



Faculteit Wetenschappen
Vakgroep Fysica en Sterrenkunde
Experimentele Deeltjesfysica

Bepaling van de verhouding van de b -quark
fragmentatiefracties f_s/f_d voor 7 TeV
proton-protonbotsingen en een meting van de
relatieve vervalbreedte van het verval $B^0 \rightarrow D^- K^+$

Pieter David

Begeleider: dr. Vladimir V. Gligorov
Promotor: prof. dr. Dirk Ryckbosch

Proefschrift ingediend met het oog op het behalen van de graad
Master in de fysica en de sterrenkunde

Academiejaar 2010–2011



Acknowledgements

I would like to thank my promotor prof. Ryckbosch and the CERN LHCb group for giving me the opportunity to stay at CERN for several months to work on the analysis described in this thesis. And I want to thank Kim for talking about the possibility of doing a master's thesis in LHCb.

I also want to thank my supervisor Vava, for all his help, and all the people I worked with: the proponents of this analysis, especially Nicola, Barbara and Piotr.

Finally I want to thank my parents and sisters for their support.

Samenvatting

Dit eindwerk beschrijft de meting van de productieverhouding f_s/f_d van B_s^0 en B^0 mesonen in het LHCb experiment aan de LHC versneller, door de studie van de hadronische vervalprocessen $B^0 \rightarrow D^- \pi^+$, $B^0 \rightarrow D^- K^+$ en $B_s^0 \rightarrow D_s^- \pi^+$. De hoge energie en luminositeit van de versneller, samen met de voor een b -fysicaexperiment uitzonderlijke hadronische triggers, voorzien het LHCb experiment van een groot aantal van deze vervalprocessen: enkel met de gegevens verzameld in 2010 zijn metingen mogelijk met precisies die in andere omgevingen veel meer tijd vergen.

De LHCb detector is een eenarmige spectrometer in de voorwaartse richting: enkel botsingsfragmenten die het interactiepunt verlaten in een relatief kleine hoek met de bundel zijn detecteerbaar. De meeste detectoronderdelen zijn na elkaar geplaatst langs de richting van de bundels, vergelijkbaar met een *fixed target* botsingsexperiment. Het eerste hoofdstuk beschrijft de LHCb detector in meer detail.

Het tweede hoofdstuk geeft een kort overzicht van de smaakstructuur van het standaard-model, een inleiding tot metingen van CP schending in het verval van neutrale mesonen en een samenvatting van een aantal concepten die nuttig kunnen zijn bij het beschouwen van de vervalprocessen van zware mesonen. Onder andere de factorisatiehypothese wordt besproken, een eenvoudige benadering die, in de gevallen waar ze verantwoord is, met een behoorlijke precisie het verval van zware hadronen beschrijft. De keuze voor de vervalprocessen $B_{d,s}^0 \rightarrow D_{(s)}^- h^+$ is ingegeven door de theoretische en experimentele resultaten beschreven in [1]. Door gebruik te maken van factorisatie, de $SU(3)$ smaaksymmetrie tussen de up-, down- en strangequark en de grote massa van de b - en c -quark ten opzichte van de typische hadronisatie-energieschaal kunnen de matrixelementen voor deze processen met voldoende precisie aan elkaar gerelateerd worden. Een meting van de verhouding van het aantal waargenomen reacties van elke soort kan dan gebruikt worden om de productieverhouding te bekomen.

De inleidende hoofdstukken worden gevolgd door een gedetailleerde beschrijving van de experimentele methode. Twee strategieën voor de bepaling van f_s/f_d worden gevolgd: bij het gebruik van de verhouding tussen $B^0 \rightarrow D^- \pi^+$ en $B_s^0 \rightarrow D_s^- \pi^+$ moet rekening worden gehouden met bijdragen van W -uitwisselingsprocessen aan het matrixelement — wat voor een grotere theoretische onzekerheid zorgt. Anderzijds is het $B^0 \rightarrow D^- K^+$ proces Cabibbo-onderdrukt ten opzichte van $B^0 \rightarrow D^- \pi^+$, waardoor de statistische onzekerheid in dat geval groter wordt.

Vervolgens wordt de selectie besproken van de waargenomen B vervalkandidaten. Bij iedere stap werden de verschillen in efficiëntie voor de verschillende vervalkanalen bestudeerd, met als resultaat een aantal correctiefactoren en systematische onzekerheden voor de gemeten verhoudingen. De selectie was zo opgezet dat de (benaderende) scheiding van pionen en kaonen — en dus van de verschillende bestudeerde vervalprocessen — met behulp van de informatie verzameld door de Cherenkovdetectoren, een aparte stap vormde, net voor de fits aan de invariante massaspectra.

Ook het modelleren van de achtergronden in deze spectra was een belangrijk onderdeel van het experimentele werk: sommige verwante vervalprocessen zijn immers erg moeilijk te elimineren, bijvoorbeeld diegene waarbij een neutraal deeltje deel uitmaakt van de finale toestand maar niet als zodanig gereconstrueerd wordt. Ook de aanwezigheid van de andere signaalver-

vallen door de beperkte efficiëntie van de pion-kaonscheiding zorgde voor belangrijke achtergronden. De vorm van de belangrijkste dergelijke bijdragen tot de invariante massaspectra werd gemodelleerd, hoofdzakelijk op basis van simulaties, waarna de grootte van elke bijdrage uit fits aan de gemeten distributies bepaald werd. Ook de vorm van het signaal werd naar gesimuleerde data gemodelleerd, waarbij twee correctiefactoren, bepaald in de $B^0 \rightarrow D^- \pi^+$ fit, gebruikt werden om meer parameters vast te houden in de $B_s^0 \rightarrow D_s^- \pi^+$ en $B^0 \rightarrow D^- K^+$ fits, waar geen betrouwbaar resultaat kon bekomen worden indien deze parameters door de fit te bepalen waren.

De onnauwkeurigheden geïntroduceerd door de gebruikte fitprocedure werden in detail bestudeerd door het uitvoeren van grote aantallen pseudo-experimenten. Daarbij werden distributies gegenereerd met licht afwijkende modellen en steeds op dezelfde wijze gefit als bij de “echte” data. Op die manier werd de gevoeligheid aan elk afzonderlijk effect bepaald, evenals de belangrijkste correlaties.

Ten slotte werden de gemeten waarden gebruikt om de f_s/f_d verhouding en de relatieve vervalbreedte van $B^0 \rightarrow D^- K^+$ te bepalen.

Chapter 1

The LHCb experiment

The Large Hadron Collider (LHC) accelerator collides two proton beams, circulating in opposite directions, in four interaction points at a centre of mass energy in the TeV range — the design energy is 14 TeV. For 2010, 2011 and 2012 data taking, the centre of mass energy is kept at 7 TeV. The bunch-crossing frequency goes up to 40 MHz. General-purpose detectors are situated at two of the interaction points. With these detectors, the ATLAS and Compact Muon Solenoid (CMS) collaborations mainly aim to observe the Higgs boson, predicted by the model of electroweak symmetry breaking, and search for signals of physics beyond the Standard Model. The collisions at the two other interaction points are studied by dedicated detectors. The ALICE detector is designed for the study of heavy-ion physics and the LHCb experiment is optimised for precision measurements of heavy quark (charm and beauty) decays.

The advantage of the environment sketched above for a heavy flavour physics experiment, with respect to b -factory experiments BELLE and BABAR, is the very large number of produced $b\bar{b}$ pairs, due to the large cross-section and luminosity — even if the latter is significantly lower than for the other LHC experiments. The expected number of $b\bar{b}$ pairs produced in 10^7 s of running at nominal conditions is expected to be 10^{12} . First, a short overview of the LHCb detector is given (a more complete description can be found in [2], the design reports, *e.g.* [3], and the technical reports published during the construction and commissioning of the detector).

1.1 The LHCb detector

The dominant $b\bar{b}$ production mechanism in pp collisions at the TeV energy scale is gluon fusion. From simulation it was expected that the gluons have a very different momentum, such that the $b\bar{b}$ pair is produced with a large total momentum in the laboratory frame. The b quarks produced in such a process make a trajectory close to the beam direction, both on the same side of the interaction point. This explains the choice for the single-arm spectrometer design of the LHCb detector: most b hadrons are observed in a small angular distance from the beam direction and in most cases both quarks of the $b\bar{b}$ pair can be observed [4]. The forward angular coverage goes from approximately 10 mrad to 300 mrad (250 mrad) in the horizontal, bending (vertical, non-bending) plane.

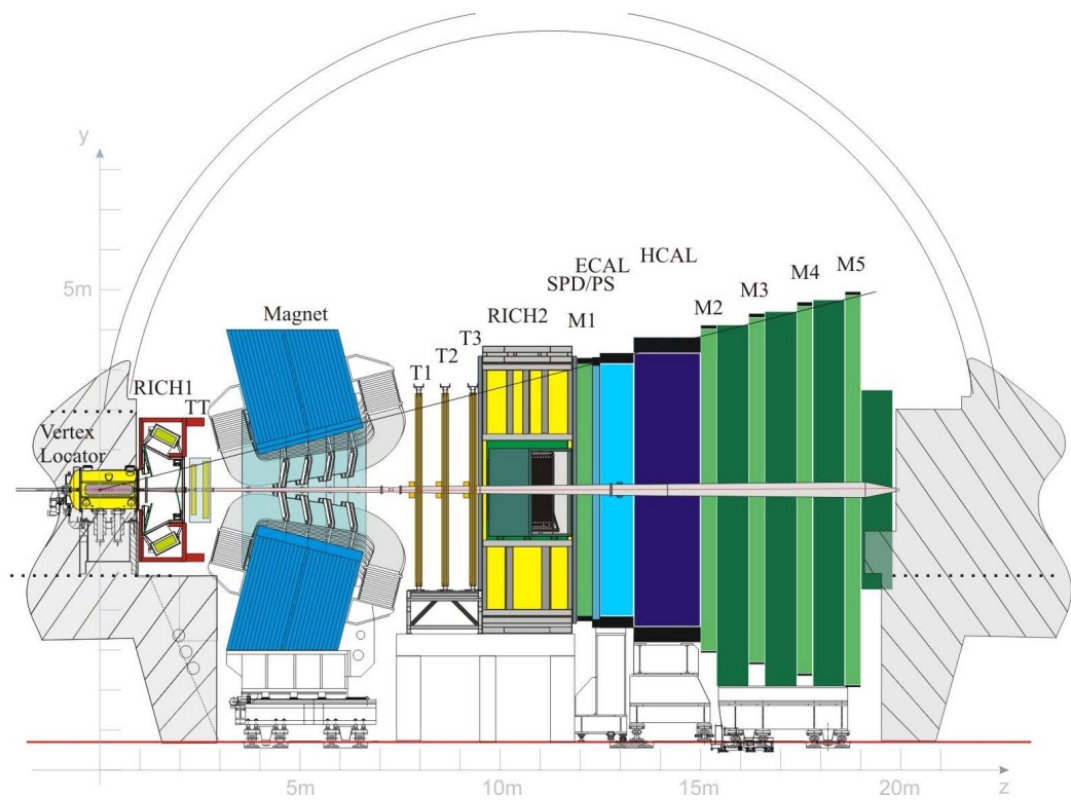


Figure 1.1: Vertical cross-section of the LHCb detector

The detector part closest to the interaction region is the Vertex Locator (VELO). This sub-detector consists of a series of silicon modules along the beam direction. Each of these modules provides a measure of the cylindrical r and φ coordinates. The distance to the interaction region during data taking is smaller than the space required by the LHC beam during injection. Therefore a mechanical system can move the two halves of the detector closer to the collision point when the beams are stable. The VELO provides very precise tracking information close to the interaction point, which is required to reconstruct the primary and secondary vertices. Vertex reconstruction is crucial for the study of the decay of the weakly decaying charm and beauty mesons, and for the measurement of their lifetimes. Figure 1.2 shows the resolution of the x component of the impact parameter (IP)¹.

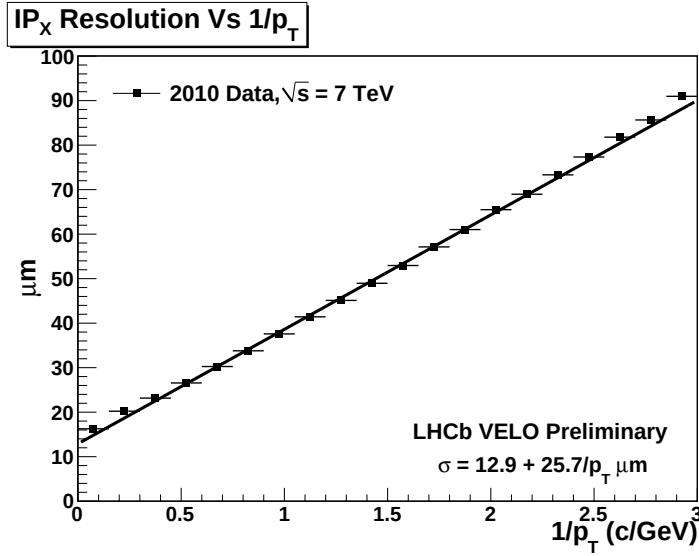


Figure 1.2: Resolution of the x component of the IP as a function of the track transverse momentum

The next detector crossed by collision products is the first ring imaging Cherenkov detector (RICH) detector. Cherenkov detectors play a crucial role in particle identification (PID): the opening angle of the light cone a particle produces in such a detector is a measure for its velocity. Complemented with the momentum measurement, this allows for an estimate of the particle mass. The information provided by the RICH detectors renders it possible to *e.g.* separate the light hadrons K and π — the reconstruction calculates a $DLL_{K\pi}$ value for the difference in log-likelihood between the kaon and pion hypothesis. The central role of PID information in the separation of different decay modes is discussed in more detail in section 3.5. The different gas mixtures in the RICH detectors make them complementary: RICH1 provides PID information in a momentum range from 1 to 60 GeV, while RICH2 gives PID information for momenta from 15 up to over 100 GeV.

¹The IP of a track with respect to a vertex is the shortest vector between the (extrapolated) track and the vertex, *i.e.* the vector between the track and the vertex in the plane perpendicular to the track that contains the vertex.

The first part of the tracking system is the tracker Turicensis (TT). The design of this sub-detector is very similar to that of the inner tracker (IT) — they are together called the silicon tracker (ST). Both use silicon microstrip sensors, with a strip pitch of about 200 μm . The TT covers the whole experiment acceptance, while the IT covers a region in the centre of the three tracking stations downstream of the magnet. The outer part of these tracking stations, where a coarser granularity can be used, forms the outer tracker (OT), a drift-time detector.

A warm dipole magnet, designed with an integrated magnetic field of 4 Tm for tracks of 10 m length is situated between the TT and the tracking stations T1, T2 and T3.

Further along the z -axis, first the second RICH detector is found, and next the calorimeters and the muon system.

The calorimeter system identifies electrons, photons and hadrons, and measures their energies and positions. It consists in turn of a number of subsystems: the scintillator pad detector (SPD), preshower detector (PS), electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). The first three are segmented in three different sections, while the HCAL is segmented in two zones. The sampling-type calorimeters ECAL and HCAL provide energy estimates, while the SPD and PS have the identification of different particle types as their main purpose. Both are scintillator detectors, with a lead layer in between. The SPD detects particles before the showering starts and thus only detects charged particles, while the PS provides a means way to separate electromagnetic and hadronic showers. For the identification of electrons, for example, the PS helps reject the charged pion background, while the SPD identifies the photon and π^0 background. The calorimeters also play an important role in the L0 trigger — providing information for the hadronic triggers is in fact one of the main purposes of the HCAL.

The muon system is composed of five rectangular stations (M1-M5), placed along the beam axis — M2 to M5 downstream of the calorimeters, M1 in front of the calorimeters to improve the trigger p_T measurement. Each station has a granularity in four regions, and the transverse dimensions scale with the distance to the interaction point. The detector technologies used are multi-wire proportional chambers for all but the inner region, where triple-gas electron multiplier (GEM) detectors are used.

The trigger system has the task to decide for which bunch crossings to read out the full detector and to select out of these the events that need to be stored for analysis, going from a 40 MHz bunch-crossing rate to a 2 kHz output event rate with a full detector readout rate of 1 MHz. This reduction is a two-step process: the hardware-based L0 trigger performs the first reduction, based on an the E_T of the highest- E_T 2×2 -cluster found in the ECAL and in the HCAL, and the highest p_T candidate found in the muon stations.

The second trigger layer, the High-Level trigger (HLT) is software-based, and consists in turn of two steps, HLT1 and HLT2. The HLT1 triggers were changed during the 2010 data taking period from the “confirmation-style” HLT1 to the 1Track HLT1 [refs]. The former tried to confirm the L0 decision by gradually adding more information [5, 6, 7, 8], the latter uses VELO information to select a few tracks to be extrapolated to the tracking stations and performs a fast fit, step by step reducing the number of candidate-tracks, finally applying selection criteria on one single “good” track. This strategy was found to be more robust for triggering hadronic B decays in case ghost tracks are present, especially when there are on average several inter-

actions per bunch crossing (pileup) [9].

HLT2 performs a selection similar in nature to an offline analysis, but using faster (and slightly lower-resolution) reconstruction algorithms. More information about the HLT2 trigger line selecting the events used in this analysis is given in the event selection section.

Chapter 2

Flavour

“Ordinary” nuclear matter, protons and neutrons, is experimentally found to be mostly composed of the light up and down quarks. In high-energy interactions, four more quarks were found. The first to be discovered was the strange quark — called strange because of the then unfamiliar pattern its weak decays showed in the detectors, with long tracks for the unstable strange particles. This was later understood in terms of the weak interaction, the decay mechanism for the lightest strange particles, that is much weaker than the strong interactions and gives the lightest strange particles a much longer lifetime than light, non-strange resonances of similar mass. The approximate SU(3) flavour symmetry between the three quarks now called “light”, that explains the symmetry of the light meson spectrum, is at the origin of the first quark model [10, 11, 12].

B^0 and B_s^0 decays show a certain qualitative similarity with the early flavour physics observation above: the most prominent interaction present in the hadrons is clearly the strong interaction — but there are no lighter bound states with the same quark content and the strong interaction cannot change the flavour. Therefore, the decay proceeds through the weak interaction and the lifetimes are relatively long. An important challenge in the study of these decays is to understand the role of the strong interaction in the decay processes.

The following sections give a short introduction to flavour physics: first the breaking of the flavour symmetry in the standard model is summarised. Next CP violation is discussed, one of the key questions in flavour physics. Finally some elements relevant for the study of heavy meson decays are summarised.

2.1 Electroweak symmetry breaking in the Standard Model

The Standard Model, that is so far very successful in describing the fundamental particles and interactions at high energies, can be formalised as a gauge theory, with the gauge symmetry group $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$. The lepton and quark fields can then be arranged as follows in representations of the gauge group: for each generation there is an SU(2) doublet of left-handed leptons and a SU(2) doublet of left-handed quarks. The right-handed fields are SU(2) singlets. The quark fields transform according to the fundamental representation of the colour gauge group, while the lepton fields are colour singlets. The $SU(2)_L \otimes U(1)_Y$ symmetry is broken

by adding a scalar field, whose vacuum expectation value spontaneously breaks the symmetry present in its potential. The gauge symmetry breaking mechanism was first formulated by Englert and Brout [13] and Higgs [14]. The precise form of the symmetry-breaking vacuum yields a relation between the electrical charge, the hypercharge and the weak isospin, which allows to assign the hypercharges that reproduce the observed particle spectrum.

Interestingly the most general Lagrangian that can be written, given the gauge symmetry group and the symmetry breaking mechanism, also contains terms generating quark masses and thus breaking the flavour symmetry.

The $SU(2)_L \otimes U(1)_Y$ gauge Lagrangian before symmetry breaking is given by

$$\mathcal{L} = \sum_f \bar{\psi}_f i \gamma^\mu \left(\partial_\mu - i g A_\mu^a \tau_a - \frac{i}{2} g' B_\mu Y_f \right) \psi_f - \frac{1}{4} F_{A,\mu\nu}^a F_A^{a,\mu\nu} - \frac{1}{4} F_{B,\mu\nu} F_B^{\mu\nu}, \quad (2.1)$$

with ψ_f the fermion fields, A_μ^a and B_μ the $SU(2)$ and $U(1)_Y$ gauge fields, respectively, with coupling constants g and g' and generators τ_a and Y . $F_{A,\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f_{abc} A_\mu^b A_\nu^c$ and $F_{B,\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ are the gauge field tensors, with f_{abc} the $SU(2)$ structure constants and implicit summation over the gauge representation indices a, b and c and the Lorentz indices μ and ν . An infinitesimal gauge transformation can be written as [15]

$$\varphi \rightarrow \left(1 + i(\Delta\alpha^a)\tau_a + i\frac{\Delta\beta}{2}Y \right) \varphi, \quad (2.2)$$

and the covariant derivative is given by

$$D_\mu = \partial_\mu - i g A_\mu^a \tau_a - \frac{i}{2} g' B_\mu Y. \quad (2.3)$$

A gauge invariant kinetic term and potential terms for a $SU(2)$ -doublet field φ are then added to the Lagrangian,

$$\mathcal{L}_H = (D_\mu \varphi)^* (D^\mu \varphi) - \mu^2 \varphi^* \varphi - \lambda (\varphi^* \varphi)^2. \quad (2.4)$$

In case $\mu^2 < 0$, all φ_0 satisfying

$$|\varphi_0| = \frac{1}{\sqrt{2}} \sqrt{\frac{-\mu^2}{\lambda}} = \frac{v}{\sqrt{2}},$$

are symmetry breaking minima $\varphi_0 \neq 0$. Choosing the vacuum

$$\varphi_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \quad \text{with} \quad Y(\varphi_0) = 1$$

breaks all generators, except for the combinations of (2.2) where $\Delta\alpha^3 = \Delta\beta$, or $Q = \tau_3 + Y/2$. The expansion of the scalar field φ around its expectation value φ_0 generates the gauge boson mass terms from the kinetic term. The difference $\varphi - \varphi_0$ describes the dynamics of the postulated massive scalar boson.

