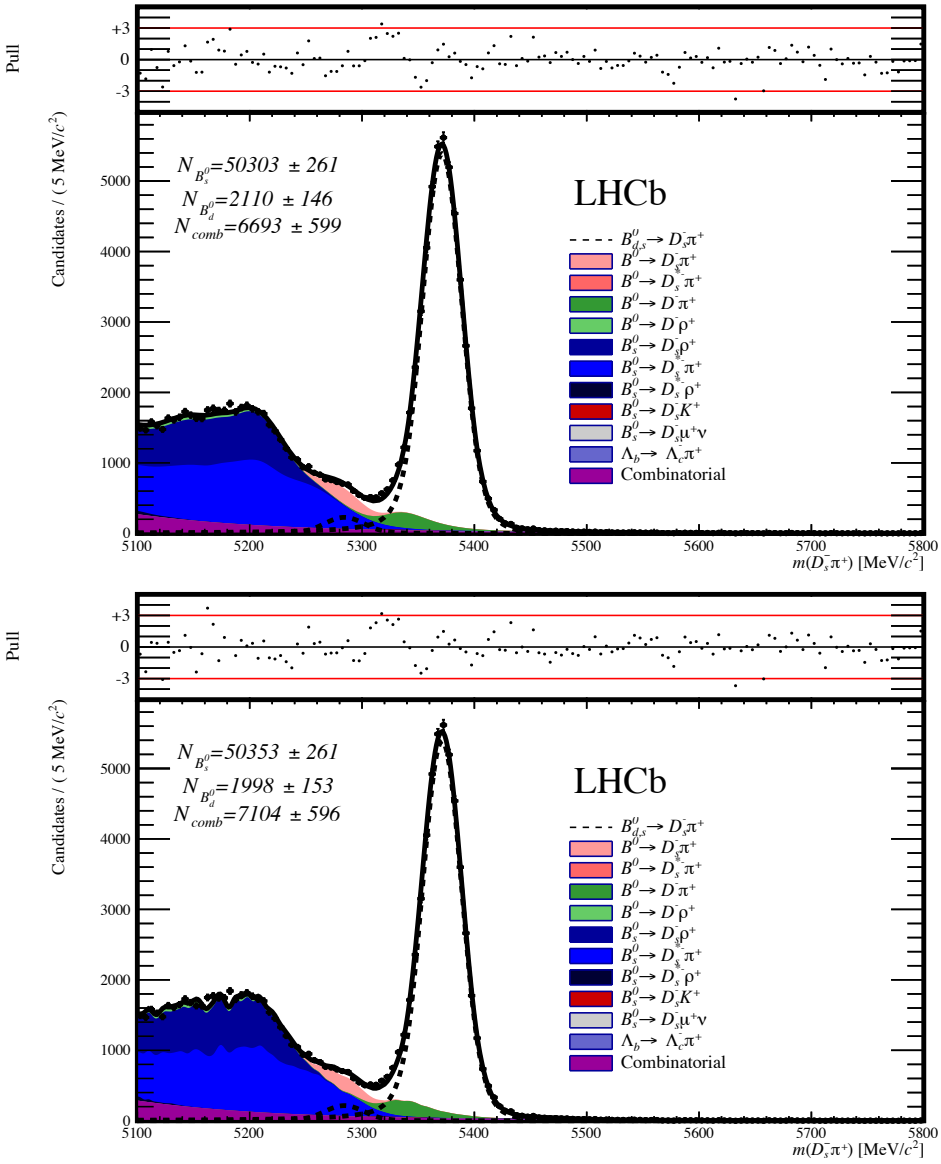

Figure 17: The fit for $B^0 \to \pi^- D_s^+$ with smoother and sharper templates, respectively.