Fermion masses can be generated without explicitly breaking the gauge symmetry, by introducing Yukawa couplings of the fermions to the scalar field. All allowed terms involving quark fields are

$$\mathcal{L}_{m,\text{quark}} = -\lambda_d^{ij} \bar{Q}_L^i \phi d_R^j - \lambda_u^{ij} \epsilon^{ab} \bar{Q}_{L,a}^i \phi_b^\dagger u_R^j + \text{hermitian conjugate terms}, \quad (2.5)$$

with ϵ^{ab} fully antisymmetric in its indices. CP conjugation of the above interchanges the couplings λ_d^{ij} and λ_u^{ij} with their complex conjugate — if these were real matrices, no CP violation would be present.

The physical degrees of freedom in these coupling matrices can be found by studying the behaviour of the Yukawa terms under a redefinition of the quark fields. The unitary matrices that diagonalise the hermitian matrices $\lambda_u \lambda_u^\dagger$ and $\lambda_d \lambda_d^\dagger$ as

$$\lambda_u \lambda_u^\dagger = U_u D_u^2 U_u^\dagger \quad \text{and} \quad \lambda_u^\dagger \lambda_u = W_u D_u^2 W_u^\dagger, \quad (2.6)$$

with D_u diagonal matrices containing only real and positive elements, give

$$\lambda_u = U_u D_u W_u^\dagger, \quad (2.7)$$

up to a global phase. Analogously λ_d gives D_d , U_d and W_d . Now performing the transformation on the fields

$$u_R^j \rightarrow W_u^{jk} u_R^k, \quad u_L^i \rightarrow U_u^{il} u_L^l, \quad d_R^j \rightarrow W_d^{jk} d_R^k \quad \text{and} \quad d_L^i \rightarrow U_d^{il} d_L^l \quad (2.8)$$

diagonalises the Yukawa couplings in (2.5) to

$$\mathcal{L}_{m,\text{quark}} = -\frac{v D_d^i}{\sqrt{2}} \bar{d}_L^i d_R^i - \frac{v D_u^i}{\sqrt{2}} \bar{u}_L^i u_R^i + \text{hermitian conjugate terms}. \quad (2.9)$$

Not all other terms of the Lagrangian (2.1) are invariant under the transformations (2.8), that mix the different components of the $SU(2)$ -doublets. The charged current interaction terms

$$\frac{g}{\sqrt{2}} W_\mu^+ \bar{u}_L \gamma^\mu d_L + \frac{g}{\sqrt{2}} W_\mu^- \bar{d}_L \gamma^\mu u_L \quad (2.10)$$

become

$$\frac{g}{\sqrt{2}} W_\mu^+ \bar{u}_L^j \gamma^\mu U_u^{*ij} U_d^{ik} d_L^k + \frac{g}{\sqrt{2}} W_\mu^- \bar{d}_L^j \gamma^\mu U_d^{*ij} U_u^{ik} u_L^k, \quad (2.11)$$

the mass eigenstates are mixed by the weak interaction, and all mixing (and CP violation) information can be embedded in the Cabibbo-Kobayashi-Maskawa matrix $V_{CKM}^{ij} = U_u^{*ki} U_d^{kj}$ — named after Cabibbo who realised that the d and s quark mix to form the partner state of the u [16] and Kobayashi and Maskawa who realised that a 3×3 matrix (for three quark generations) can generate CP violation [17]. The physical number of (real) degrees of freedom is four in that case: unitarity reduces the eighteen of an arbitrary complex matrix to nine and the relative phases

between the quark fields can be used to eliminate five more. The standard parametrisation is given by

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (2.12)$$

with $s_{ij} = \sin \theta_{ij}$ and $c_{ij} = \cos \theta_{ij}$. Another frequently used parametrisation, inspired by the hierarchy $s_{13} \ll s_{23} \ll s_{12} \ll 1$, the so-called Wolfenstein parametrisation [18], is defined by

$$s_{12} = \lambda \quad s_{23} = A\lambda^2 \quad s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta) \quad (2.13)$$

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \begin{pmatrix} -\frac{\lambda^4}{8} & 0 & 0 \\ \frac{A^2\lambda^5}{2}(1 - 2(\rho + i\eta)) & -\frac{\lambda^4}{8}(1 + 4A^2) & 0 \\ \frac{A\lambda^5}{2}(\rho + i\eta) & \frac{A\lambda^4}{2}(1 - 2(\rho + i\eta)) & -\frac{A^2\lambda^4}{2} \end{pmatrix} + \mathcal{O}(\lambda^6). \quad (2.14)$$

In both parametrisations, there is only one parameter that can generate imaginary parts for CKM matrix elements (δ and η) — CP violation is present if this parameter is nonzero or equivalently if the area of the unitarity triangle is nonzero.

The unitarity of the CKM matrix also requires that

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0, \quad (2.15)$$

which can be written as

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} = 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}, \quad (2.16)$$

where the first equality is the definition of $\bar{\rho} + i\bar{\eta}$, and the second the standard model prediction: the consistency of all measurements with one value of $\bar{\rho} + i\bar{\eta}$ is the ultimate test of the CKM mechanism — the fit is visualised using allowed and excluded regions for the point representing $\bar{\rho} + i\bar{\eta}$ in the complex plane, and the relation (2.16) is depicted as a triangle. The angles of the unitarity triangle

$$\alpha = \text{Arg} \left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right) \quad \beta = \text{Arg} \left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right) \quad \gamma = \text{Arg} \left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right) \quad (2.17)$$

supplemented with β_s

$$\beta_s = \text{Arg} \left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*} \right). \quad (2.18)$$

are commonly used to express experimental results in a convention-independent way. Using the approximation up to λ^4 in (2.14) — where only V_{ub} , V_{td} and V_{ts} have a phase — the CKM matrix can be written using these phases and the absolute values of the matrix elements:

$$V_{CKM} = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}|e^{-i\gamma} \\ -|V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}|e^{-i\beta} & -|V_{ts}|e^{i\beta_s} & |V_{tb}| \end{pmatrix} + O(\lambda^5) \quad (2.19)$$

An important observation from the previous discussion is the fact that the flavour symmetry breaking only enters the strong interaction via the quark mass terms.

2.2 Neutral meson mixing and CP violation

Neutral meson systems play a central role in the experimental study of CP violation. The combination of mixing between the charge conjugated flavour eigenstates and the CP violating phases allows to determine the angles of the unitarity triangle. One of the goals of the study of CP violation is to provide more insight in the baryon asymmetry in the early universe, the excess of matter over antimatter that after annihilation remains as the universe we see today. This asymmetry could be explained by the presence baryon number violation, violation of the C and CP symmetry and interactions in the absence of thermal equilibrium [19]. The level of CP violation observed in the standard model so far, however, is in the current models not sufficient to generate the observed baryon asymmetry.

An example of a CP violation measurement is found in the $B_s^0 \rightarrow D_s^- K^+$ decay which will be used for a γ determination in the LHCb experiment. The γ value extracted from processes like this that are dominated by tree-level diagrams, is expected not to be affected by new physics contributions — the consistency with a measurement using loop diagram dominated decays should provide information about the flavour structure of new physics.

No conservation law forbids the mixing of the B^0 and $\bar{B}^0$ mesons. The most general Hamiltonian matrix in the basis $\{|B^0\rangle, |\bar{B}^0\rangle\}$ can be written as

$$H = M - \frac{i}{2}\Gamma, \quad (2.20)$$

with M and Γ hermitian 2×2 matrices. CPT invariance requires the B^0 and $\bar{B}^0$ particles to have equal masses and lifetimes, which translates into

$$\langle B^0 | \hat{H} | B^0 \rangle = \langle \bar{B}^0 | \hat{H} | \bar{B}^0 \rangle, \quad (2.21)$$

or equivalently $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$. Diagonalising this matrix, and taking advantage of the freedom to choose the phases when defining the CP eigenstates, these can be parametrised as

$$\begin{pmatrix} |B_L^0\rangle \\ |B_H^0\rangle \end{pmatrix} = \begin{pmatrix} p & q \\ p & -q \end{pmatrix} \begin{pmatrix} |B^0\rangle \\ |\bar{B}^0\rangle \end{pmatrix} \quad \text{with} \quad |p|^2 + |q|^2 = 1. \quad (2.22)$$

With the notation $\lambda_{H,L} = (m \pm \frac{\Delta m}{2}) - \frac{i}{2} (\Gamma \pm \frac{\Delta \Gamma}{2})$ for the eigenvalues of the Hamiltonian matrix in the symmetry basis, the time evolution in the flavour basis is easily found, *e.g.* by using the evolution operator $\hat{U}(t) = \exp[-i\hat{H}t]$ ¹

$$\begin{pmatrix} \langle B^0 | \hat{U}(t) B^0 \rangle & \langle B^0 | \hat{U}(t) \bar{B}^0 \rangle \\ \langle \bar{B}^0 | \hat{U}(t) B^0 \rangle & \langle \bar{B}^0 | \hat{U}(t) \bar{B}^0 \rangle \end{pmatrix} = \begin{pmatrix} p & p \\ q & -q \end{pmatrix} e^{-im t} e^{-\frac{\Gamma t}{2}} \begin{pmatrix} e^{i\frac{\Delta m t}{2}} e^{\frac{\Delta \Gamma t}{4}} & 0 \\ 0 & e^{-i\frac{\Delta m t}{2}} e^{-\frac{\Delta \Gamma t}{4}} \end{pmatrix} \begin{pmatrix} \frac{1}{2p} & \frac{1}{2q} \\ \frac{1}{2p} & -\frac{1}{2q} \end{pmatrix} \quad (2.23)$$

$$g_+(t) = \langle B^0 | \hat{U}(t) B^0 \rangle = \langle \bar{B}^0 | \hat{U}(t) \bar{B}^0 \rangle = \frac{1}{2} \exp \left[-im t - \frac{\Gamma t}{2} \right] \cdot \left(\exp \left[\left(i\Delta m + \frac{\Delta \Gamma}{2} \right) \frac{t}{2} \right] + \exp \left[- \left(i\Delta m + \frac{\Delta \Gamma}{2} \right) \frac{t}{2} \right] \right) \quad (2.24)$$

$$g_-(t) = \frac{q}{p} \langle B^0 | \hat{U}(t) \bar{B}^0 \rangle = \frac{p}{q} \langle \bar{B}^0 | \hat{U}(t) B^0 \rangle = \frac{1}{2} \exp \left[-im t - \frac{\Gamma t}{2} \right] \cdot \left(\exp \left[\left(i\Delta m + \frac{\Delta \Gamma}{2} \right) \frac{t}{2} \right] - \exp \left[- \left(i\Delta m + \frac{\Delta \Gamma}{2} \right) \frac{t}{2} \right] \right), \quad (2.25)$$

so the probability for a B^0 ($\bar{B}^0$) meson to have oscillated into its antiparticle after a time t , or, alternatively, to be in the same state, is given by

$$|\langle \bar{B}^0 | \hat{U}(t) B^0 \rangle|^2 = \left| \frac{q}{p} \right|^2 |g_-(t)|^2, \quad (2.26)$$

$$|\langle B^0 | \hat{U}(t) \bar{B}^0 \rangle|^2 = \left| \frac{p}{q} \right|^2 |g_-(t)|^2 \quad \text{with} \quad (2.27)$$

$$|g_-(t)|^2 = \frac{e^{-\Gamma t}}{2} \left(\cosh \frac{\Delta \Gamma t}{2} - \cos \Delta m t \right) \quad (2.28)$$

and

$$|\langle B^0 | \hat{U}(t) B^0 \rangle|^2 = |\langle \bar{B}^0 | \hat{U}(t) \bar{B}^0 \rangle|^2 = |g_+(t)|^2 = \frac{e^{-\Gamma t}}{2} \left(\cosh \frac{\Delta \Gamma t}{2} + \cos \Delta m t \right). \quad (2.29)$$

CP violation can show up in time-dependent decay rates in different ways. A first possibility is CP violation in mixing. An example where only this type is visible, is a flavour-specific decay, where the charge of the decay products determines the state (B^0 or $\bar{B}^0$) of the particle when decaying. Time-dependent analyses of these decays allow to extract Γ , $\Delta \Gamma$, Δm and $|p/q|$ using (2.26) to (2.29); CP violation is present if $|p/q| \neq 1$.

¹Natural units, *i.e.* $\hbar = c = 1$, are used

CP violation in the decay is observed when the norm of the amplitude of a B^0 decay is different from its charge conjugated decay. This can be realised by the interference of two diagrams, without requiring either of them to break CP violation separately. A third possible source of CP violation is the interference between mixing and decay — in case the amplitudes for a B^0 decay and its charge conjugated process have a different argument and mixing is present, different decay rates may be observed. If we consider the case where both a final state f and its CP conjugate $\bar{f}$ are accessible from both a B^0 and a $\bar{B}^0$ decay, the probability for B^0 and $\bar{B}^0$ to decay to each of these states after a time t is given by

$$\begin{aligned}
P_{B^0 \rightarrow f}(t) &= |\mathcal{M}_{B^0 \rightarrow f} \langle B^0 | \hat{U}(t) B^0 \rangle + \mathcal{M}_{\bar{B}^0 \rightarrow f} \langle \bar{B}^0 | \hat{U}(t) B^0 \rangle|^2 \\
&= |\mathcal{M}_{B^0 \rightarrow f}|^2 \left| g_+(t) + \left(\frac{\mathcal{M}_{\bar{B}^0 \rightarrow f}}{\mathcal{M}_{B^0 \rightarrow f}} \cdot \frac{q}{p} \right) \cdot g_-(t) \right|^2 \\
P_{B^0 \rightarrow \bar{f}}(t) &= |\mathcal{M}_{B^0 \rightarrow \bar{f}}|^2 \left| g_+(t) + \left(\frac{\mathcal{M}_{\bar{B}^0 \rightarrow \bar{f}}}{\mathcal{M}_{B^0 \rightarrow \bar{f}}} \cdot \frac{q}{p} \right) \cdot g_-(t) \right|^2 \\
P_{\bar{B}^0 \rightarrow f}(t) &= |\mathcal{M}_{\bar{B}^0 \rightarrow f}|^2 \left| g_+(t) + \left(\frac{\mathcal{M}_{B^0 \rightarrow f}}{\mathcal{M}_{\bar{B}^0 \rightarrow f}} \cdot \frac{p}{q} \right) \cdot g_-(t) \right|^2 \\
P_{\bar{B}^0 \rightarrow \bar{f}}(t) &= |\mathcal{M}_{\bar{B}^0 \rightarrow \bar{f}}|^2 \left| g_+(t) + \left(\frac{\mathcal{M}_{B^0 \rightarrow \bar{f}}}{\mathcal{M}_{\bar{B}^0 \rightarrow \bar{f}}} \cdot \frac{p}{q} \right) \cdot g_-(t) \right|^2.
\end{aligned}$$

Even if $|p/q| = 1$ and $|\mathcal{M}_{B^0 \rightarrow f}| = |\mathcal{M}_{\bar{B}^0 \rightarrow \bar{f}}|$ and $|\mathcal{M}_{\bar{B}^0 \rightarrow f}| = |\mathcal{M}_{B^0 \rightarrow \bar{f}}|$, CP violation can be generated if there is a phase difference between $\frac{\mathcal{M}_{\bar{B}^0 \rightarrow f} q}{\mathcal{M}_{B^0 \rightarrow f} p}$ and $\frac{\mathcal{M}_{B^0 \rightarrow \bar{f}} p}{\mathcal{M}_{\bar{B}^0 \rightarrow \bar{f}} q}$ — the same condition is found using either $P_{B^0 \rightarrow f}(t) \neq P_{\bar{B}^0 \rightarrow \bar{f}}(t)$ or $P_{B^0 \rightarrow \bar{f}}(t) \neq P_{\bar{B}^0 \rightarrow f}(t)$.

2.2.1 $B_s^0 \rightarrow D_s^- K^+$

Using the decay rate expressions and assuming no CP violation in mixing ($|q/p| = 1$), the time-dependent asymmetry between *e.g.* $B_s^0 \rightarrow D_s^- K^+$ and $\bar{B}_s^0 \rightarrow D_s^- K^+$ can be found

$$\frac{\Gamma(B_q^0 \rightarrow f_q)(t) - \Gamma(\bar{B}_q^0 \rightarrow f_q)(t)}{\Gamma(B_q^0 \rightarrow f_q)(t) + \Gamma(\bar{B}_q^0 \rightarrow f_q)(t)} = \frac{\left(1 - |\xi_q|^2\right) \cos(\Delta m_q t) + 2 \operatorname{Im}(\xi_q) \sin(\Delta m_q t)}{\left(1 + |\xi_q|^2\right) \cosh(\Delta \Gamma_q t/2) - 2 \operatorname{Re}(\xi_q) \sinh(\Delta \Gamma_q t/2)} \quad (2.30)$$

where

$$\xi_q = -\frac{\mathcal{M}_{\bar{B}_q^0 \rightarrow f}}{\mathcal{M}_{B_q^0 \rightarrow f}} \left(\frac{q}{p} \right)_{B_q^0}. \quad (2.31)$$

The tree level diagrams straightforwardly give the CKM matrix elements for the different $B_s^0 \rightarrow D_s^- K^+$ decays:

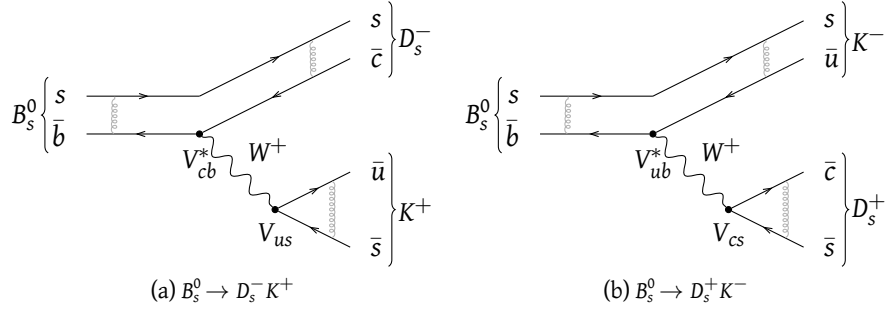


Figure 2.1: Tree level diagrams for the decays $B_s^0 \rightarrow D_s^- K^+$ and $B_s^0 \rightarrow D_s^+ K^-$

$$\mathcal{M}_{B_s^0 \rightarrow D_s^- K^+} \sim V_{cb}^* V_{us} \quad \mathcal{M}_{B_s^0 \rightarrow D_s^+ K^-} \sim V_{ub}^* V_{cs} \quad (2.32)$$

$$\mathcal{M}_{\bar{B}_s^0 \rightarrow D_s^- K^+} \sim V_{cs}^* V_{ub} \quad \mathcal{M}_{\bar{B}_s^0 \rightarrow D_s^+ K^-} \sim V_{us}^* V_{cb} \quad (2.33)$$

Such that for $f = D_s^- K^+$

$$\xi_s = -e^{-i\varphi_s} \left| \frac{V_{ub} V_{cs}^*}{V_{cb}^* V_{us}} \right| e^{-i\gamma} \left| \frac{A_1}{A_2} \right| e^{i\beta_{\text{strong}}} \quad (2.34)$$

$$\bar{\xi}_s = -e^{-i\varphi_s} \left| \frac{V_{cb} V_{us}^*}{V_{ub}^* V_{cs}} \right| e^{-i\gamma} \left| \frac{A_2}{A_1} \right| e^{-i\beta_{\text{strong}}}, \quad (2.35)$$

where φ_s is the weak phase of $\left(\frac{q}{p}\right)_{B_q^0}$ and β_{strong} is the phase difference between the strong interaction matrix elements. In the product $\xi_s \cdot \bar{\xi}_s$ all hadronic factors cancel, which allows for a clean extraction of $\gamma + \varphi_s$. The angle φ_s can be measured independently.

More complicated strategies involving the charge-conjugate final state, the untagged rates and the decays $B^0 \rightarrow D^- \pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$, make a more precise measurement possible [20, 21, 22].

2.3 Description of hadronic B decays

The theoretical description of the decay of heavy quarks is complicated by the structure of the hadrons where these decays can be studied: it has so far not been possible to fully calculate the properties of hadrons from first principles in the framework of quantum chromodynamics (QCD), the colour SU(3) gauge theory that provides a good description of the high-energy strong processes. Due to the number of quarks present and the gauge group, the evolution of the coupling constant shows an increasing value when going to low energy scales. This makes it impossible to truncate the perturbative series when calculating matrix elements — which is an important factor of the success of quantum electrodynamics (QED) for the description of electromagnetic interactions. The presence of well-separated energy scales in the problem, $m_{W^\pm} \approx 80 \text{ GeV}$ for the charged-current weak interaction, the b quark mass of around 5 GeV and the typical energy scale of non-perturbative QCD Λ_{QCD} of several hundreds MeV, makes a number of physically motivated approximations possible.