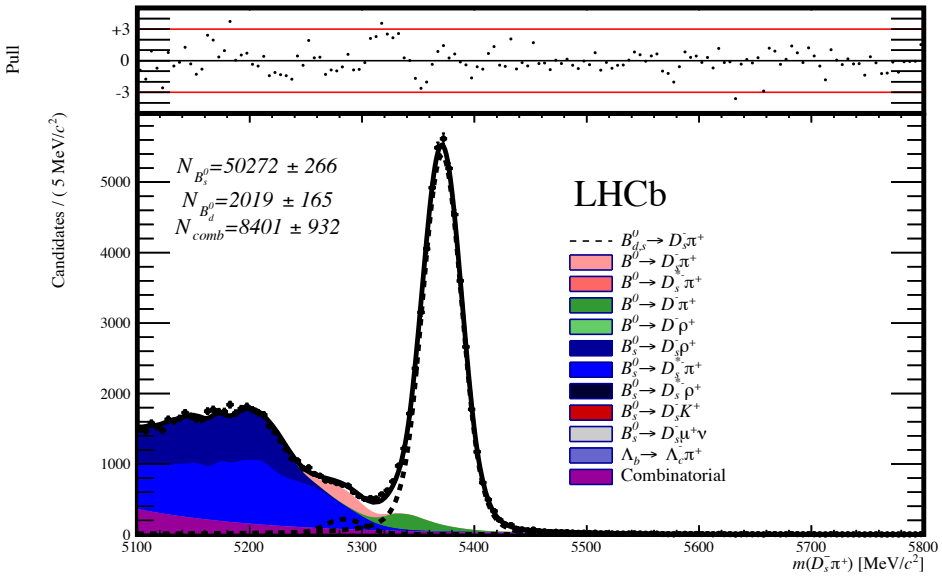

Figure 18: The fit for $B^0 \to \pi^- D_s^+$ with only $B_s^0 \to D_s^- \rho^+$, $B_s^0 \to D_s^{*-} \pi^+$, $B^0 \to D^- \pi^+$ and $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ as backgrounds.

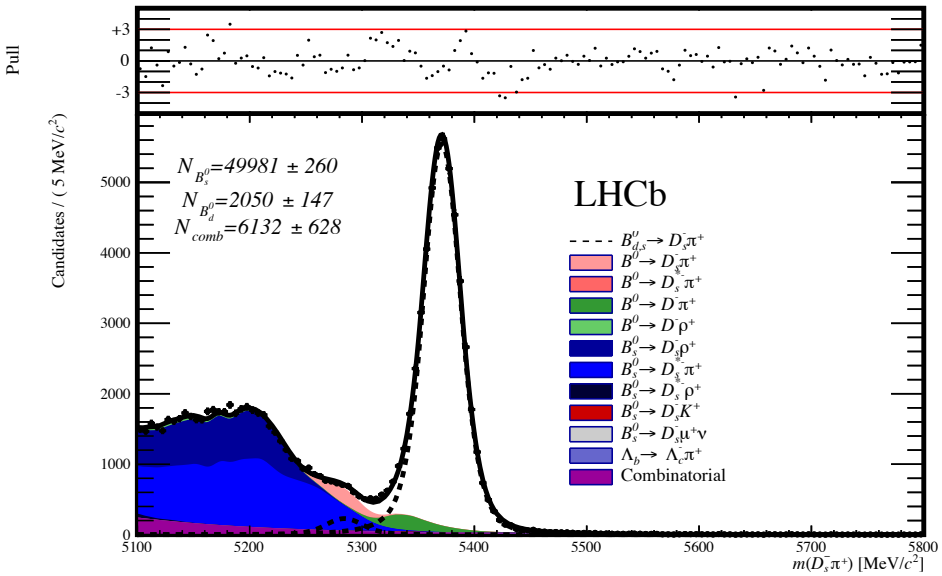

Figure 19: The fit for $B^0 \to \pi^- D_s^+$ with scaled tail parameters.

B Systematic uncertainty tables

Table 14: Fit results for $B^0 \to \pi^- D_s^+$ in BDT region 2 with an exponential combinatorial background.

	Magnet Up	Magnet Down
Parameters	Fit Results	Fit Results
$N_{B_s^0 \to D_s^- \pi^+}$	$23\,881 \pm 179$	$26\,435 \pm 190$
$N_{B^0 \to \pi^- D_s^+}$	$1\,034 \pm 103$	988 ± 108
$N_{B_s^0 \to D_s^- \rho^+}$	$8\,323 \pm 522$	$8\,690 \pm 560$
$N_{B_s^0 \to D_s^{*-} \pi^+}$	$12\,997 \pm 535$	$15\,493 \pm 571$
$N_{B_s^0 \to D_s^{*-} \rho^+}$	104 ± 23	111 ± 25
$N_{B^0 \to \pi^- D_s^{*+}}$	148 ± 28	162 ± 31
$N_{B^0 \to D^- \pi^+}$	$1\,211 \pm 83$	$1\,477 \pm 91$
$N_{B^0 \to D^- \rho^+}$	366 ± 36	401 ± 40
$N_{B_s^0 \to D_s^{\mp} K^{\pm}}$	41 ± 4	45 ± 4
$N_{\Lambda_b^0 \to \Lambda_c^+ \pi^-}$	197 ± 19	234 ± 20
p_1	-0.00784 ± 0.00036	-0.00828 ± 0.00035
$N_{\text{combinatorial}}$	$3\,352 \pm 418$	$3\,551 \pm 430$
Common Parameters		
$\overline{x}$	$5\,371.25 \pm 0.10$	
σ	17.64 ± 0.09	
$\text{mean}_{B^0 \to \pi^- D_s^+}$	$5\,284.2 \pm 0.0$	
σ	17.64 ± 0.09	
Fixed Parameters		
α_1	1.5966	
α_2	-1.5879	
n_1	1.70	
n_2	6.98	
$\text{Fract}_{\text{Sig}-\text{CBs}}$	0.50	
$\alpha_1(B^0 \to \pi^- D_s^+)$	1.5966	
$\alpha_1(B^0 \to \pi^- D_s^+)$	-1.5879	
$n_1(B^0 \to \pi^- D_s^+)$	1.70	
$n_2(B^0 \to \pi^- D_s^+)$	6.98	
$\text{Fract}_{\text{Sig}-\text{CBs}}(B^0 \to \pi^- D_s^+)$	0.50	