2.3.1 Effective field theory

A first simplification can be performed by constructing an effective field theory, integrating out contributions from short-distance physics, at momentum scales higher than the normalisation point μ of the effective theory which is typically chosen at a few GeV, well above Λ_{QCD} and small compared to the heavy quark mass. The procedure is inspired by the Wilson picture of renormalisation: by integrating out the degrees of freedom of the full theory (in the typical example of renormalisation where the existence of an exact theory is assumed, the exact theory with scale-independent coupling constants), a theory is obtained that provides a complete description of all processes involving momenta below the renormalisation point, with scale-dependent coupling constants. The processes involving virtual momenta above the renormalisation point generate a set of local operators. A well-known example of this is provided by the Fermi theory of the weak interaction: the W exchange diagram in the neutron decay can be replaced by an effective four-fermion vertex that comes with an effective coupling constant. Performing the expansion of the W propagator, a series of corrections can be obtained, which are suppressed by increasing powers of k^2/m_W^2 , with k the typical energy scale of the decay. In heavy quark theory, the same happens with processes involving virtual gluons that carry large momenta. The *operator product expansion* obtained for decay matrix elements is formally given by

$$A(B \rightarrow f) = \frac{G_F}{\sqrt{2}} \sum_i \lambda_i C_i(\mu) \langle f | \mathcal{O}_i | B \rangle(\mu), \quad (2.36)$$

with G_F the Fermi constant, λ_i the CKM matrix elements, C_i the Wilson coefficients (running coupling constants) and $\mathcal{O}_i$ local operators. Higher-dimensional operators, the remnant of processes involving many high-momentum virtual particles, come with negative coupling constants — this is not problematic because there is a cut-off scale — which naturally generate suppression factors $\sim \mu/M$ when integrating up to μ , with M representative for the energy scale of the virtualities. [23, 24, 25].

The heavy quark fields play a special role in the effective theory. Despite the fact that the normalisation point is chosen well below the heavy quark mass, these cannot fully be integrated out because they will be present in initial and final states. The virtualities containing heavy quarks can however be integrated out, removing the process of heavy quark pair creation from the theory. This can be done directly [25] using a Gaussian integration of the generating function

$$Z_{\text{heavy}}(\eta, \bar{\eta}) = \int \mathcal{D}Q \mathcal{D}\bar{Q} \exp i \left[\int d^4x [\bar{Q} (i\not{D} - m_Q) Q] + \bar{\eta}Q + \bar{Q}\eta \right] \quad (2.37)$$

Using the four-velocity of the heavy hadron, $v = \frac{p_{\text{hadron}}}{m_{\text{hadron}}}$, the heavy quark field can be split into its projections on the $\not{v}$ eigenbase:

$$\varphi_v = \frac{1 + \not{v}}{2} Q \quad \not{v} \varphi_v = \varphi_v \quad (2.38)$$

$$\chi_v = \frac{1 - \not{v}}{2} Q \quad \not{v} \chi_v = -\chi_v, \quad (2.39)$$

This corresponds to the separation of the two degrees of freedom describing a spin 1/2 particle from the two degrees of freedom describing its antiparticle in the not fully relativistic quantum mechanical application of the Dirac equation, where virtual heavy quark-antiquark pair formation is neglected. D_μ can be decomposed into

$$D_\mu = v_\mu(v \cdot D) + D_\mu^\perp \quad \text{with} \quad D_\mu^\perp = (g_{\mu\nu} - v_\mu v_\nu)D^\nu, \quad (2.40)$$

such that $\{\not{D}^\perp, \not{v}\} = 0$. The heavy quark action becomes

$$S_{\text{heavy}} = \int d^4x \left[\bar{\varphi}(iv \cdot D - m_Q)\varphi - \bar{\chi}(iv \cdot D + m_Q)\chi + \bar{\varphi}i\not{D}^\perp\chi + \bar{\chi}i\not{D}^\perp\varphi \right]. \quad (2.41)$$

Factoring out the free-particle propagation, analogous to the interaction picture in perturbation theory,

$$\varphi_v = e^{-im_Q v \cdot x} h_v \quad \text{and} \quad \chi_v = e^{-im_Q v \cdot x} H_v, \quad (2.42)$$

the residual momentum $k = p - m_Q v$ of the heavy quark is only of the order Λ_{QCD} . The action for these fields becomes

$$S_{\text{heavy}} = \int d^4x \left[i\bar{h}(v \cdot D)h - \bar{H}(i(v \cdot D) + 2m_Q)H + i\bar{h}\not{D}^\perp H + i\bar{H}\not{D}^\perp h \right], \quad (2.43)$$

where the field H represents the heavy degree of freedom. In the heavy quark limit, only propagators for h need to be calculated, so H can be integrated out explicitly. Using a Gaussian integral:

$$\begin{aligned} \mathcal{L}_{\text{heavy}} &= i\bar{h}(v \cdot D)h - \bar{H}(i(v \cdot D) + 2m_Q) \left(H - \frac{i}{i(v \cdot D) + 2m_Q} \not{D}^\perp h \right) \\ &\quad + i\bar{h}\not{D}^\perp \left[H - \frac{i}{i(v \cdot D) + 2m_Q} \not{D}^\perp h + \frac{i}{i(v \cdot D) + 2m_Q} \not{D}^\perp h \right] \end{aligned} \quad (2.44)$$

$$\begin{aligned} &= i\bar{h}(v \cdot D)h - \bar{h}\not{D}^\perp \frac{1}{i(v \cdot D) + 2m_Q} \not{D}^\perp h \\ &\quad - \left(\bar{H} - i\bar{h}\not{D}^\perp \frac{1}{i(v \cdot D) + 2m_Q} \right) (i(v \cdot D) + 2m_Q) \left(H - \frac{1}{i(v \cdot D) + 2m_Q} i\not{D}^\perp h \right) \end{aligned} \quad (2.45)$$

where the last line generates a determinant when integrating over $H, \bar{H}$. The tree-level relations are found by expanding the fraction in the second term of (2.45), with the result

$$\mathcal{L} = \bar{h}_v i v D h_v + \bar{h}_v i \not{D}^\perp \frac{1}{2m_Q + i v D} i \not{D}^\perp h_v \quad (2.46)$$

$$= \bar{h}_v i v D h_v + \frac{1}{2m_Q} \bar{h}_v (i \not{D}^\perp)^2 i h_v + \left(\frac{1}{2m_Q} \right)^2 \bar{h}_v (i \not{D}^\perp) (-i v D) (i \not{D}^\perp) h_v + \dots \quad (2.47)$$

$$Q(x) = e^{-im_Q v \cdot x} \left[1 + \frac{1}{2m_Q + i v D} i \not{D}^\perp \right] h_v \quad (2.48)$$

$$= e^{-im_Q v \cdot x} \left[1 + \frac{1}{2m_Q} \not{D}^\perp + \left(\frac{1}{2m_Q} \right)^2 (-i v D) \not{D}^\perp + \dots \right] h_v \quad (2.49)$$

In this way an expansion in $1/m_Q$ of the QCD matrix elements can be obtained, which allows to calculate corrections to the heavy quark limit in a systematic way.

The Heavy Quark limit

The heavy quark limit of the Lagrangian above shows a number of additional symmetries that can help forming a physical picture of a hadron containing a heavy quark. First, the lowest-order terms for the c and b quark in (2.47)

$$\mathcal{L}_{\text{heavy}} = \bar{b}_v(v \cdot D)b_v + \bar{c}_v(v \cdot D)c_v \quad (2.50)$$

show a $SU(2)_{\text{HF}}$ symmetry, with b_v and c_v — with the same velocity — forming a doublet. This is the *heavy flavour symmetry*: the quark-gluon interaction does only depend on the quark flavour through its mass. In the limit where both the beauty and charm mass are infinite, the strong interaction does not see any difference any more. The heavy quark is then an infinitely inert source of a static colour field, that carries a spin of $1/2$ which can be coupled with the spin of the remainder of the hadron, yielding relations between *e.g.* the pseudoscalar and vector meson matrix elements. This is called the heavy quark spin symmetry.

2.3.2 Symmetry arguments

Useful insights and approximate relations between branching ratios for different decays can be obtained from $SU(3)$ flavour symmetry and the $SU(2)$ isospin and $SU(2)$ U -spin symmetry that are part of it [26]; a fundamental representation for each of these symmetries can be formed using the light quarks (u, d, s) , (u, d) and (d, s) , respectively. The light mesons $\pi^\pm, \pi^0, K^\pm$ and $K^0, \bar{K}^0$ are part of the $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$ octet representation of $SU(3)$ (together with the η and η' singlet-octet admixture states), while the heavy mesons $B^\pm, B^0, \bar{B}^0, B_s^0, \bar{B}_s^0, D^\pm, D^0, \bar{D}^0$ and $D_s^\pm$ can be classified into a $\mathbf{3}$ and a $\bar{\mathbf{3}}$ representation for each heavy flavour, according to their light quark content. The diagrams contributing to the $B \rightarrow P\bar{D}$ decays, containing the quark transition $\bar{b} \rightarrow \bar{c}u\bar{q}$, with $q \in \{d, s\}$, can then be found from $SU(3)$ representation multiplications: the weak Hamiltonian transforms as an octet representation, the combination with the initial state $\mathbf{3}$ gives $\mathbf{8} \otimes \mathbf{3} = \mathbf{3} \oplus \bar{\mathbf{6}} \oplus \mathbf{15}$. The final state forms exactly the same representations, so there are three $SU(3)$ amplitudes, transforming as $\mathbf{3}, \bar{\mathbf{6}}$ and $\mathbf{15}$ for these decays. These can be connected to the colour-allowed tree diagrams (T), colour-suppressed tree diagrams (C) and exchange diagrams (E), see fig. 2.2. Additionally to the $SU(3)$ invariant matrix elements, the $SU(3)$ breaking (but isospin invariant) contributions can be constrained. Table 2.1 summarises the results for the decays relevant to this analysis. The $SU(3)$ invariant contributions can easily be found from trying to match the initial and final states in the diagrams with the decays at hand: the annihilation topology is only present in charged B decays, the penguin and penguin annihilation topologies cannot contain a single charm quark in the final state and the colour-suppressed topology produces a neutral charm meson. That leaves the colour-allowed and exchange diagrams, where the exchange diagram cannot contribute to the $B^0 \rightarrow D^-K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ decays that have four different valence quark flavours in the final state — it does correct the $B_s^0 \rightarrow D_s^- K^+$ and $B^0 \rightarrow D^- \pi^+$ matrix elements.

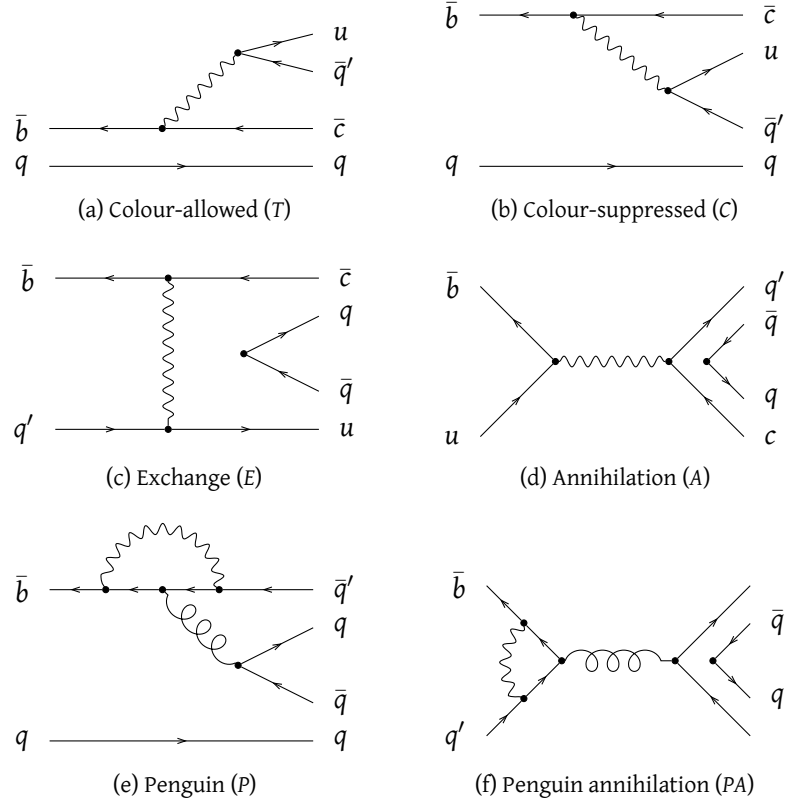


Figure 2.2: Diagrams describing $B \rightarrow P\bar{D}$ and $B \rightarrow PP$ decays; the P and PA diagram do not contribute to $B \rightarrow P\bar{D}$ decays, but are given for completeness since these are important for other B decays. q denotes a light quark; q' denotes a light down-type quark

Table 2.1: Decomposition of the amplitudes for $B_{d,s}^0 \rightarrow D_{(s)}^- h^+$ decays

Decay	SU(3) invariant	SU(3) breaking
$B^0 \rightarrow D^- \pi^+$	$T + E$	
$B^0 \rightarrow D^- K^+$	T	T_1
$B_s^0 \rightarrow D_s^- \pi^+$	T	T_2
$B_s^0 \rightarrow D_s^- K^+$	$T + E$	$T_1 + T_2 + E_1 + E_2$

2.3.3 Factorisation for the decays $B^0 \rightarrow D^- \pi^+$, $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$

A hypothesis that was already used before the heavy quark expansion was formulated, is the factorisation hypothesis. For some weak decays of heavy quarks, the effective four-fermion operator matrix elements can, to a good approximation, be factorised into the product of two current matrix elements, similar to the currents coupling to the W boson in the weak interaction Lagrangian. Effective theories make a more systematic approach possible: the level up to which factorisation provides a good description can be determined, and the expected corrections can be quantified. In [24] the contributions to decay matrix elements for decays into heavy-light final states are ranked according to suppression with powers of 0.2, based on the heavy quark expansion, the structure of a heavy hadrons, decay constants, CKM matrix elements *etc.* Some of the at first sight non-factorisable diagrams turn out to cancel each other, or they can be taken into account as a correction to a factor in a factorised term...

The colour-allowed tree topology is found to be a good candidate for factorisation, while the exchange topology is not. The consequence for the f_s/f_d measurement from $B_{d,s}^0 \rightarrow D_{(s)}^- h^+$ decays is that the correction factors accounting for non-factorisable corrections are relatively small for $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$, while they form a large source of theoretical uncertainty for the $B^0 \rightarrow D^- \pi^+$ decay. The latter decay is however not Cabibbo-suppressed, which makes it attractive for a measurement with a limited data set. The most recent theoretical and experimental information prior to this measurement is described in detail in [1], with the interesting observation that the experimental precision currently reachable allows to test the approximate theories.

Chapter 3

The f_s/f_d analysis

The b quark fragmentation fractions $f_d, f_u, f_s, f_\Lambda \dots$ describe the probability for a b quark produced in a pp interaction to form a B_q meson (where $q \in \{u, d, s, \Lambda \dots\}$). Any B_s^0 branching ratio measurement depends on knowledge of the production rate of B_s^0 mesons, which is commonly determined using an absolute measurement of a B^0 production rate and the f_s/f_d ratio. An important example is the measurement of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay ratio [27, 28], one of the key probes for new physics in LHCb. In the standard model this is a very rare decay, as it contains the $b \rightarrow s\gamma$ transition which only proceeds via loop diagrams, see for example fig. 3.1. Therefore it is very sensitive to contributions from new physics.

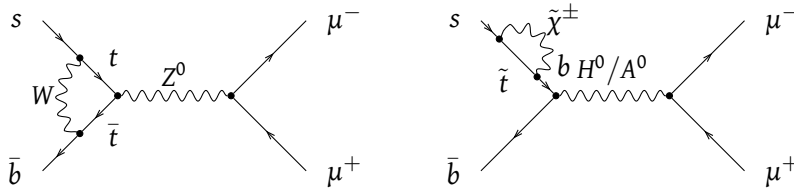


Figure 3.1: Examples of diagrams contributing to the $B_s^0 \rightarrow \mu^+ \mu^-$ decay in the standard model (left) and in the MSSM (right) [27]

Another reason why an analysis of these decay channels is interesting, is the measurement of the CKM-angle γ using $B_s^0 \rightarrow D_s^- K^+$ decays. That analysis will use a simultaneous fit to some of the same decay channels, and as such directly benefits from the experience developed during this work.

3.1 Analysis strategy

Recently it has been proposed that the argument of measuring B_s^0 branching ratios by f_s/f_d to determine the production rate can be reversed: the measurement of the decays $B^0 \rightarrow D^- K^+$ ¹,

¹The notations $B^0 \rightarrow D^- \pi^+$ include the charge conjugated mode $\bar{B}^0 \rightarrow D^+ \pi^-$ in this report, since an untagged analysis is performed

$B^0 \rightarrow D^- \pi^+$, and $B_s^0 \rightarrow D_s^- \pi^+$ can be used to determine f_s/f_d . The method using the ratio of the $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ yields is theoretically the cleanest [29], whereas the ratio of the number of $B^0 \rightarrow D^- \pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ decays [1] leads to a smaller statistical uncertainty (the $B^0 \rightarrow D^- K^+$ branching ratio is Cabibbo-suppressed with respect to the $B^0 \rightarrow D^- \pi^+$ ratio).

The $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ decays receive only contributions from colour-allowed tree topologies, cf. section 2.3.2. The ratio of branching fractions is therefore theoretically well understood and suitable to measure the B_s^0 and B^0 production rate [29]. The ratio of branching ratios, assuming the factorisation approach, is then given by

$$\frac{\text{BR}(B_s^0 \rightarrow D_s^- \pi^+)}{\text{BR}(B^0 \rightarrow D^- K^+)} \approx \frac{\tau_{B_s^0}}{\tau_{B^0}} \left| \frac{V_{ud}}{V_{us}} \right|^2 \times \left(\frac{f_\pi}{f_K} \right)^2 \left[\frac{F_0^{(s)}(m_\pi^2)}{F_0^{(d)}(m_K^2)} \right]^2 \left| \frac{a_1(D_s \pi)}{a_1(DK)} \right|^2, \quad (3.1)$$

where the subsequent factors on the right-hand side are the ratio of the lifetimes, CKM matrix elements, decay constants, form factors and factors representing the deviation from naive factorisation. The observed yield ratio is given by

$$\frac{N_{B_s^0 \rightarrow D_s^- \pi^+}}{N_{B^0 \rightarrow D^- K^+}} = \frac{f_s}{f_d} \frac{\text{BR}(B_s^0 \rightarrow D_s^- \pi^+)}{\text{BR}(B^0 \rightarrow D^- K^+)} \frac{\epsilon_{B_s^0 \rightarrow D_s^- \pi^+}}{\epsilon_{B^0 \rightarrow D^- K^+}}, \quad (3.2)$$

with $\epsilon_{B_s^0 \rightarrow D_s^- \pi^+}$ the detection efficiency for the $B_s^0 \rightarrow D_s^- \pi^+$ decay, similarly for $B^0 \rightarrow D^- K^+$. Taking into account all numerical factors, f_s/f_d can be extracted from the observed yields using

$$\frac{f_s}{f_d} = 0.0743 \cdot \frac{\tau_{B^0}}{\tau_{B_s^0}} \left[\frac{1}{\mathcal{N}_a \mathcal{N}_F} \frac{\epsilon_{B^0 \rightarrow D^- K^+}}{\epsilon_{B_s^0 \rightarrow D_s^- \pi^+}} \frac{N_{B_s^0 \rightarrow D_s^- \pi^+}}{N_{B^0 \rightarrow D^- K^+}} \right], \quad (3.3)$$

where

$$\mathcal{N}_a \equiv \left| \frac{a_1(D_s \pi)}{a_1(DK)} \right|^2 = 1.00 \pm 0.02 \quad [1] \quad (3.4)$$

$$\mathcal{N}_F \equiv \left[\frac{F_0^{(s)}(m_\pi^2)}{F_0^{(d)}(m_K^2)} \right]^2 = 1.24 \pm 0.08 \quad [30]. \quad (3.5)$$

The $B^0 \rightarrow D^- K^+$ decay is, compared to $B^0 \rightarrow D^- \pi^+$, Cabibbo-suppressed. Hence a sizeable statistical uncertainty is expected when using the 2010 data set alone. The $B^0 \rightarrow D^- \pi^+$ decay occurs more abundantly, but is theoretically less clean since it receives contributions from the (non-factorisable) exchange topology. The equation (3.3) is adjusted to [1]

$$\frac{f_s}{f_d} = 0.982 \times \frac{\tau_{B^0}}{\tau_{B_s^0}} \times \left[\frac{1}{\tilde{\mathcal{N}}_a \mathcal{N}_F \mathcal{N}_E} \frac{\epsilon_{D\pi}}{\epsilon_{D_s \pi}} \frac{N_{D_s \pi}}{N_{D\pi}} \right], \quad (3.6)$$

where the additional $\mathcal{N}_E = 0.966 \pm 0.075$ correction factor takes this W -exchange contribution into account, and $\tilde{\mathcal{N}}_a = |a_1(D_s \pi)/a_1(D\pi)|^2$. Both methods are exploited.

In addition to the f_s/f_d measurement, the branching ratios of the suppressed modes $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- K^+$ can be measured. The accuracy reached by previous experiments,

$$\frac{\text{BR}(B^0 \rightarrow D^- K^+)}{\text{BR}(B^0 \rightarrow D^- \pi^+)} = 0.068 \pm 0.015 \pm 0.007 \quad (3.7)$$

and

$$\frac{\text{BR}(B_s^0 \rightarrow D_s^- K^+)}{\text{BR}(B_s^0 \rightarrow D_s^- \pi^+)} = 0.097 \pm 0.018 \pm 0.009, \quad (3.8)$$

is limited. A competitive measurement by LHCb is therefore possible even with a limited data set.

Since the relative yields of the modes $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$ are fitted for, the extraction of the ratio of branching ratios is trivial. The main systematic uncertainties come from the background modelling and the PID efficiency corrections;

$$\frac{\text{BR}(B^0 \rightarrow D^- K^+)}{\text{BR}(B^0 \rightarrow D^- \pi^+)} = \frac{N_{DK}}{N_{D\pi}} \times \frac{\epsilon_{D\pi}}{\epsilon_{DK}}. \quad (3.9)$$

The measured decay modes have the same topology and very similar final states. Therefore, they can be selected using identical kinematic (*e.g.* track momenta) and topological (*e.g.* impact parameter of tracks with respect to the primary vertex) criteria, with the following considerations:

- the favoured decay modes of D_s^- and D^- are $\pi^- K^- K^+$ and $\pi^- \pi^- K^+$ respectively. The reconstructed $D_{(s)}^-$ mass changes dramatically when between the two cases — the $D_{(s)}^-$ and $B_{d,s}^0$ mass are found by assigning a mass hypothesis to the $\pi^\pm$ and $K^\pm$ decay products: the energy of each of these light mesons is calculated from the Einstein relation $E^2 = \vec{p}^2 + m^2$ and the resulting four-momenta are added to obtain that of the parent particle.
- the lifetimes of the D_s^- and D^- meson are approximately 0.5 ps and 1.0 ps respectively. This difference implies that any cut on the $D_{(s)}^-$ flight distance or on the impact parameter of the $D_{(s)}^-$ children with respect to the $D_{(s)}^-$ decay vertex, will have a different efficiency for the B^0 and B_s^0 modes. A correction for this difference is required.

The data set used corresponds to approximately 35 pb⁻¹ of data. “Simulated events” refer to Monte Carlo events simulated using the full LHCb detector simulation, unless specified differently.

3.2 Event selection

The experimental values entering the f_s/f_d formula above are the *relative* branching ratios of the decay channels considered. Studies of the event selection were focused on maximising the signal significance, and on developing a good understanding of any possible efficiency differences between the decay modes — these should however be small, as the selection is designed

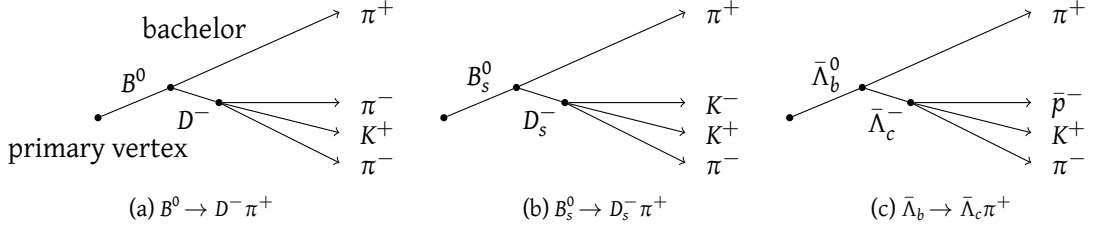


Figure 3.2: Reconstruction of the decay channels $B^0 \rightarrow D^- \pi^+$, $B_s^0 \rightarrow D_s^- \pi^+$ and $\Lambda_b \rightarrow \Lambda_c \pi^-$

to only use topological information, which is expected to be almost identical across the decay channels.

A number of selections are applied between the signals recorded by the detector and the sample for the final invariant mass fit. The first selection is performed online by the trigger system, described in section 3.2.2. After the reconstruction a so-called “stripping selection” is performed, which selects a number of sub samples from the triggered fully reconstructed data set — this selection is executed centrally for the experiment, in order to avoid all individual analyses being run on the complete data sample. The characteristics of this selection step are described in section 3.2.3. Finally a selection specific to this analysis is performed, optimising the significance of the signal with respect to the combinatorial background. This is described in section 3.2.4.

The signal mode efficiencies were evaluated using simulated events, generated under conditions corresponding to the 2010 data taking circumstances.

3.2.1 Reconstruction efficiencies

Between the signal modes, the only difference in the final state is the flavour (and mass) of final state hadrons — the difference between $B^0 \rightarrow D^- \pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ is the D^- (D_s^-) child particle being a π^- or a K^- , respectively; the difference between $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ is the PID of the D^- (D_s^-) child and of the bachelor particle. The small differences in reconstruction efficiency are, for this reason, arising from tracking efficiency differences between π^- and K^- , different momentum spectra of the produced B^0 and B_s^0 mesons, and the lifetime difference of the D^- and D_s^- mesons (1.0 ps and 0.5 ps respectively).

On reconstructed events, the reconstruction efficiencies for $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$ decays are found to be $22.80 \pm 0.07 \%$ and $22.16 \pm 0.07 \%$, while those for $B_s^0 \rightarrow D_s^- \pi^+$ and $B_s^0 \rightarrow D_s^- K^+$ are found to be $21.85 \pm 0.03 \%$ and $21.40 \pm 0.07 \%$. The decays where the bachelor particle is a pion are around 2 – 3 % relatively more likely to be reconstructible than those where it is a kaon. This is mainly caused by the different K^- and π^- lifetimes, while the effects of different material interactions are much less than 1 % [31]. Since these effects are well-understood, the ratio in the simulated events is used as a correction factor and no systematic uncertainty is assigned.

3.2.2 Trigger

A clear trigger strategy has to be specified in order to avoid or at least understand biases caused by this selection. Fully data-driven efficiency studies can be performed, if sufficiently large data samples are available, by using the distinction between Trigger on Signal (TOS) and Trigger independent of Signal (TIS) selected decays: a particle or track² is TOS with respect to a selection (a trigger line, for example) if the requirements of the selection are fulfilled by signals that are associated to that particle or track. Analogously, a particle or track is TIS with respect to a selection if the requirements can be met without using any of the signals associated to that particle or track — it is also possible to have an object that is neither TIS or TOS, but Trigger on Both (TOB): if neither signals belonging to that object and signals not associated to that object fulfil the selection requirements separately, but their combination does; more formally: an object is TOB if it is retained but neither TIS nor TOS. A TIS sample is unbiased, up to the correlation between the B^0 and of other signals in the event (the momentum of the two b quarks produced as a pair is for example correlated), and allows to estimate the trigger efficiency from the sub sample that is also TOS. TIS triggered events can enlarge the data sample usable for the final analysis.

L0

The main L0 trigger for these decay modes is the hadron transverse energy (E_T) trigger, that selects events using the 2×2 hadronic calorimeter cluster with the highest E_T present in the event. Small differences in the B^0 and B_s^0 momentum spectra amount to percent-level efficiency differences, where the K^- efficiency is slightly higher. On the other hand, the probability to decay in flight is substantially higher for K^- than for π^- . Allowing muon TOS triggers would therefore favour the modes containing more kaons. Not considering muon TOS has the additional advantage of substantially reducing the semileptonic B backgrounds.

The strategy for the L0 trigger layer finally is: either L0 global TIS or L0 hadron TOS.

HLT1

The HLT1 triggers underwent a substantial change part of the way through 2010 running, from the “confirmation” style of HLT1 [5, 6, 7, 8] to the new 1-Track HLT1 [9]. The first period cannot simply be disregarded because of the much lower integrated luminosity recorded, because the L0 trigger efficiency was much higher.

Using the same arguments as for L0, HLT1 muon TOS triggers are not considered. Electron and photon TOS triggers were not considered either, although the biases could be expected to be very small since they contribute to only a few percent of the events.

This results in a HLT1 strategy of either HLT1 global TIS or HLT1 Single/DiHadron TOS (first generation) or HLT1-TrackAllL0 TOS (second generation).

²A technically rigorous definition can be found in [32]

HLT2

This analysis uses only the so-called HLT2 topological lines, which also underwent a substantial change during the 2010 running. They were redesigned to be inherently inclusive, *i.e.* to allow one or more tracks to be omitted from the trigger candidate. This eliminates the possibility to apply a cut on the mass of the B candidate *etc.* Tight cuts on the reconstructed vertex quality must also be avoided in order to trigger efficiently on B decays with long-lived resonances. Thanks to its inclusive nature, it can replace a number of other specific HLT2 lines. Two major differences from the exclusive selections are the use of a corrected mass variable, using the missing transverse momentum, instead of widening the B mass window, and a requirement for the sum of the decay products p_T values — rather than a requirement for each of them. See [33] for a detailed description of the HLT2 topological trigger line. Its two most important aspects with regard to this analysis are its high efficiency on $B_{d,s}^0 \rightarrow D_{(s)}^- h^+$ decays³ and the independence of this efficiency on the types of particles.

Trigger efficiency for TCK 0x02e002a

The efficiency of each part of the trigger chain is evaluated from simulated events. The results of this evaluation for the dominant trigger configuration used during 2010 data taking are given in table 3.1. Each efficiency is given relative to the events selected by the previous stage: the L0 efficiency is given with respect to all reconstructed events, the HLT1 TOS efficiency is given with respect to all L0 selected reconstructed events, *etc.* The overall absolute efficiency cannot be extracted from simulated events because the global event cuts⁴ applied cannot meaningfully be reproduced. Since this analysis is concerned with ratio measurements however, the efficiency *differences* are of more interest than the absolute efficiencies.

The L0 efficiencies are found to be very similar for all four of the decay modes under analysis; this is expected because although the B_s^0 mesons take on average 100 MeV more of transverse momentum from the primary vertex than the B^0 mesons, the poor energy resolution (around 20 % for transverse energies around the cut value of 3.6 GeV used for most of the 2011 data taking) of the hadronic calorimeter means that this extra momentum has a negligible effect on the trigger efficiency. Nevertheless, the efficiency for $B_s^0 \rightarrow D_s^- K^+$ is significantly higher than that for $B^0 \rightarrow D^- \pi^+$. Investigating further it turns out modes in which the bachelor track is a kaon tend to have significantly higher efficiencies than those where the bachelor is a pion because the bachelor contributes most of the overall L0 efficiency, and because kaons (at least in simulated events) have a higher efficiency than pions at L0. This is an effect in the 2 % range, and a precise understanding of it is impossible because of the limited sample of simulated events available with which to study it.

On the other hand the HLT1 and HLT2 efficiencies show a significant difference between the B_s^0 and B^0 modes, although they are equal within each B flavour. This can be explained by the lifetime difference between the D_s^- and D^- mesons. HLT1 looks for a single high IP

³For offline-selected simulated $B_{d,s}^0 \rightarrow D_{(s)}^- h^+$ events, 85 % to 90 % of the events passing L0 and HLT1 also pass the HLT2 topological trigger

⁴A selection on global properties of the event, *e.g.* a maximum multiplicity in the SPD, is applied prior to running the HLT algorithms, in order to reduce the CPU load

Table 3.1: Trigger efficiencies for the channels $B_{d,s}^0 \rightarrow D_{(s)}^- h^+$ for TCK 0x02e002a. The efficiency of the L0 stage is given with respect to all events reconstructed in the detector, while the efficiency of the HLT1 and HLT2 stages are relative to reconstructed events selected by L0 and the ensemble of L0 and HLT1 respectively. The efficiency of the topological trigger is a logical OR of the two, three, and four body topological triggers

Decay Mode	L0 TOS Eff.	HLT1 TOS Eff.	HLT2 Topo TOS Eff.
$B^0 \rightarrow D^- \pi^+$	$27.8 \pm 0.2 \%$	$61.7 \pm 0.5 \%$	$64.2 \pm 0.6 \%$
$B^0 \rightarrow D^- K^+$	$28.5 \pm 0.2 \%$	$62.4 \pm 0.5 \%$	$64.0 \pm 0.6 \%$
$B_s^0 \rightarrow D_s^- \pi^+$	$28.3 \pm 0.2 \%$	$58.2 \pm 0.5 \%$	$72.8 \pm 0.6 \%$
$B_s^0 \rightarrow D_s^- K^+$	$28.7 \pm 0.2 \%$	$58.6 \pm 0.5 \%$	$73.1 \pm 0.6 \%$

track, and hence prefers B^0 decays into D^- mesons, since the D^- lives on average about two times longer than the D_s^- and hence gives a greater chance that one of its decay products has a high impact parameter with respect to the primary vertex. Conversely in HLT2 the topological trigger builds a vertex between the bachelor $B_{d,s}^0$ child and the children of the D^- and D_s^- . Here the shorter lifetime of the $D_s^\pm$ is an advantage because its decay products are produced closer to the bachelor and are hence more likely to form a common vertex.

All this means that the trigger efficiencies must be corrected, in order to compare the yields of the different modes under analysis. As only a relative efficiency is needed, this can be done safely from the simulated events, but it must be done for each type of trigger used in the analysis.

3.2.3 Stripping

The stripping selection is based on the requirements for final-state particle tracks and B and D decay vertices. The range of invariant masses for D and B is chosen so as to cover both B_s^0 and B^0 decay modes.

The full list of selection requirements applied at the stripping stage is given in table 3.2. The candidates are built for events with not more than 180 “long” tracks⁵. The D candidates are built from triplets of tracks, and the D candidate is combined with the bachelor track to form the neutral B candidate. The direction angle (DIRA) is defined as the angle between the vector connecting the production and decay vertex of the D^- and its momentum, as found by summing the momenta of its decay products. The distance of closest approach (DOCA) between two tracks is the smallest distance between the extrapolations of these tracks. The vertex χ^2 and $\text{IP}\chi^2$ values give a measure for the significance for the vertex fit or distance calculation respectively: the significance for the hypothesis that the tracks indeed originate from the same vertex, and the significance for the hypothesis that the particle is produced at the vertex.

The efficiencies of the stripping for the different decay modes is evaluated on simulated events, with respect to reconstructible events passing the TOS trigger chain outlined in section 3.2.2, and found to be $75.6 \pm 0.8 \%$ and $76.6 \pm 0.8 \%$ for $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$,

⁵“long tracks” are defined as tracks traversing the full tracking system, from the VELO up to the T stations

Table 3.2: Stripping selection cuts for the $B_{d,s}^0 \rightarrow D_{(s)}^- h^+$ channels. The “Inv. mass” cuts refer to mass windows opened around the nominal D or B mass

Cut Type	Parameter	Value
Global Event Cut	Number of long tracks in the event	≤ 180
D selection	χ^2/ndf for child tracks	< 5
	p_T for child tracks	$> 250 \text{ MeV}$
	p for child tracks	$> 2000 \text{ MeV}$
	$\text{IP}\chi^2$ for child tracks (all)	> 4
	$\text{IP}\chi^2$ for child tracks (max)	> 40
	Inv. mass for D child combination	$\pm 110 \text{ MeV}$
	Inv. mass for D after vertex fit	$\pm 100 \text{ MeV}$
	D vertex fit χ^2/ndf	< 12
	p_T for D	$> 1500 \text{ MeV}$
	D DIRA	> 0.9
	DOCA between D child tracks	$< 1.5 \text{ mm}$
B selection	bachelor p_T	$> 500 \text{ MeV}$
	bachelor p_T	$> 500 \text{ MeV}$
	bachelor p	$> 5000 \text{ MeV}$
	bachelor $\text{IP}\chi^2$	> 16
	B vertex fit χ^2/ndf	< 12
	B $\text{IP}\chi^2$	< 25
	B proper time	$> 0.2 \text{ ps}$
	B DIRA	> 0.9998
	Inv. mass for B child combination	$\pm 500 \text{ MeV}$

respectively, and $75.6 \pm 0.8 \%$ and $76.6 \pm 0.8 \%$ for $B^0 \rightarrow D_s^- \pi^+$ and $B_s^0 \rightarrow D_s^- K^+$, respectively.

3.2.4 Offline selection

Different multivariate selection techniques have been studied, as well as an existing rectangular cut-based selection which served as a reference. The cut-based selection was trained over three years ago, using an obsolete version of the LHCb simulation software. This ensures it is not overtrained. Only the $D_{(s)}^-$ mass window has been increased in order to account for the different mass resolution in real data, and for the radiative tails of the D^- (and, to a lesser extent, D_s^-).

The same reasoning as for the trigger and stripping selections holds: the different lifetime of the D^- and D_s^- mesons can cause a different selection efficiency. This was again studied using simulated events. The efficiencies for $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$ (with respect to reconstructible events), are found to be $32.7 \pm 0.2 \%$ and $32.2 \pm 0.2 \%$, while for $B_s^0 \rightarrow D_s^- \pi^+$ and $B_s^0 \rightarrow D_s^- K^+$ the numbers $32.0 \pm 0.1 \%$ and $32.1 \pm 0.2 \%$ are found. No correction factor or systematic uncertainty needs to be assigned, as this selection is purely used as a reference.

Multivariate selection techniques can give a better performance, since the correlations between the input variables can be exploited. The TMVA software package [34] was used to study various multivariate selection methods, in particular the Fisher Discriminant and the Boosted Decision Tree (BDT).

The linear discriminant (Fisher Discriminant) technique produces a linear combination of the input variables, separating signal and background as much as possible. For better performance, a gaussianisation of the input variables is performed before building the Fisher discriminant (“Fisher Gauss”).

The BDT is a binary tree structured classifier: repeated yes/no decisions are taken on a single variable at a time until convergence criteria are fulfilled. This way, the input variable space is split into many regions, which are eventually classified as signal or background according to the majority of the training events found in the region. The boosting of a decision tree refers to the extension of this concept from one tree to a so-called “forest”, an ensemble of trees. The “gradient” boosting method is used, since this method is less sensitive to over-training (BDTG).

Multivariate selection methods are based on machine learning techniques: the selection is optimised using training samples of background and signal events. An independent sample is used to evaluate the performance after the training stage. In all cases, the final selection is performed by applying a cut on the single “response” variable of the multivariate algorithm.

A small sub sample of the data, corresponding to an integrated luminosity of 2 pb^{-1} , was used as a training sample.

The complementary properties of the Fisher Gauss and the BDTG explain why these methods were studied: the Fisher discriminant is more robust for small training samples, whereas the BDT has a better performance in presence of multiple background sources. The training samples were chosen according to these properties: the Fisher Gauss was trained with a background sample from the upper B^0 mass sideband — eliminating the partially reconstructed B^0 decay backgrounds — and additional selections to obtain a pure combinatorial background sample. The BDTG, on the other hand, was trained using both B^0 mass sidebands from the full

Table 3.3: Input variables for the MVA analyses

Variable	Variable Importance (e-01)
$D_{(s)}^- p_T$	1.479
bachelor p_T	1.460
$B_{d,s}^0 p_T$	1.442
$D_{(s)}^-$ child p_T (min)	1.377
$B_{d,s}^0$ IP χ^2	0.830
$D_{(s)}^-$ vertex χ^2/ndf	0.767
$B_{d,s}^0$ lifetime	0.627
$B_{d,s}^0$ vertex χ^2/ndf	0.620
$B_{d,s}^0$ pointing angle to the primary vertex	0.472
$D_{(s)}^-$ IP χ^2	0.316
bachelor IP χ^2	0.306
$D_{(s)}^-$ -child IP χ^2 (min)	0.304

sample after the stripping selection. The signal sample was taken from simulated $B^0 \rightarrow D^- \pi^+$ events. The selection was optimised with respect to the combinatorial background only, as treating the partially reconstructed backgrounds (see section 3.4) could be done more consistently in the invariant mass fits.

The variables used in the multivariate analyses are listed in table 3.3, sorted according to the “variable importance”, evaluated during the execution of the BDT algorithm. This ranking is derived by counting how often the variable are used to split decision tree nodes, and by weighting each split occurrence by the achieved separation gain-squared and by the number of events in the node. This measure of the variable importance can be used for a single decision tree as well as for a forest [34].

Note that the $D_{(s)}^-$ lifetime is not used in the optimisation, in order not to introduce a difference originating from the different D^- and D_s^- lifetimes in the reconstruction of the B^0 and B_s^0 decay respectively.

The selection performance was evaluated by computing the signal significance in the usual way,

$$\text{Significance} = \frac{S}{\sqrt{S+B}}, \quad (3.10)$$

scaling the signal yield by a factor 1/20, however, to align the training on a $B^0 \rightarrow D^- \pi^+$ signal with the optimisation for the less abundant $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ modes — the factor 1/20 was chosen as a compromise between those two.

The BDT response variable was transformed using a function $f(x) = 1/(1 + e^{-x})$, to bring it into a more usable shape for comparison to the Fisher Gauss when varying the cut value in the range $[0, 1]$.