Table 15: Fit results for $B_s^0 \to D_s^- \pi^+$ in BDT region 2 with smoother MC shapes.

	Magnet Up	Magnet Down
Parameters	Fit Results	Fit Results
$N_{B_s^0 \to D_s^- \pi^+}$	$23\,876 \pm 179$	$26\,427 \pm 190$
$N_{B^0 \to \pi^- D_s^+}$	$1\,073 \pm 100$	$1\,037 \pm 106$
$N_{B_s^0 \to D_s^- \rho^+}$	$8\,447 \pm 525$	$8\,847 \pm 562$
$N_{B_s^0 \to D_s^{*-} \pi^+}$	$12\,948 \pm 533$	$15\,407 \pm 568$
$N_{B_s^0 \to D_s^{*-} \rho^+}$	103 ± 23	111 ± 25
$N_{B^0 \to \pi^- D_s^{*+}}$	147 ± 28	162 ± 31
$N_{B^0 \to D^- \pi^+}$	$1\,202 \pm 84$	$1\,467 \pm 91$
$N_{B^0 \to D^- \rho^+}$	365 ± 36	400 ± 40
$N_{B_s^0 \to D_s^{\mp} K^{\pm}}$	41 ± 4	45 ± 4
$N_{\Lambda_b^0 \to \Lambda_c^+ \pi^-}$	204 ± 19	237 ± 21
p_0	0.00000 ± 0.00837	0.00000 ± 0.00404
p_1	-0.00788 ± 0.00038	-0.00830 ± 0.00036
$N_{\text{combinatorial}}$	$3\,244 \pm 418$	$3\,449 \pm 429$
Common Parameters		
$\overline{x}$	$5\,371.27 \pm 0.10$	
σ	17.64 ± 0.09	
$\text{mean}_{B^0 \to \pi^- D_s^+}$	$5\,284.2 \pm 0.0$	
σ	17.64 ± 0.09	
Fixed Parameters		
α_1	1.5966	
α_2	-1.5879	
n_1	1.70	
n_2	6.98	
$\text{Fract}_{\text{Sig}-\text{CBs}}$	0.50	
$\alpha_1(B^0 \to \pi^- D_s^+)$	1.5966	
$\alpha_1(B^0 \to \pi^- D_s^+)$	-1.5879	
$n_1(B^0 \to \pi^- D_s^+)$	1.70	
$n_2(B^0 \to \pi^- D_s^+)$	6.98	
$\text{Fract}_{\text{Sig}-\text{CBs}}(B^0 \to \pi^- D_s^+)$	0.50	

Table 16: Fit results for $B_s^0 \to D_s^- \pi^+$ in BDT region 2 with sharper MC shapes.