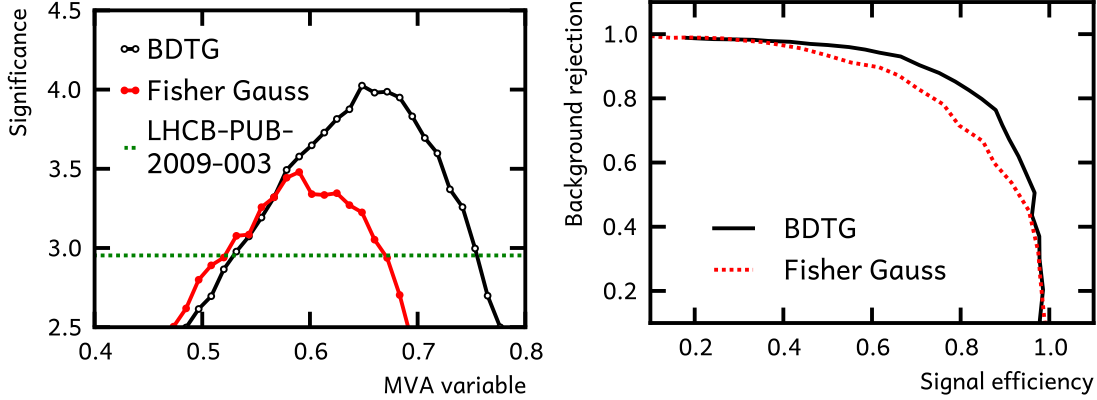


Figure 3.3: Multivariate selection significance as a function of the response variable cut value (left) and background rejection versus signal efficiency (right)

In fig. 3.3 the performance of the cut-based, BDTG and Fisher Gauss selections is shown as a function of the cut value on the response variable, together with the Receiver Operating Characteristic (ROC) curve. This curve was obtained by extracting the number of signal and background events from the fit in the signal region, for different cut values of the multivariate analysis. Both multivariate methods show an improvement of the signal significance over the rectangular cut selection: the Fisher Gauss performs 15 % better, while the BDTG offers a 25 % improvement.

For the optimal cut value for the BDTG selection, 0.65, the selection efficiencies for $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$ decays, with respect to reconstructible events passing the TOS trigger selection (see section 3.2.2) and the stripping selection (see section 3.2.3), are found to be $75.5 \pm 0.8 \%$ and $76.3 \pm 0.8 \%$, respectively. The efficiencies for $B_s^0 \rightarrow D_s^- \pi^+$ and $B_s^0 \rightarrow D_s^- K^+$ are found to be $75.2 \pm 0.4 \%$ and $75.4 \pm 0.8 \%$. As expected from the choice of input variables, the efficiencies are identical for all channels; therefore, no correction factor for this selection step is applied in the final yield ratio evaluation.

3.2.5 Data-Simulation comparisons

A comparison of the distributions of the selection variables between data and the full detector simulation was made, using events passing the TOS trigger chain from table 3.1 and the cut-based offline selection. This method was chosen because it could be applied to data and simulated events in an identical way, and because the resulting invariant mass distribution is sufficiently clean for no further background subtraction to be required. A 2σ window around the $B_{d,s}^0$ mass was selected, in order to further reduce the background contributions. The distributions of simulated and data events are then inspected visually, after normalisation to unity. As can be seen from figs. 3.4 to 3.8 there is generally a good agreement between the distribution from data and from simulation. The impact parameter χ^2 distribution for the B candidate in data is shifted to higher values: this is because the simulation of the VELO gives better performance for the impact parameter resolution than is seen in the data.

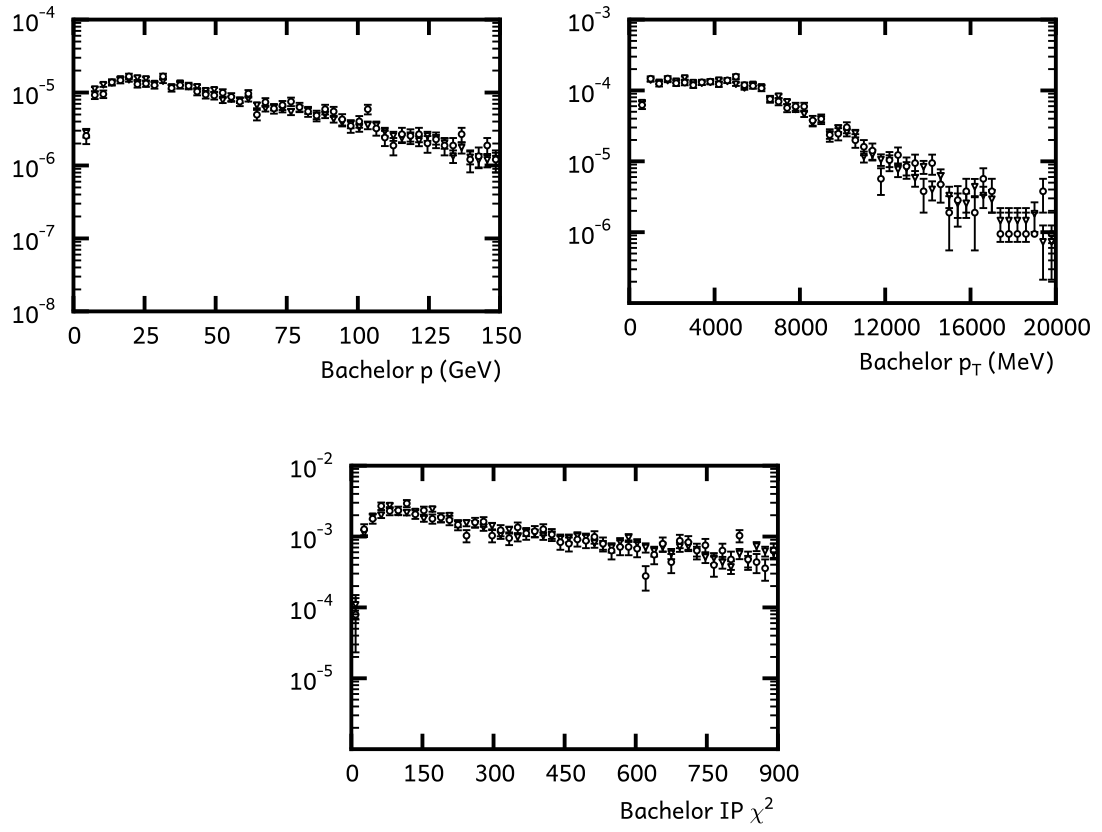


Figure 3.4: Comparison between data (circles) and simulated event (triangles) distributions for selection variables relating to the bachelor. Left: momentum; right: transverse momentum; bottom: impact parameter χ^2 to the primary vertex

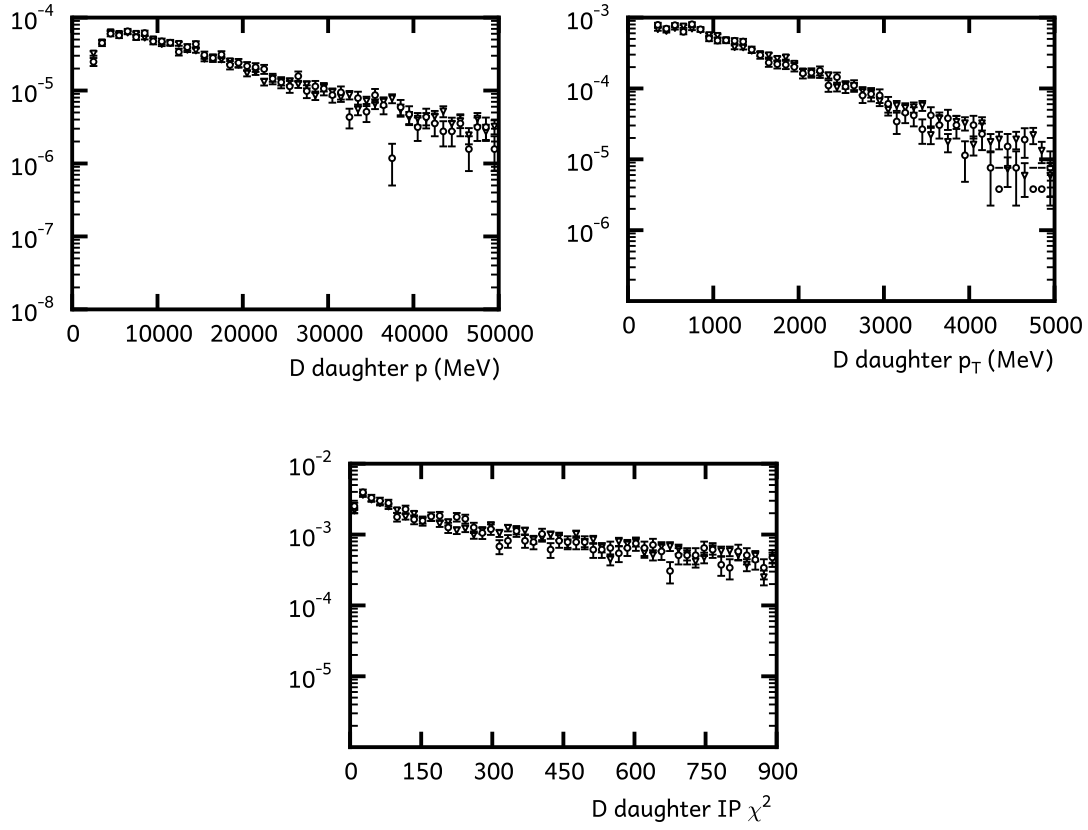


Figure 3.5: Comparison between data (circles) and simulated event (triangles) distributions for selection variables relating to the first π^- daughter of the D^- . Left: momentum; right: transverse momentum; bottom: impact parameter χ^2 to the primary vertex

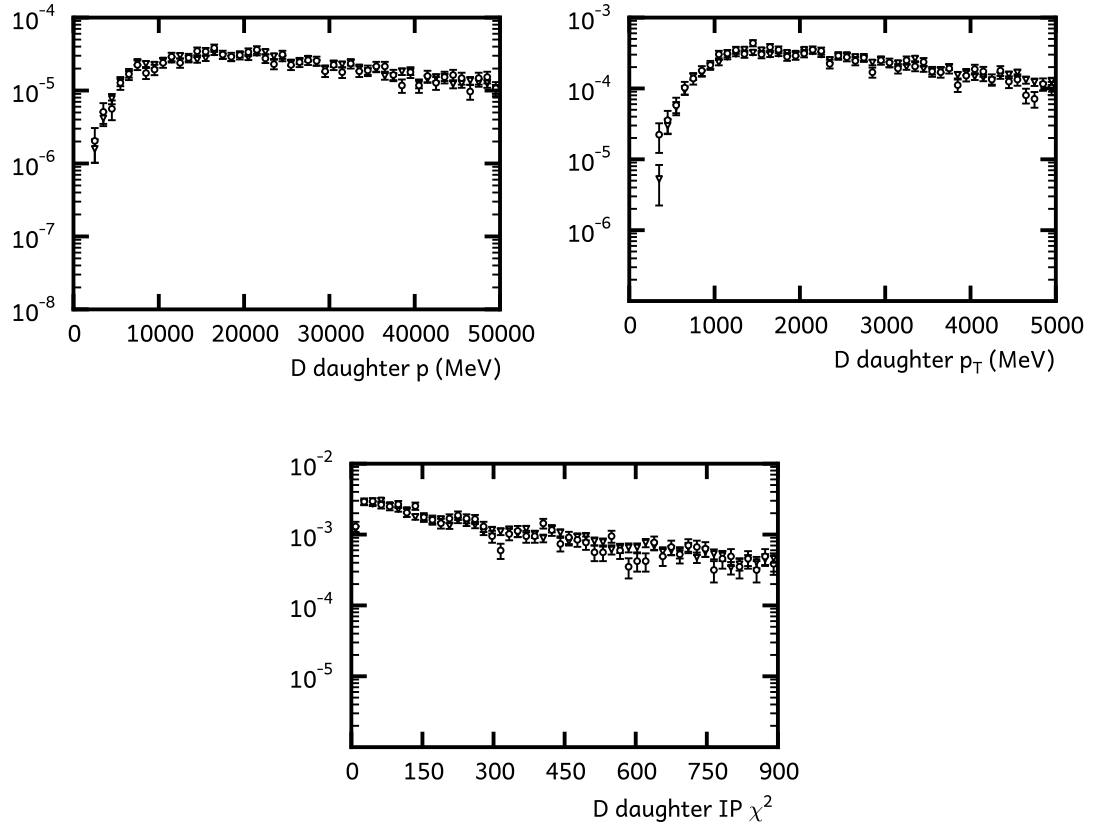


Figure 3.6: Comparison between data (circles) and simulated event (triangles) distributions for selection variables relating to the second π^- daughter of the D^- . Left: momentum; right: transverse momentum; bottom: impact parameter χ^2 to the primary vertex

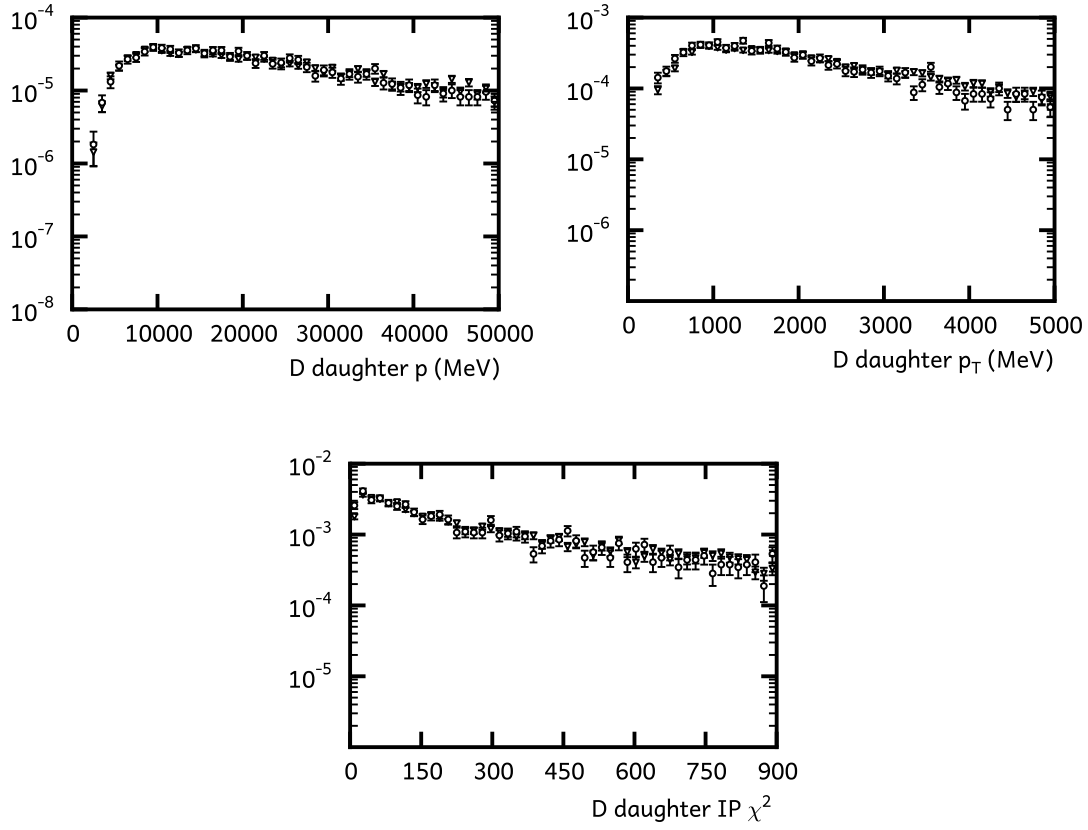


Figure 3.6: Comparison between data (circles) and simulated event (triangles) distributions for selection variables relating to the K^- daughter of the D^- . Left: momentum; right: transverse momentum; bottom: impact parameter χ^2 to the primary vertex

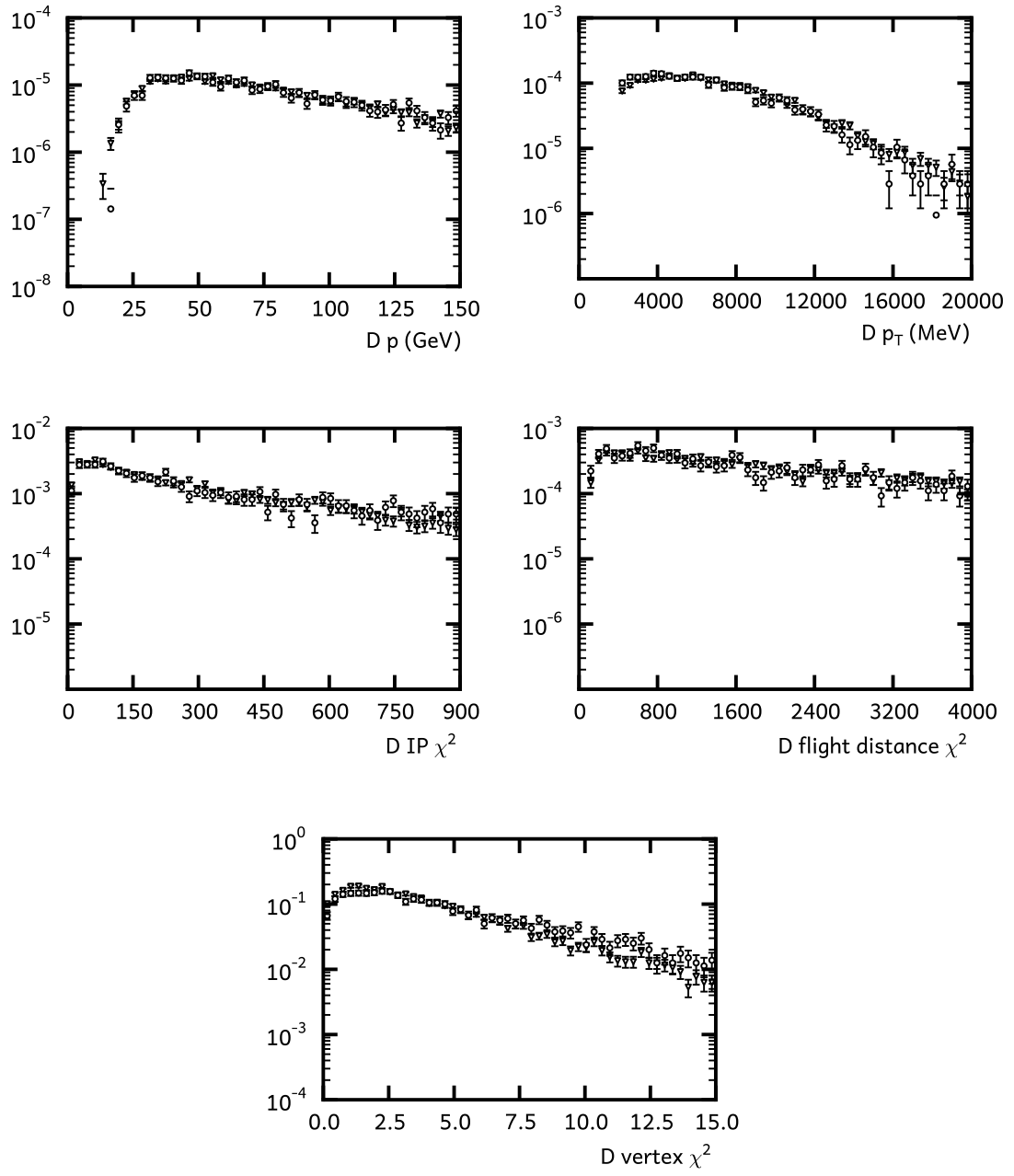


Figure 3.7: Comparison between data (circles) and simulated event (triangles) distributions for selection variables relating to the D^- candidate. Top left: momentum; top right: transverse momentum; middle left: impact parameter χ^2 to the primary vertex; middle right: flight distance χ^2 from the primary vertex; bottom: χ^2 of the D^- vertex

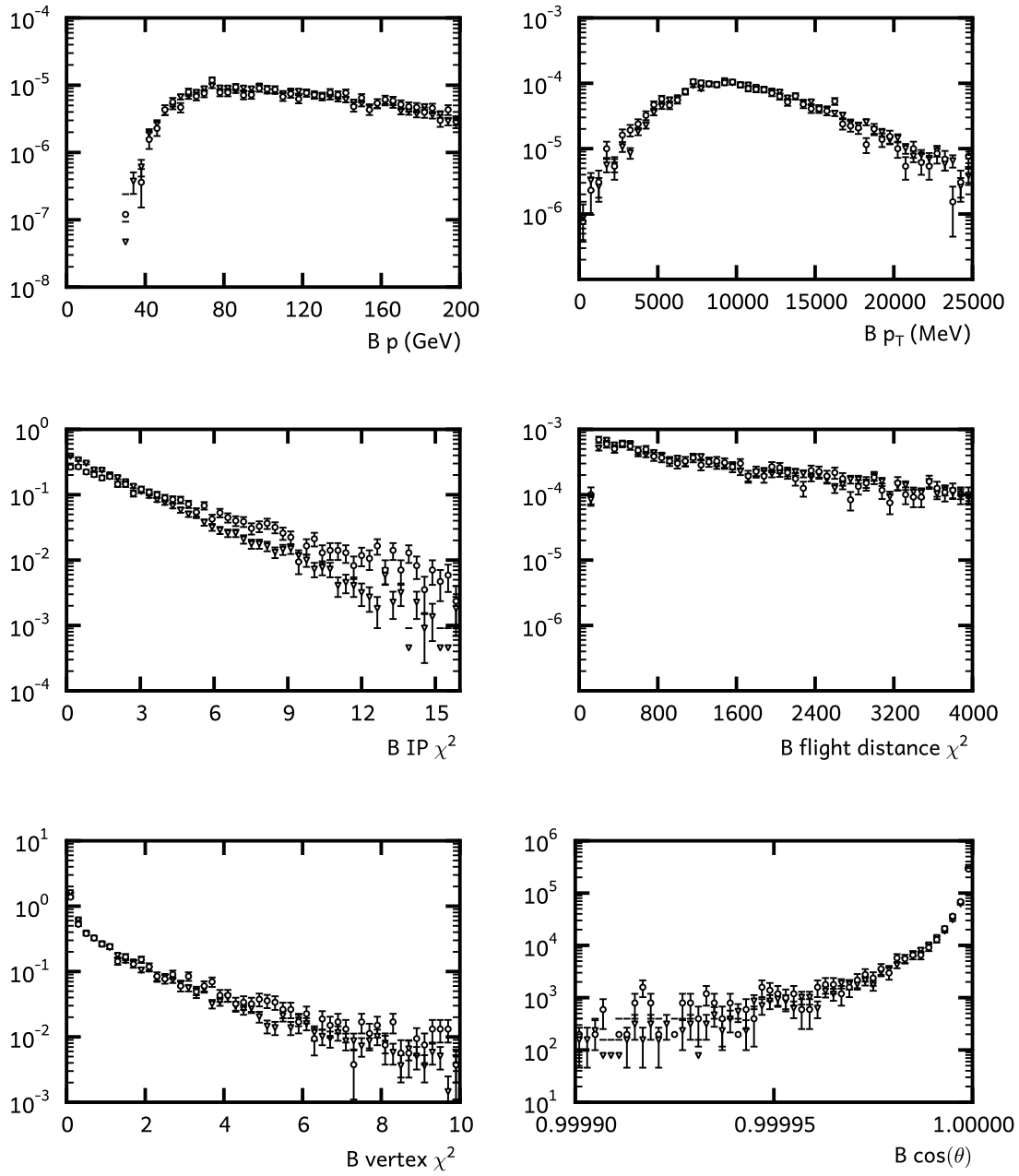


Figure 3.8: Comparison between data (circles) and simulated event (triangles) distributions for selection variables relating to the B^0 candidate. Top left: momentum; top right: transverse momentum; middle left: impact parameter χ^2 to the primary vertex; middle right: flight distance χ^2 from the primary vertex; bottom left: χ^2 of the B^0 vertex; bottom right: cosine of the direction angle with respect to the primary vertex

It would be very surprising if the data-simulation comparison strongly depended on the B flavour, but nevertheless this assumption can be verified by explicitly comparing the distributions for Λ_b baryons. This decay is selected alongside the signal modes, due to the absence of PID selection up to this stage, and can be isolated by recomputing the invariant mass under the appropriate mass hypotheses for the decay products, and applying a hard PID cut on the proton-kaon separation to distinguish $\Lambda_b \rightarrow \Lambda_c \pi^-$ from $B_s^0 \rightarrow D_s^- \pi^+$ — the invariant masses are close to each other. The distributions show a similar behaviour as those for the B^0 mesons (the figures are left out for brevity).