	Magnet Up	Magnet Down
Parameters	Fit Results	Fit Results
$N_{B_s^0 \to D_s^- \pi^+}$	$12\,354 \pm 145$	$13\,781 \pm 153$
$N_{B^0 \to \pi^- D_s^+}$	232 ± 97	469 ± 102
$N_{B_s^0 \to D_s^- \rho^+}$	$5\,160 \pm 470$	$5\,671 \pm 498$
$N_{B_s^0 \to D_s^{*-} \pi^+}$	$8\,724 \pm 580$	$9\,065 \pm 618$
$N_{B_s^0 \to D_s^{*-} \rho^+}$	73 ± 16	82 ± 18
$N_{B^0 \to \pi^- D_s^{*+}}$	86 ± 16	94 ± 18
$N_{B^0 \to D^- \pi^+}$	648 ± 56	685 ± 62
$N_{B^0 \to D^- \rho^+}$	28 ± 3	30 ± 3
$N_{B_s^0 \to D_s^{\mp} K^{\pm}}$	20 ± 2	21 ± 2
$N_{\Lambda_b^0 \to \Lambda_c^+ \pi^-}$	185 ± 19	202 ± 20
p_0	0.25276 ± 0.02584	0.22368 ± 0.02474
p_1	-0.00446 ± 0.00031	-0.00442 ± 0.00029
$N_{\text{combinatorial}}$	$17\,333 \pm 624$	$19\,488 \pm 663$
Common Parameters		
$\overline{x}$	$5\,371.28 \pm 0.15$	
σ	17.47 ± 0.15	
$\text{mean}_{B^0 \to \pi^- D_s^+}$	$5\,284.2 \pm 0.0$	
σ	17.47 ± 0.15	
Fixed Parameters		
α_1	1.6371	
α_2	-1.9537	
n_1	1.37	
n_2	3.13	
$\text{Fract}_{\text{Sig}-\text{CBs}}$	0.50	
$\alpha_1(B^0 \to \pi^- D_s^+)$	1.6371	
$\alpha_1(B^0 \to \pi^- D_s^+)$	-1.9537	
$n_1(B^0 \to \pi^- D_s^+)$	1.37	
$n_2(B^0 \to \pi^- D_s^+)$	3.13	
$\text{Fract}_{\text{Sig}-\text{CBs}}(B^0 \to \pi^- D_s^+)$	0.50	

Table 17: Fit results for $B^0 \to \pi^- D_s^+$ in BDT region 2 without some of the smaller backgrounds.

	Magnet Up	Magnet Down
Parameters	Fit Results	Fit Results
$N_{B_s^0 \to D_s^- \pi^+}$	$23\,856 \pm 183$	$26\,416 \pm 194$
$N_{B^0 \to \pi^- D_s^+}$	$1\,031 \pm 114$	988 ± 120
$N_{B_s^0 \to D_s^- \rho^+}$	$8\,622 \pm 635$	$9\,046 \pm 679$
$N_{B_s^0 \to D_s^{*-} \pi^+}$	$12\,635 \pm 578$	$15\,128 \pm 614$
$N_{B^0 \to D^- \pi^+}$	$1\,214 \pm 84$	$1\,483 \pm 92$
$N_{\Lambda_b^0 \to \Lambda_c^+ \pi^-}$	193 ± 20	230 ± 21
p_0	0.00000 ± 0.00376	0.00000 ± 0.00295
p_1	-0.00837 ± 0.00029	-0.00876 ± 0.00029
$N_{\text{combinatorial}}$	$4\,102 \pm 647$	$4\,299 \pm 672$
Common Parameters		
$\overline{x}$	$5\,371.24 \pm 0.10$	
σ	17.63 ± 0.09	
$\text{mean}_{B^0 \to \pi^- D_s^+}$	$5\,284.2 \pm 0.0$	
σ	17.63 ± 0.09	
Fixed Parameters		
α_1	1.5966	
α_2	-1.5879	
n_1	1.70	
n_2	6.98	
$\text{Fract}_{\text{Sig}-\text{CBs}}$	0.50	
$\alpha_1(B^0 \to \pi^- D_s^+)$	1.5966	
$\alpha_1(B^0 \to \pi^- D_s^+)$	-1.5879	
$n_1(B^0 \to \pi^- D_s^+)$	1.70	
$n_2(B^0 \to \pi^- D_s^+)$	6.98	
$\text{Fract}_{\text{Sig}-\text{CBs}}(B^0 \to \pi^- D_s^+)$	0.50	

Table 18: Fit results for $B^0 \to \pi^- D_s^+$ in BDT region 2 with scaled tail parameters.