3.2.6 Veto of peaking backgrounds

As neither the rectangular cut selection nor the multivariate selections explicitly requires a certain separation between the D and B decay vertices, a small fraction of $B^0 \rightarrow J/\psi K^{*0}$ decays can be reconstructed as $B^0 \rightarrow D^- \pi^+$. Because of the small difference between the muon and pion masses, this contribution will have a peaked shape under the $B^0 \rightarrow D^- \pi^+$, but not under the $B^0 \rightarrow D^- K^+$ hypothesis. Events where two of the final-state tracks are identified as muons, and these two tracks have an invariant mass within 40 MeV from the J/ψ mass, are therefore removed from the data sample.

One also expects final states with two muon tracks outside the J/ψ window to contribute to the background, for example originating from semileptonic decays where a D decay product decayed in flight (before hitting the muon stations), or a signal decay with two in-flight decays. No veto is needed for these cases, since they do not exhibit a sharply peaking behaviour in the vicinity of the B^0 mass.

3.2.7 Treatment of multiple candidates

Approximately 2 % of the events surviving the full offline selection contain multiple candidates, usually *clones*, an artifact of the track reconstruction where the hits of one particle causes two tracks to be reconstructed. In this case, the candidate with the best vertex χ^2 is retained. This selection, together with the $D_{(s)}^-$ and $B_{(s)}^0$ invariant mass windows and $\text{DLL}_{K\pi}$ cuts, reduces the contribution of the “double-misidentified” background, where two $D_{(s)}$ decay products are assigned the wrong mass hypothesis, cf. fig. 3.2b — the reconstruction output contains all possible combinations compatible with the track charges.

3.2.8 Final efficiencies

As can be seen from the preceding discussion, the reconstruction and selection efficiencies are very similar between decay modes of the same B flavour, while there are significant differences between decay modes of different B flavours. The largest difference is found in the trigger efficiency. The overall efficiency numbers corresponding to events triggered by the TCK 0x02e002a are collected for convenience in table 3.4

From the previous discussions, two distinct corrections were found that need to be computed: the reconstruction and the trigger efficiency.

Table 3.4: Overall reconstruction, trigger, stripping, and BDTG selection efficiencies for the decay modes $B^0 \rightarrow D^- \pi^+$, $B^0 \rightarrow D^- K^+$, $B_s^0 \rightarrow D_s^- \pi^+$ and $B_s^0 \rightarrow D_s^- K^+$. Each efficiency is given relative to the previous stages. The trigger efficiency listed here is for the TOS trigger chain considered in table 3.1.

Decay Mode	Reconstruction	Trigger	Stripping	Selection
$B^0 \rightarrow D^- \pi^+$	$22.80 \pm 0.07 \%$	$11.0 \pm 0.1 \%$	$75.6 \pm 0.8 \%$	$75.5 \pm 0.8 \%$
$B^0 \rightarrow D^- K^+$	$22.16 \pm 0.07 \%$	$11.3 \pm 0.1 \%$	$76.6 \pm 0.8 \%$	$76.3 \pm 0.8 \%$
$B_s^0 \rightarrow D_s^- \pi^+$	$21.85 \pm 0.03 \%$	$12.0 \pm 0.05 \%$	$75.6 \pm 0.4 \%$	$75.2 \pm 0.4 \%$
$B_s^0 \rightarrow D_s^- K^+$	$21.40 \pm 0.07 \%$	$12.3 \pm 0.1 \%$	$75.5 \pm 0.8 \%$	$75.4 \pm 0.8 \%$

Reconstruction efficiency correction

As anticipated in section 3.2.1, the difference between decay modes involving a kaon or a pion bachelor must be corrected for. From the simulated data, a correction factor of 1.029 ± 0.004 is found for the B^0 modes, while a factor 1.021 ± 0.004 is found for the B_s^0 modes. Part of this difference is due to kaon decay inside the detector, hence the efficiency difference will depend on low momentum, and be more pronounced at low momentum.

The numbers quoted above describe the difference between the modes with a pion and a kaon bachelor — given the similar momentum spectrum of B^0 and B_s^0 , this can be averaged to 1.025 ± 0.003 .

For the f_s/f_d measurement from $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$, the difference is caused by the different momentum spectrum of the bachelor particle and the D decay product. The factor from simulated data is $\epsilon_{B^0 \rightarrow D^- K^+} / \epsilon_{B_s^0 \rightarrow D_s^- \pi^+} = 1.014 \pm 0.004$, significantly different from unity — the higher efficiency for $B^0 \rightarrow D^- K^+$ is expected, because the bachelor has a harder momentum spectrum than the D products.

For the other f_s/f_d measurement, from $B^0 \rightarrow D^- \pi^+$ and $B_s^0 \rightarrow D_s^- \pi^+$, a larger efficiency difference is expected — $B^0 \rightarrow D^- \pi^+$ has a pion in the D decay where $B_s^0 \rightarrow D_s^- \pi^+$ has a relatively low-momentum kaon. The average of $\epsilon_{B^0 \rightarrow D^- \pi^+} / \epsilon_{B_s^0 \rightarrow D_s^- \pi^+}$ and $\epsilon_{B^0 \rightarrow D^- K^+} / \epsilon_{B_s^0 \rightarrow D_s^- K^+}$ can be used, as they both show exactly this difference: 1.043 ± 0.004 and 1.036 ± 0.004 are found, averaged to 1.039 ± 0.003 .

Overall selection efficiency correction

The computation of the trigger efficiency correction has the additional complexity that it can depend on the trigger configuration used. Moreover, the TIS events, which are selected with equal efficiency across all signal modes, need to be accounted for. The corrections needed to be extracted from simulated events, because not enough events are available for a data-driven (TISTOS, [32]) determination. The data was split into categories (HLT global TIS, HLT1 hadron and HLT2 topological TOS and HLT1 1Track and HLT2 topological TOS, for all different configurations) after which the efficiencies were determined for each category separately and averaged according to the contribution of each category to the final data sample. The efficiency was computed as that of $B^0 \rightarrow D^- \pi^+$ with respect to $B_s^0 \rightarrow D_s^- \pi^+$. The equivalence of $B^0 \rightarrow$

$D^-\pi^+$ and $B^0 \rightarrow D^-K^+$ has been demonstrated for the main trigger configuration key (TCK) and there is no physical reason to expect a difference for the other configurations. The weighted trigger efficiency ratio between B_s^0 and B^0 modes was found to be 1.05 ± 0.02 .

3.3 Signal model

The sample selected by the offline selection is finally split up into sub samples for the extraction of the yields, based on PID information. Due to the fact that binary cuts on the $DLL_{K\pi}$ are used to do this, it is unavoidable that events of one decay mode are misidentified, and crossfeed into the sample for the extraction of another decay mode — with a different mass hypothesis. Therefore, it is necessary to model the invariant mass shape of all signal modes, both under the correct and under wrong mass hypotheses.

3.3.1 Signal shape under the correct mass hypothesis

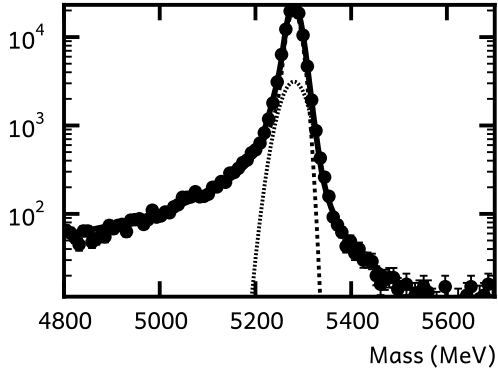
The two main effects determining the shape of the signal mass distribution are the detector resolution and radiative losses in the B and D meson decays. Both the B and the D decay products can radiate photons, which are not reconstructed as part of the B decay [35]. This results in a lower invariant mass for the measured final state — the mass distribution has a tail on the lower side, which can be described by a Crystal Ball (CB) function, a Gaussian with one exponential tail [36, 37, 38].

The estimation of the signal shape was done on large samples of simulated events. The best description was found using the sum of two CB functions, with tails in the opposite direction, with the same mean value. The description of the upper mass region by another CB is purely empirical, and does not imply any asymmetric detector effects: it is not surprising that a model where both outer regions are described by separate, mostly independent parameters, is able to describe the distribution more accurately — a similar-sized contribution on the lower mass side could easily be absorbed by the radiative tail. The sensitivity of the final result to the signal model was studied, and found to be small. The result of this procedure is shown in fig. 3.9; the fitted parameters values can be found in table 3.5.

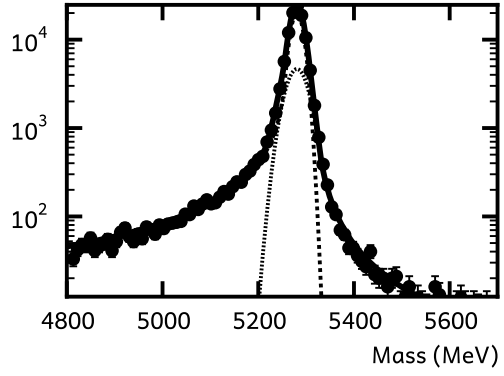
3.3.2 Signal shape under a wrong mass hypothesis

The crossfeed between different decay channels is present for all possible combinations. The difference between the branching ratios, complemented with PID cuts, makes some contributions negligible however. This is discussed in more detail in section 3.5.

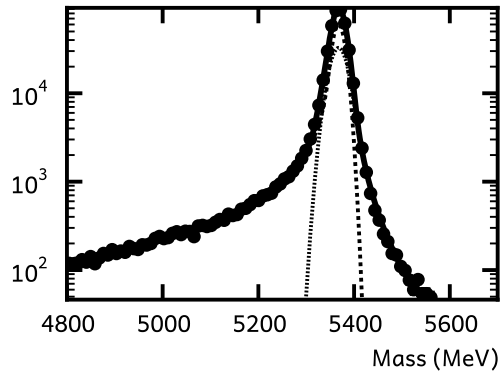
The invariant mass shapes cannot straightforwardly be found by selecting a clean sample of the crossfeeding decay mode using $DLL_{K\pi}$ cuts and changing the mass hypothesis, because the $DLL_{K\pi}$ distributions depend on a number of variables: most importantly p , next p_T and the total number of tracks in the event...The invariant mass distribution under the wrong mass hypothesis also depends on the bachelor momentum, cf. fig. 3.10. A reweighting procedure is used to extract the shapes from data, correcting for the momentum dependence. The dependence on other variables is much smaller, so no reweighting according to these is performed



(a) $B^0 \rightarrow D^- \pi^+$



(b) $B^0 \rightarrow D^- K^+$



(c) $B_s^0 \rightarrow D_s^- \pi^+$

Figure 3.9: Signal shapes, fitted to the distribution of fully simulated events

Table 3.5: Fitted values of parameters for the sum of two CB functions describing the signal shape of $B^0 \rightarrow D^- \pi^+$, $B^0 \rightarrow D^- K^+$, $B_s^0 \rightarrow D_s^- \pi^+$ and $B_s^0 \rightarrow D_s^- K^+$. Both CB functions use the same mean value. The parameter “fraction” is the fraction of the first CB function in the total normalisation

Parameter	$B^0 \rightarrow D^- \pi^+$	$B^0 \rightarrow D^- K^+$	$B_s^0 \rightarrow D_s^- \pi^+$	$B_s^0 \rightarrow D_s^- K^+$
Mean	5280.47 ± 0.06	5280.75 ± 0.06	5367.61 ± 0.04	5367.58 ± 0.05
α_1	1.79 ± 0.02	1.82 ± 0.03	1.56 ± 0.01	1.70 ± 0.01
α_2	-2.08 ± 0.04	-2.04 ± 0.04	-1.74 ± 0.07	-1.94 ± 0.18
n_1	1.06 ± 0.02	1.09 ± 0.02	1.24 ± 0.01	1.19 ± 0.01
n_2	1.26 ± 0.08	1.48 ± 0.08	3.04 ± 0.37	2.43 ± 0.55
σ_1	14.22 ± 0.15	13.29 ± 0.16	12.99 ± 0.05	12.28 ± 0.10
σ_2	26.46 ± 0.90	22.91 ± 0.80	18.95 ± 0.39	18.34 ± 0.25
fraction	0.79 ± 0.02	0.73 ± 0.02	0.59 ± 0.02	0.55 ± 0.01

— an appropriate systematic is assigned.

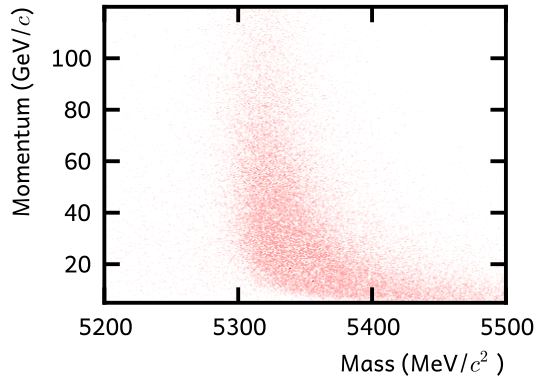


Figure 3.10: Illustration of the correlation between momentum and the invariant mass under the $B^0 \rightarrow D^- K^+$ hypothesis for $B^0 \rightarrow D^- \pi^+$ events (signal MC)

As an example, the procedure is outlined here for the $B^0 \rightarrow D^- \pi^+$ crossfeed into $B^0 \rightarrow D^- K^+$. In that case, the bachelor π^+ is misidentified as a K^+ . The right pane of fig. 3.11 illustrates that it is impossible to completely suppress the $B^0 \rightarrow D^- \pi^+$ background without losing a significant part of the signal sample — as much signal events as possible are needed for a good mass fit. The invariant mass shape under the $B^0 \rightarrow D^- K^+$ hypothesis is obtained from data as follows:

- First, a clean sample of $B^0 \rightarrow D^- \pi^+$ is extracted from a tight mass window around the B^0 mass and by applying a cut $\text{DLL}_{K\pi} < 0$ on the bachelor particle.

The efficiency, *i.e.* the probability for a K to have a $\text{DLL}_{K\pi}$ value above the cut value, and misidentification, *i.e.* the probability for a π to have a $\text{DLL}_{K\pi}$ value above the cut value,

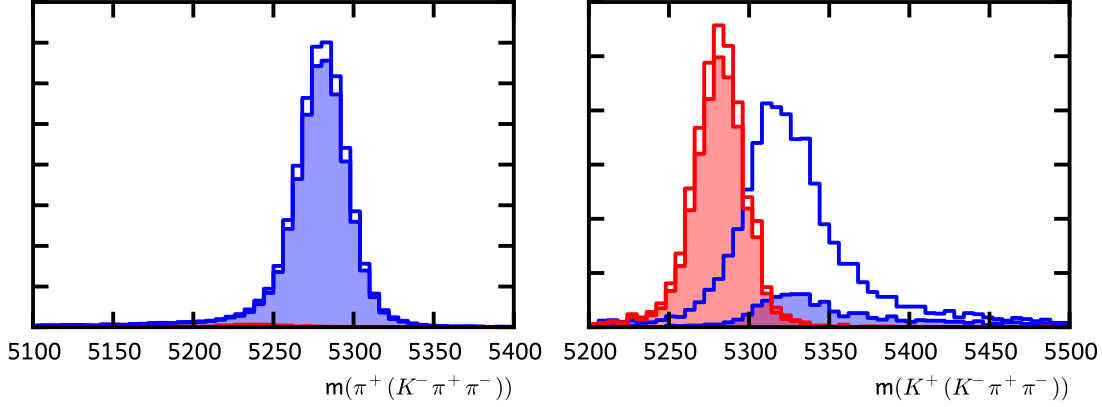


Figure 3.11: Effect of a $\text{DLL}_{K\pi}$ cut on the $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$ invariant mass distribution (purely illustrative, using the world-average measurement of the relative branching ratio and simulated events). The curves from top to bottom are for a $\text{DLL}_{K\pi}$ cut at 0 and at 5. Left: sub sample below the cut value, under the $B^0 \rightarrow D^- \pi^+$ mass hypothesis; Right: sub sample above the cut value, under the $B^0 \rightarrow D^- K^+$ mass hypothesis

for this $\text{DLL}_{K\pi}$ cut as a function of momentum are obtained from a D^* sample, used to calibrate the RICH detector (corresponding to the same running conditions as the data sample). These probabilities are shown in fig. 3.12.

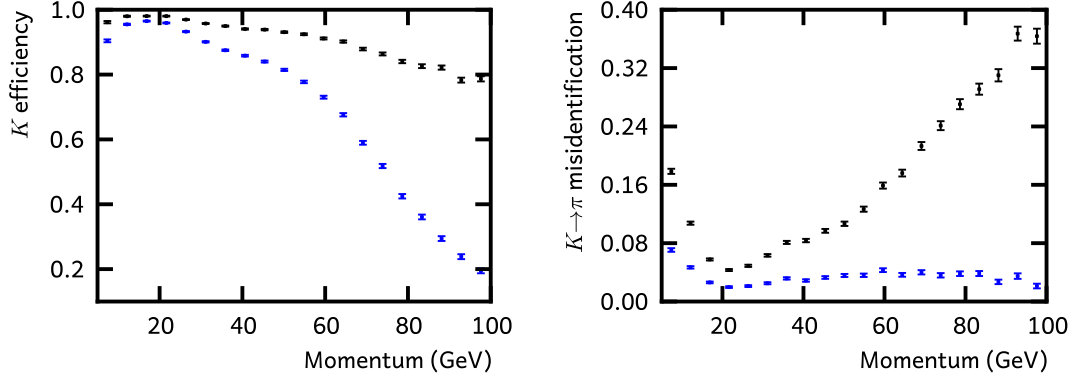


Figure 3.12: $\text{DLL}_{K\pi}$ cut calibration: efficiency to correctly identify a kaon (left) and to misidentify a π as a K (right), for $\text{DLL}_{K\pi}$ cut values of 0 (black, upper points) and 5 (blue, lower points)

- The “clean $B^0 \rightarrow D^- \pi^+$ ” sample now has a momentum distribution that is different from the momentum distribution before the PID cut. We assume that the momentum distributions for the bachelor in $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$ is the same: simulated events show a difference of a few percent. Taking into account the much smaller size of the $B^0 \rightarrow D^- K^+$ contribution, this is a reasonable approximation. The ratio of the momen-

tum distributions of this sample and the sample before the $DLL_{K\pi}$ cut (still in the tight mass window of the previous step) is used to reweight these events according to the bachelor momentum, restoring the $B^0 \rightarrow D^- \pi^+$ bachelor momentum distribution before any $DLL_{K\pi}$ cut, see fig. 3.13.

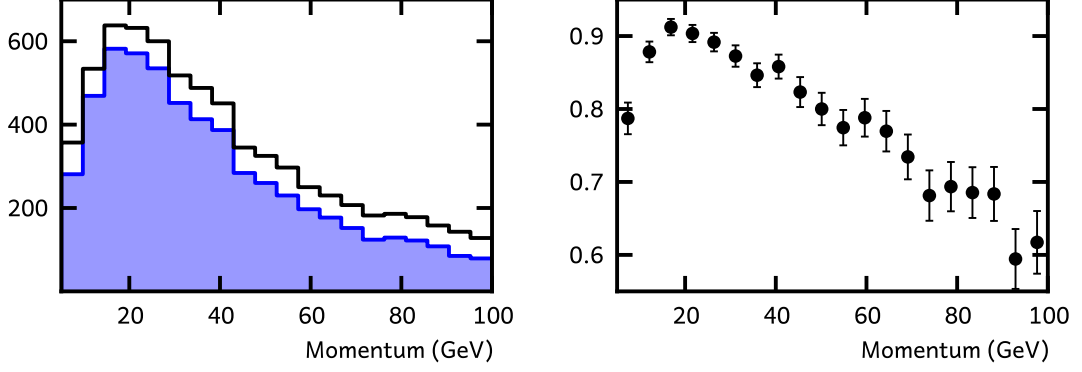


Figure 3.13: Momentum distribution a sample selected in a tight $B^0 \rightarrow D^- \pi^+$ signal window, and its sub sample with $DLL_{K\pi} < 0$ (left), and the ratio of these distributions (right)

- Next, the misidentification probabilities from the RICH calibration, for the cut value used to isolate the $B^0 \rightarrow D^- K^+$ sample, are used to reweight the $B^0 \rightarrow D^- \pi^+$ sample again according to the bachelor momentum. The mass is recalculated under the $B^0 \rightarrow D^- K^+$, and the result is our approximation to the $B^0 \rightarrow D^- \pi^+$ crossfeed. The misidentification probabilities depend on the $DLL_{K\pi}$ cut value, a value of 5 was chosen to reduce the $B^0 \rightarrow D^- \pi^+$ background. For this value, the shape does not change dramatically by the reweighting (see fig. 3.14). For a $DLL_{K\pi}$ cut value of 0, however, the difference is much more pronounced. This is not unexpected from the misidentification probabilities, which are flat over a large part of the momentum range.
- Finally, kernel estimation [39] is used to construct a probability density function for the invariant mass.

3.4 Background model

The offline selection discussed in section 3.2.4 was trained only to reject pure combinatorial background, where four random tracks (*i.e.* not originating from a real B or D decay) are reconstructed as a $B_{d,s}^0 \rightarrow D_{(s)}^- h^+$ decay. Other contributions that are unavoidably present in the data sample, are prompt charm background (where a correctly reconstructed D is combined with a random bachelor track), partially reconstructed B decays, where one or more final-state particles are not reconstructed as part of the B decay: neutral particles, for example. The crossfeed of other “signal” modes is discussed in section 3.3.2 and section 3.5.

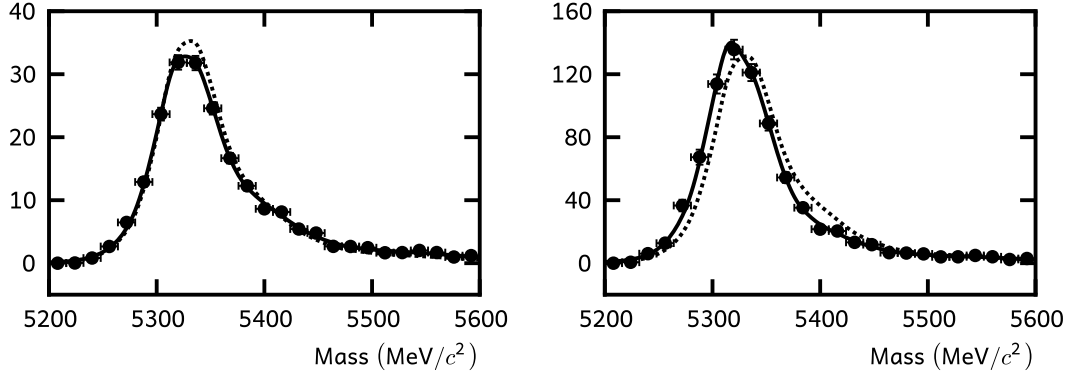


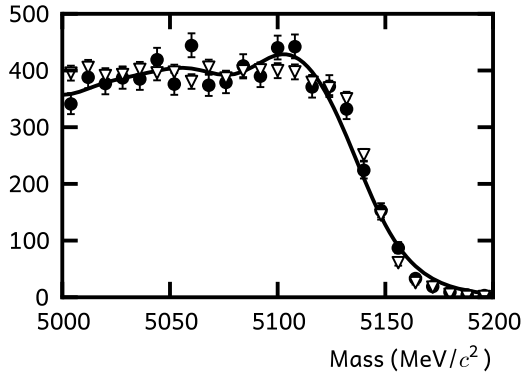
Figure 3.14: $B^0 \rightarrow D^- \pi^+$ background shape to $B^0 \rightarrow D^- K^+$, when a $DLL_{K\pi}$ cut at 5 (left) or at 0 (right) is used to isolate the bachelor K . The dashed line represents the distribution before reweighting.

Below an early background model is discussed. For the final background model generator-level studies were performed to identify the most important backgrounds and samples of the specific modes were generated with the full detector simulation — the backgrounds are then processed using exactly the same selection criteria as the data. This model is nevertheless an interesting exercise: the general features of all background shapes can be reproduced from kinematic arguments and resolution effects.

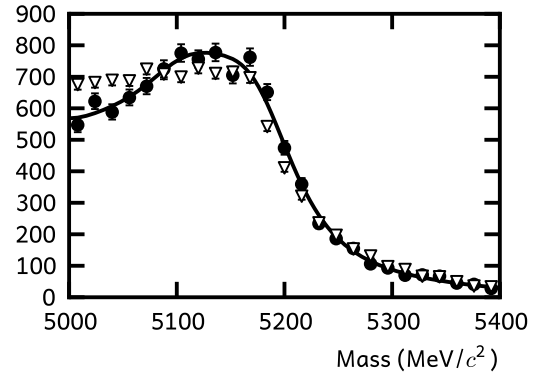
3.4.1 Phase space background model

Two prime examples of partially reconstructed backgrounds to the $B^0 \rightarrow D^- \pi^+$ are $B^0 \rightarrow D^{*-} \pi^+$ and $B^0 \rightarrow D^- \rho^+$ decays. In the first case, when the π^0 (dominant) or γ from the $D^{*-} \rightarrow D^- \pi^0$ or $D^{*-} \rightarrow D^- \gamma$ decay is not reconstructed; in the second case, the π^0 from the $\rho^+ \rightarrow \pi^+ \pi^0$ decay is not seen. In both cases, the invariant mass distribution is shifted to lower mass values than would be obtained from the fully reconstructed decay. In the case where the undetected particle carries a low momentum however, *e.g.* $B^0 \rightarrow D^- \rho^+$, a significant tail of the distribution entering the signal region is possible.

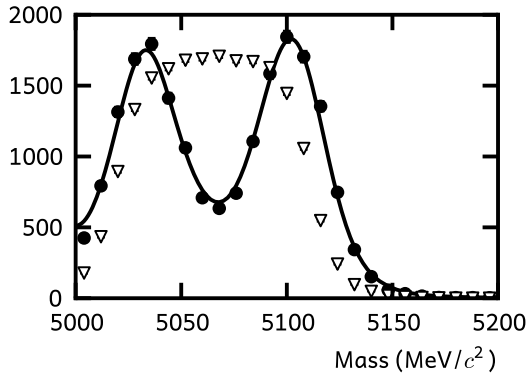
Since these backgrounds were expected to heavily complicate the mass fits, a simple model was developed to estimate their shapes in a very early stage, long before the specific simulation samples were prepared. In this model, only the phase space distribution was taken into account. Although this is a rudimentary approximation, the invariant mass shape and momentum distribution were, for most of the decay channels considered, reproduced in a sufficiently accurate way. An important exception to this is the $B^0 \rightarrow D^{*-} \pi^+$ decay, where an angular dependence is generated by the spin of the D^{*-} particle. A simple correction favouring small (forward and backward) angles between the momenta and the D^{*-} spin and the momenta indeed showed a rough double peak structure. See fig. 3.15 for a comparison between simple toy simulation and real simulation shapes for the $B^0 \rightarrow D^{*-} \pi^+$ and $B^0 \rightarrow D^- \rho^+$ backgrounds.



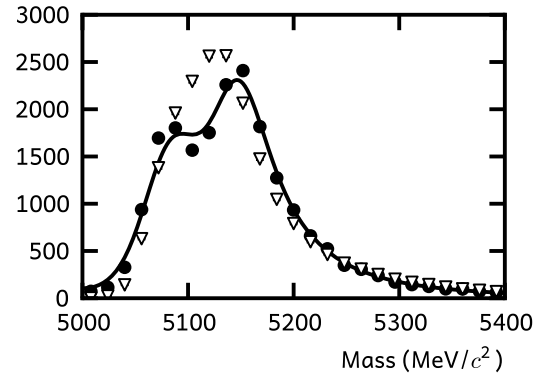
(a) $B^0 \rightarrow D^- \rho^+$ background to $B^0 \rightarrow D^- \pi^+$



(b) $B^0 \rightarrow D^- \rho^+$ background to $B^0 \rightarrow D^- K^+$



(c) $B^0 \rightarrow D^{*-} \pi^+$ background to $B^0 \rightarrow D^- \pi^+$



(d) $B^0 \rightarrow D^{*-} \pi^+$ background to $B^0 \rightarrow D^- K^+$

Figure 3.15: Comparison between the toy partially reconstructed model (triangles), the full simulation data after selection (circles) and the kernel estimated version of the latter.

3.5 PID separation

A good understanding of the PID information is of crucial importance to this analysis. The influence of the $DLL_{K\pi}$ cut value on the crossfeed shape has been shown in section 3.3.2. More details about the use of $DLL_{K\pi}$ cuts are discussed in this section.

The data sample after the offline selection is split into four sub samples for the invariant mass fits, based on $DLL_{K\pi}$ cuts for the bachelor particle and the D decay product that is different between the B^0 and B_s^0 decay modes. This approach is used in preference to using the full $DLL_{K\pi}$ likelihood in a simultaneous fit to the mass and the $DLL_{K\pi}$ distribution, because that would require a good understanding of the $DLL_{K\pi}$ distributions of all backgrounds — the combinatorial background is problematic in this regard: it contains a mixture of pions, kaons and ghosts, with possibly different momentum distributions. If these issues are not accounted for, the simultaneous fit may become biased [40]. Given the relatively small amount of data available and the time constraints, it was not felt worthwhile to invest in such studies at present.

The $D_{(s)}^-$ candidates must be classified as D^- or D_s^- . Two issues need to be considered:

- the crossfeed between D^- and D_s^- ;
- the contamination of $\Lambda_b \rightarrow \Lambda_c \pi^-$, where $\Lambda_c \rightarrow p K^- \pi^+$ to both D^- and D_s^- .

3.5.1 Crossfeed between D^- and D_s^-

The $B_s^0 \rightarrow D_s^- \pi^+$ background under the $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$ signal is reduced using loose $DLL_{K\pi}$ cuts on the D^- decay products: $DLL_{K\pi} < 10$ for the pions and $DLL_{K\pi} > 0$ for the kaons. This, combined with the difference in branching ratios, reduces the $B_s^0 \rightarrow D_s^- \pi^+$ contribution under $B^0 \rightarrow D^- \pi^+$ to a few percent of the signal yield. It is not included as a separate contribution in the fit, and a systematic uncertainty is assigned. The $B_s^0 \rightarrow D_s^- \pi^+$ background to $B^0 \rightarrow D^- K^+$ is suppressed by the aforementioned D child $DLL_{K\pi}$ cuts, and additionally by the bachelor PID cut, down to the same level as in the $B^0 \rightarrow D^- \pi^+$ sample.

The opposite, $B^0 \rightarrow D^- \pi^+$ background under $B_s^0 \rightarrow D_s^- \pi^+$, is more dangerous. A hard $DLL_{K\pi}$ cut at 5 is applied on the opposite-sign kaon. The remaining background shape is estimated using the same data-drive technique as for the bachelor $DLL_{K\pi}$ cut separating π and K bachelor modes, *cf.* section 3.3.2, and included in the mass fit, see table 3.7.

3.5.2 Λ_c contamination to D^- and D_s^-

The situation is very similar to the above. Different approaches are used for D^- and D_s^- :

D^-

Simulated events are used to evaluate the relative reconstruction efficiency in the signal window, with the result that this efficiency is a factor 30 and 50 smaller than for respectively $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$. Taking into account the branching and production ratios and the bachelor $DLL_{K\pi}$ cut for the $B^0 \rightarrow D^- K^+$, a 2 % contamination under the $B^0 \rightarrow D^- \pi^+$ mass peak and a 1 % contamination under the $B^0 \rightarrow D^- K^+$ mass peak remains. From here on, the

treatment is the same as for the D_s^- crossfeed: this background is not included in the fit, and a systematic uncertainty is assigned.

D_s^-

Due to the smaller $B_s^0 \rightarrow D_s^- \pi^+$ yield, the $\Lambda_b \rightarrow \Lambda_c \pi^-$ background is no longer negligible. Any DLL_{Kp} cut that removes a significant amount of this background, also impacts signal, however — this is because the RICH proton-kaon separation is extremely poor below 10 GeV, and a significant fraction of the signal is situated in this region. No DLL_{Kp} cut is applied, and the $\Lambda_b \rightarrow \Lambda_c \pi^-$ background is accounted for in the fit, in the same way as the partially reconstructed backgrounds.

3.6 Invariant mass fits

All components are now in place to extract the raw yields of the $B^0 \rightarrow D^- \pi^+$, $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ decay modes. Each yield is extracted in a separate fit.

3.6.1 Description of the fit models

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

For the $B^0 \rightarrow D^- \pi^+$ decay channel the fit procedure is straightforward. Most of the ingredients of the fit model have been described before: the included partially reconstructed backgrounds are $B^0 \rightarrow D^- \rho^+$ and $B^0 \rightarrow D^{*-} \pi^+$, and an exponential shape is used to model the remaining combinatorial and prompt charm background contribution. The signal CB tail parameters are fixed from simulation, *cf.* section 3.3.1. The widths are left free in this fit. The other fitted parameters are the common mean and relative fraction of the two signal CB functions, the (small negative) exponential shape parameter (a in e^{am}) and the sizes of all included contributions (signal, combinatorial background and the partially reconstructed backgrounds $B^0 \rightarrow D^- \rho^+$ and $B^0 \rightarrow D^{*-} \pi^+$).

The final fit is shown in fig. 3.16; the final values for the fitted parameters are in table 3.6.

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

The most important difference from the $B^0 \rightarrow D^- \pi^+$ fit is the momentum-reweighting procedure applied to the $B^0 \rightarrow D^- \pi^+$ crossfeed and the partially reconstructed backgrounds involving a bachelor pion. The combinatorial background is further reduced by the $\text{DLL}_{K\pi}$ cut, which makes it possible to model it using a uniform distribution. The signal shape tails are again fixed from simulation; the widths are now fixed from simulation, corrected using the difference between simulation and data observed in the $B^0 \rightarrow D^- \pi^+$ fit. The fitted parameters in this case are then: the common mean and relative fraction of the two signal CB functions and the sizes of all included contributions (signal, combinatorial background, misidentified $B^0 \rightarrow D^- \pi^+$, and the partially reconstructed $B^0 \rightarrow D^{*-} K^+$ and $B^0 \rightarrow D^- \rho^+$). The final fit is shown in fig. 3.17; the final values for the fitted parameters are in table 3.7.

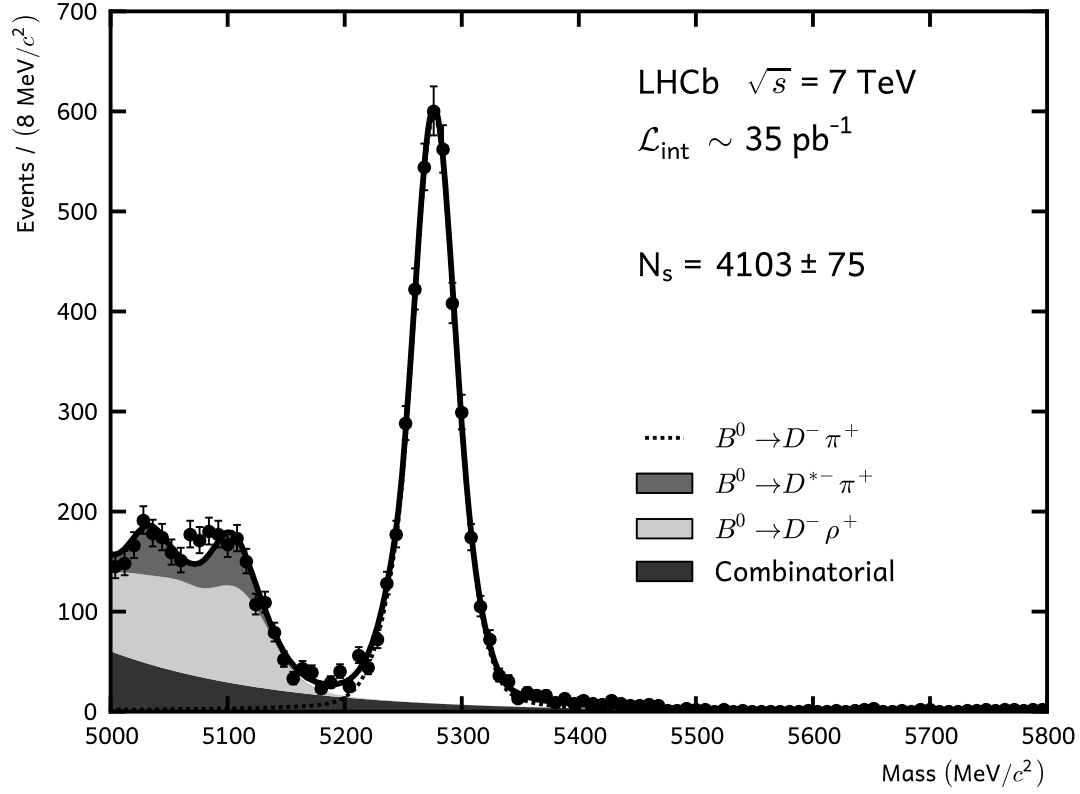


Figure 3.16: Invariant mass fit to the $B^0 \rightarrow D^- \pi^+$ sample

Table 3.6: $B^0 \rightarrow D^- \pi^+$ fitted parameters. The N_i parameters are the sizes of different contributions. The parameter “frac” is the relative fraction of the signal CB functions, while “mean” is their common mean. a_1 characterises the combinatorial background exponential

Parameter	Fitted value	Error
$N_{B^0 \rightarrow D^- \pi^+}$	4103	75
frac	0.58	0.06
mean	5276.3	0.4
σ_1	15.1	1.0
σ_2	27.1	1.2
$N_{\text{Combinatorial}}$	1037	148
a_1	$-7.2 \cdot 10^{-3}$	$0.5 \cdot 10^{-3}$
$N_{B^0 \rightarrow D^- \rho^+}$	1631	198
$N_{B^0 \rightarrow D^{*-} \pi^+}$	535	137

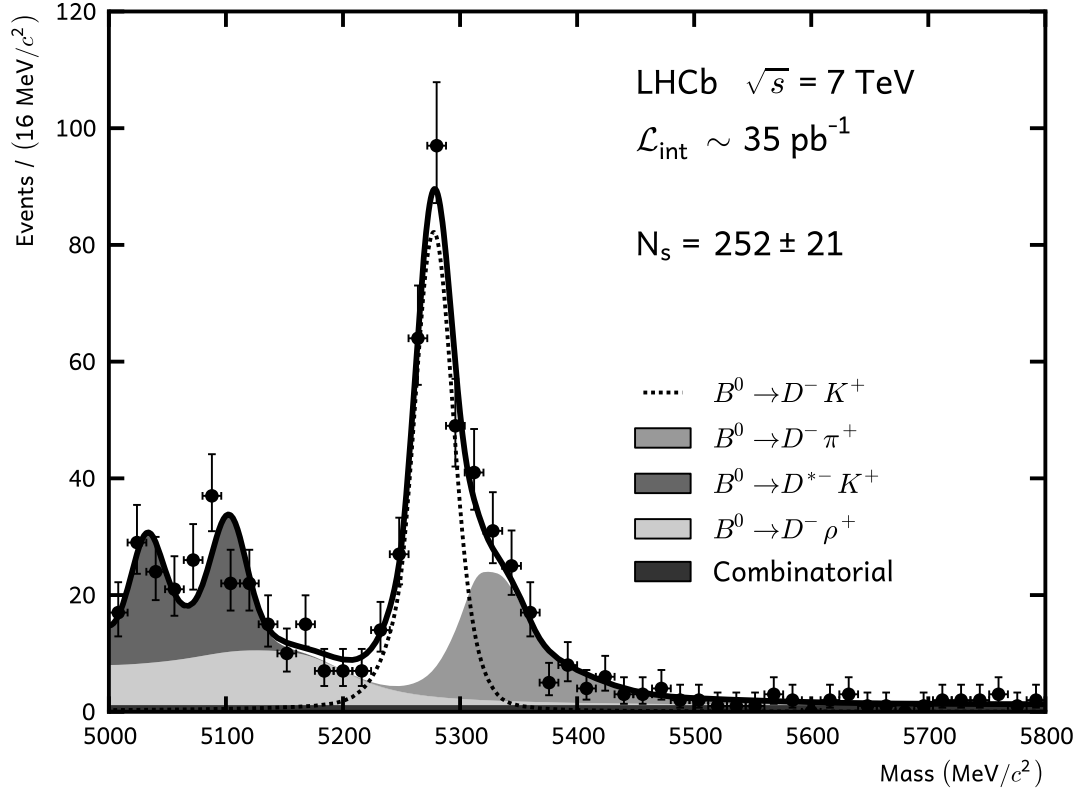


Figure 3.17: Invariant mass fit to the $B^0 \rightarrow D^- K^+$ sample

Table 3.7: $B^0 \rightarrow D^- K^+$ fitted parameters. The names are as in table 3.6

Parameter	Fitted value	Error
$N_{B^0 \rightarrow D^- K^+}$	252	21
frac	0.38	0.17
mean	5277.5	1.8
$N_{\text{Combinatorial}}$	58	14
$N_{B^0 \rightarrow D^- \pi^+}$	131	19
$N_{B^0 \rightarrow D^- \rho^+}$	125	24
$N_{B^0 \rightarrow D^{*-} K^+}$	123	19

$$B_s^0 \rightarrow D_s^- \pi^+$$

As mentioned earlier, the analysis on data was performed in steps: first, the offline selection was optimised on a small sub sample, next the full statistics fits in the B^0 signal region were performed. The B_s^0 signal region was only looked into after studies on “cocktail” simulation samples, in order to develop an approximate fit model. This allowed to see at an early stage that separating the signal from the $B^0 \rightarrow D^- \pi^+$ background, using the invariant mass fit with “template” distributions and floating the sizes of all contributions, was not possible. The $B^0 \rightarrow D^- \pi^+$ shape is mostly situated in the same region as the signal, and correlates with the $\Lambda_b \rightarrow \Lambda_c \pi^-$ contribution due to the right tail coming from the misidentification of the D pion as a kaon. Attempts to extract the $B^0 \rightarrow D^- \pi^+$ contribution from the fit showed the limitations of the background description and the fitting approach; the correlations between the signal, $B^0 \rightarrow D^- \pi^+$ and $\Lambda_b \rightarrow \Lambda_c \pi^-$ numbers — which are to be expected from the shapes — make the extracted signal yield sensitive to small effects in the template shapes and hence less reliable. The fit behaviour was studied by fixing the size of the $B^0 \rightarrow D^- \pi^+$ background to values around the true value — this showed the correlations between the $B^0 \rightarrow D^- \pi^+$, $\Lambda_b \rightarrow \Lambda_c \pi^-$ and signal $B_s^0 \rightarrow D_s^- \pi^+$ sizes. Furthermore, the absence of a minimum in the best fit negative log-likelihood for different values of the $B^0 \rightarrow D^- \pi^+$ size showed that fitting for this contribution is not possible in a reliable way. This was resolved by applying a Gaussian constraint to the number of $B^0 \rightarrow D^- \pi^+$ events, calculated using the reweighting procedure described in section 3.3.2, but now for the PID cut applied to the D or D_s^+ child particle that is a π in the B^0 case and a K in the B_s^0 case.