	Magnet Up	Magnet Down
Parameters	Fit Results	Fit Results
$N_{B_s^0 \to D_s^- \pi^+}$	$23\,728 \pm 178$	$26\,254 \pm 189$
$N_{B^0 \to \pi^- D_s^+}$	$1\,053 \pm 101$	999 ± 107
$N_{B_s^0 \to D_s^- \rho^+}$	$8\,318 \pm 525$	$8\,625 \pm 562$
$N_{B_s^0 \to D_s^{*-} \pi^+}$	$13\,597 \pm 531$	$16\,153 \pm 566$
$N_{B_s^0 \to D_s^{*-} \rho^+}$	108 ± 23	115 ± 26
$N_{B^0 \to \pi^- D_s^{*+}}$	150 ± 28	164 ± 31
$N_{B^0 \to D^- \pi^+}$	$1\,168 \pm 84$	$1\,429 \pm 91$
$N_{B^0 \to D^- \rho^+}$	367 ± 36	402 ± 40
$N_{B_s^0 \to D_s^{\mp} K^{\pm}}$	41 ± 4	45 ± 4
$N_{\Lambda_b^0 \to \Lambda_c^+ \pi^-}$	188 ± 19	225 ± 20
p_1	-0.00765 ± 0.00043	-0.00814 ± 0.00040
$N_{\text{combinatorial}}$	$2\,944 \pm 442$	$3\,187 \pm 446$
Common Parameters		
$\overline{x}$	$5\,371.10 \pm 0.10$	
σ	16.57 ± 0.09	
$\text{mean}_{B^0 \to \pi^- D_s^+}$	$5\,284.2 \pm 0.0$	
σ	16.57 ± 0.09	
Fixed Parameters		
α_1	1.1330	
α_2	-1.0539	
n_1	5.51	
n_2	140.00	
$\text{Fract}_{\text{Sig}-\text{CBs}}$	0.50	
$\alpha_1(B^0 \to \pi^- D_s^+)$	1.1200	
$\alpha_1(B^0 \to \pi^- D_s^+)$	-1.0500	
$n_1(B^0 \to \pi^- D_s^+)$	6.42	
$n_2(B^0 \to \pi^- D_s^+)$	139.91	
$\text{Fract}_{\text{Sig}-\text{CBs}}(B^0 \to \pi^- D_s^+)$	0.50	

References

[1] P. A. M. Dirac, *The Quantum theory of electron*, Proc. Roy. Soc. Lond. **A117** (1928) 610.

[2] A. D. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[3] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[4] R. L. Garwin, L. M. Lederman, and M. Weinrich, *Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon*, Phys. Rev. **105** (1957) 1415.

[5] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the k_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.

[6] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[7] M. Kobayashi and T. Maskawa, *Cp-violation in the renormalizable theory of weak interaction*, Progress of Theoretical Physics **49** (1973), no. 2 652, `arXiv:http://ptp.oxfordjournals.org/content/49/2/652.full.pdf+html`.

[8] A. Pich, *The Standard Model of Electroweak Interactions*, `arXiv:1201.0537`.

[9] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[10] J. Charles *et al.*, *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, Phys. Rev. **D91** (2015), no. 7 073007, `arXiv:1501.05013`.

[11] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.

[12] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, *QCD factorization for exclusive, nonleptonic B meson decays: General arguments and the case of heavy light final states*, Nucl. Phys. **B591** (2000) 313, `arXiv:hep-ph/0006124`.

[13] P. Ball and R. Zwicky, *New results on $B \to$ pi, K, eta decay formfactors from light-cone sum rules*, Phys. Rev. **D71** (2005) 014015, `arXiv:hep-ph/0406232`.

[14] I. Caprini, L. Lellouch, and M. Neubert, *Dispersive bounds on the shape of anti-$B \to D^{(*)}$ lepton anti-neutrino form-factors*, Nucl. Phys. **B530** (1998) 153, `arXiv:hep-ph/9712417`.

[15] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots available at http://www.slac.stanford.edu/xorg/hfag/.

[16] L. J. Bel, K. De Bruyn, R. Fleischer, M. Mulder, and N. Tuning, *Anatomy of $B \to D\bar{D}$ Decays*, arXiv:1505.01361.

[17] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008), no. 08 S08001.

[18] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008), no. 08 S08003.

[19] The CMS Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008), no. 08 S08004.

[20] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008), no. 08 S08002.

[21] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[22] LHCb collaboration, *LHCb VELO (public page)*, http://lhcb-vd.web.cern.ch/lhcb-vd/html/project.htm.

[23] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions of $B_s^0 \to D_s^* K^+$ and $B_s^0 \to D^* \pi^+$*, LHCb-PAPER-2015-008, in preparation.

[24] A. Martin Sanchez, P. Robbe, and M.-H. Schune, *Performances of the LHCb L0 Calorimeter Trigger*, Tech. Rep. LHCb-PUB-2011-026. CERN-LHCb-PUB-2011-026, CERN, Geneva, Jun, 2012.

[25] L. Carson, R. Koopman, I. Sepp, and N. Tuning, *Measurement of the p_T and η dependences of Λ_b^0 production, and $\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \pi^-)$, using hadronic decays*, Internal LHCb documentation. .

[26] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286.