From these studies on cocktail samples it was also possible to get a first idea of the partially reconstructed backgrounds in the mass range below the B_s^0 mass. Two contributions that were expected are those coming from $B_s^0 \rightarrow D_s^{*-} \pi^+$ and $B_s^0 \rightarrow D_s^- \rho^+$ decays. The difficulty in fitting for their sizes comes from the fact that these are mostly determined by the shape and size of the backgrounds observed on the left of the signal region, where they look very similar, but the $B_s^0 \rightarrow D_s^- \rho^+$ enters the signal region much more than the $B_s^0 \rightarrow D_s^{*-} \pi^+$. This is an important source of uncertainty in the fit model, given the fact that the good description of this part of the mass window suffers from sizeable uncertainties. The current best measurements of the ratio of $B_s^0 \rightarrow D_s^- \rho^+$ and $B_s^0 \rightarrow D_s^{*-} \pi^+$ branching fractions have large uncertainties. In addition there is an observed difference in shape between the data and simulated event cocktail samples, a double peak structure which is supposed to come from a $B_s^0 \rightarrow D_s^{(1)-} \pi^+$ decay where the $D_s^{(1)-}$ in turn decays into $D_s^\pm \pi^0$.

The backgrounds very close to the signal mentioned above require the misidentification of a π or a p from respectively a D or Λ_c decay as a K . Using $DLL_{K\pi}$ cuts, the other (smaller) $B^0 \rightarrow D^- \pi^+$ -like backgrounds, $B^0 \rightarrow D^- \rho^+$ and $B^0 \rightarrow D^{*-} \pi^+$ can be suppressed to a level where including them in the fit has no influence on the signal yield anymore: the shapes are very similar to the corresponding B_s^0 decays, and these small contributions can easily be absorbed by the templates for those backgrounds and the combinatorial background.

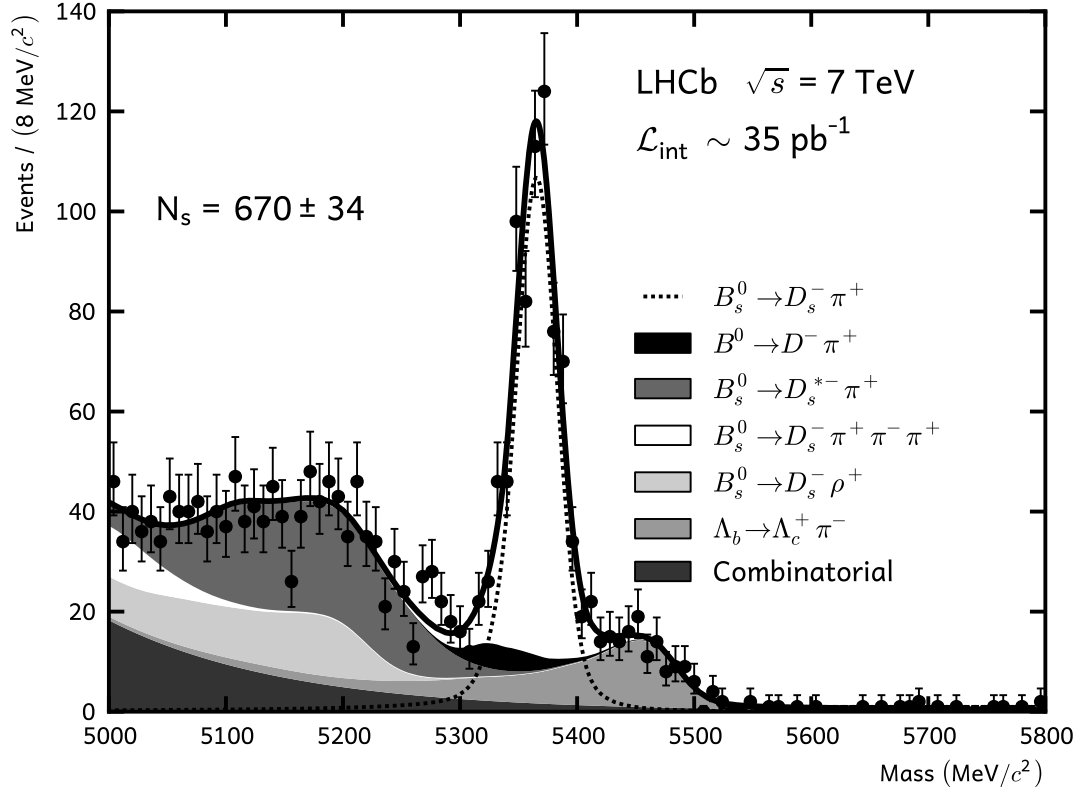


Figure 3.18: Invariant mass fit to the $B_s^0 \rightarrow D_s^- \pi^+$ sample

Table 3.8: $B_s^0 \rightarrow D_s^- \pi^+$ fitted parameters. The names are as in table 3.6; $frac_{\text{part}}$ is the ratio of $B_s^0 \rightarrow D_s^- \rho^+$ and the $B_s^0 \rightarrow D_s^{*-} \pi^+$ contribution

Parameter	Fitted value	Error
$N_{B_s^0 \rightarrow D_s^- \pi^+}$	670	34
$frac$	0.50	0.14
$mean$	5365.3	1.0
$N_{\text{Combinatorial}}$	345	240
a_1	$-6.6 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$
$N_{B^0 \rightarrow D^- \pi^+}$	48	18
$N_{\Lambda_b \rightarrow \Lambda_c^+ \pi^-}$	287	36
$N_{B_s^0 \rightarrow D_s^- \pi^+ \pi^- \pi^-}$	79	79
N_{part}	668	121
$frac_{\text{part}}$	0.34	0.06

3.6.2 Validation of the fit models

Toy fits

The main method used to study the fit stability and sensitivity to some of our assumptions in constructing the fit model, are toy fits: a large number of samples were generated from a “generation model” and fitted with a “fit model”. The accuracy and stability of the fit model can be studied by using the fit model as a generator model: the fluctuations in the samples are then just those due to the limited number of events. The distribution of the fitted parameter values and their errors is studied. For a sufficiently large number of toy fits, the value distribution is expected to be Gaussian, centred around the “true” value used in the generator model, with a width corresponding to the uncertainties calculated by the fit algorithms — more precisely: the so-called pull distribution, computed by taking for each toy fit the difference between the fitted value and the true value, divided by the uncertainty on the fitted value, should be a standard Gaussian (centred around zero, with width one).

Fit model systematic uncertainty

The concept outlined in the previous section can be extended to evaluate the systematic uncertainties originating from approximations in the fit models and uncertainties on input parameters. The response of the pull distribution to changes in the generation model straightforwardly quantifies those systematics. A first example is found in the $B^0 \rightarrow D^- \pi^+$ fit: the $B_s^0 \rightarrow D_s^- \pi^+$ and $\Lambda_b \rightarrow \Lambda_c \pi^-$ contributions, expected at a level of a few percent of the signal yield, were not fitted for. To account for this, those backgrounds were added in the generation model for the toy samples. The variation of the signal yield pull parameters, μ and σ of a Gaussian function fitted to the pull distribution of 2000 toy fits for each point, are shown in fig. 3.19.

The trends in the mean signal yield can be understood qualitatively as follows: the $\Lambda_b \rightarrow \Lambda_c \pi^-$ background is situated just on the right side of the signal peak (with a lower-end tail due to misidentification), while the $B_s^0 \rightarrow D_s^- \pi^+$ shape peaks left under the signal, with a similar tail: there is no other contribution to fit to the $\Lambda_b \rightarrow \Lambda_c \pi^-$ than the signal, while the $B_s^0 \rightarrow D_s^- \pi^+$ background can be absorbed by the partially reconstructed backgrounds.

Expected background sizes of 2 for $\Lambda_b \rightarrow \Lambda_c \pi^-$ and 1 % for $B_s^0 \rightarrow D_s^- \pi^+$ give an uncertainty on the signal of around 0.7σ . Taking into account the uncertainties on these numbers, a 1.5 % systematic uncertainty on the signal yield is found.

Additionally, the dependence of the signal yield on the combinatorial background parametrisation was studied, by generating this background contribution according to a linear distribution. The signal yield value and pull distributions are shown in fig. 3.20.

The $B_s^0 \rightarrow D_s^- \pi^+$ fit model systematic uncertainty was studied similarly to the $B^0 \rightarrow D^- \pi^+$ case. The main source of systematic uncertainty in this case is due to the fixed values for the σ parameters of the signal shape CB functions. These were varied over a large range in the toy sample generating model. The result is shown in fig. 3.21 — note that the range for σ_2 is much larger. This is done to take into account the larger uncertainty on the value extract from the signal shape fits. The negative trend with increasing σ can be understood from the wider

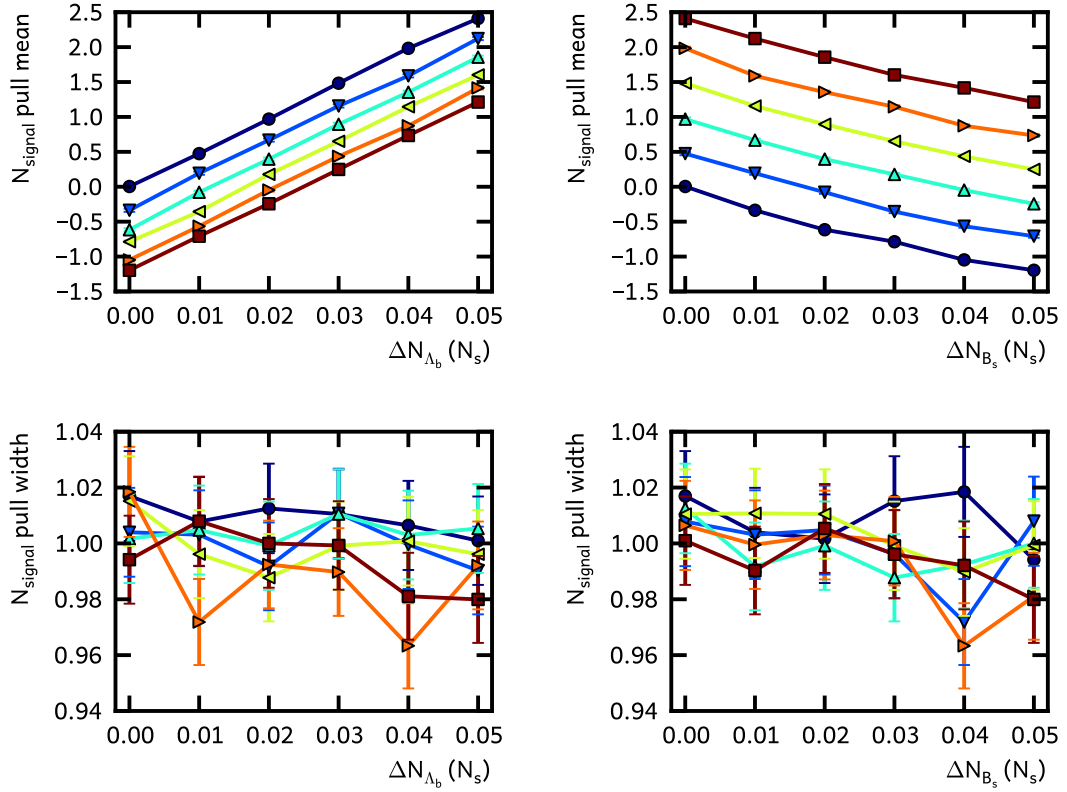


Figure 3.19: Change of $B^0 \rightarrow D^- \pi^+$ signal yield pull μ and σ , when changing the generated amount of $B_s^0 \rightarrow D_s^- \pi^+$ and $\Lambda_b \rightarrow \Lambda_c \pi^-$ in the toy samples

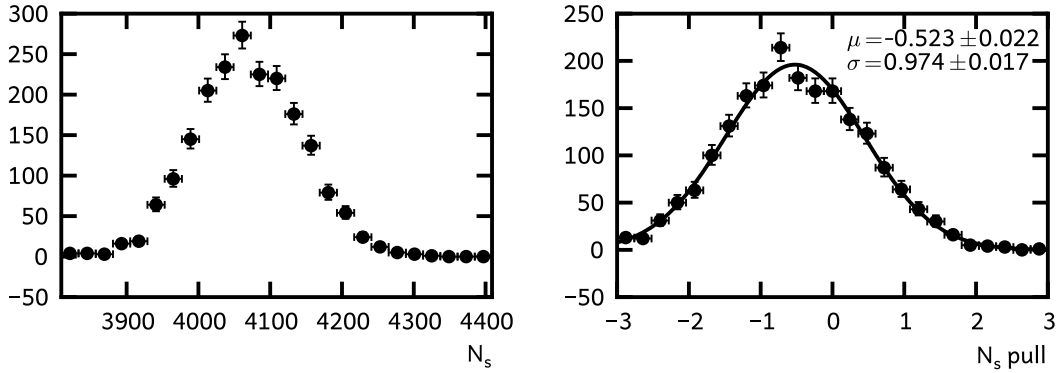


Figure 3.20: $B^0 \rightarrow D^- \pi^+$ signal yield value and pull distribution when generating with a linear instead of an exponential combinatorial background shape

generated signal being absorbed by the backgrounds, such that the fitted signal has a smaller area.

The combinatorial background is again generated from a linear shape — which has only a very small effect on the signal yield. A combined fit model systematic uncertainty of 2 % is found.

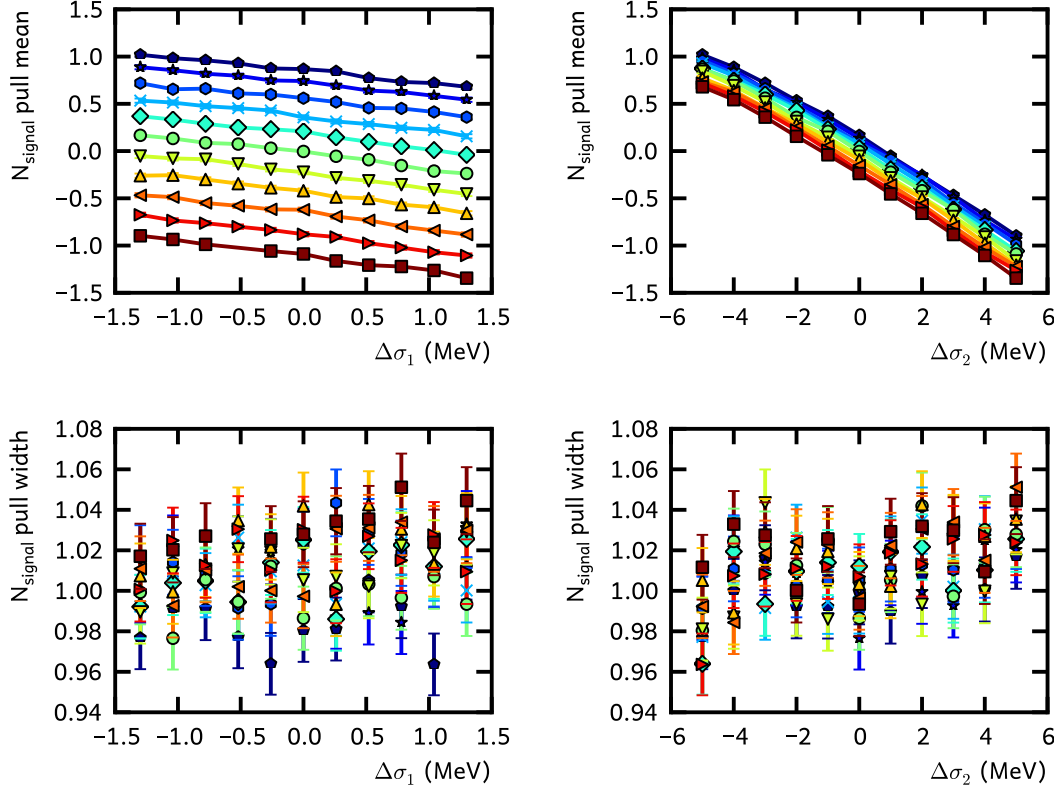


Figure 3.21: Change of the $B_s^0 \rightarrow D_s^- \pi^+$ signal yield pull μ and σ when changing the σ parameters of the signal CB functions

For the $B^0 \rightarrow D^- K^+$ fit, the same procedure as for $B^0 \rightarrow D^- \pi^+$ was applied to evaluate the systematic uncertainty associated to not including the $\Lambda_b \rightarrow \Lambda_c \pi^-$ and $B_s^0 \rightarrow D_s^- \pi^+$ backgrounds, see fig. 3.24. The trends are again easily understood qualitatively: the $\Lambda_b \rightarrow \Lambda_c \pi^-$ background is absorbed by the $B^0 \rightarrow D^- \pi^+$ contribution, which gives a more shallow dependence — the $B_s^0 \rightarrow D_s^- \pi^+$ background couples with the partially reconstructed backgrounds left of the signal. A systematic uncertainty of 1.5 % was found. For this fit, the combinatorial background was fixed to a uniform distribution, as described before. An exponential distribution was inserted in the generator model to study the effect of this approximation. The slope was varied up to the value found in the $B^0 \rightarrow D^- \pi^+$ fit. A systematic uncertainty of 0.5 % was found fig. 3.23

The systematic uncertainty associated to the fixed CB σ parameters was studied in the same

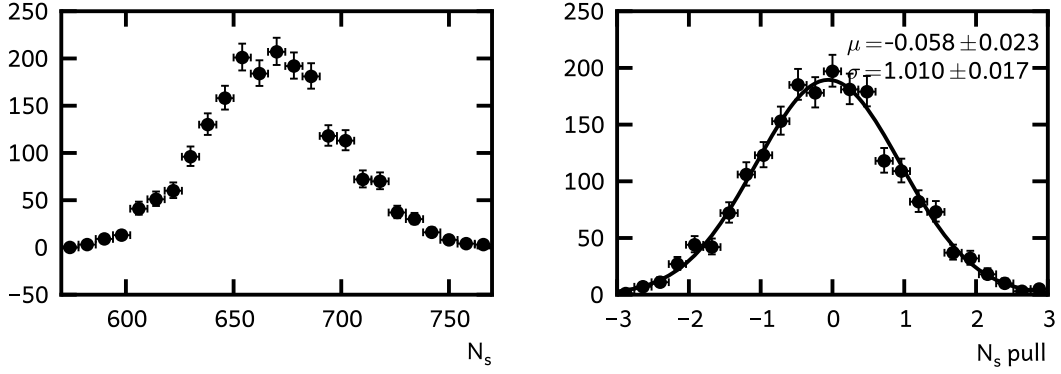


Figure 3.22: $B^0 \rightarrow D_s^- \pi^+$ signal yield value and pull distribution when generating with a linear instead of an exponential combinatorial background shape

way as for $B_s^0 \rightarrow D_s^- \pi^+$. A final uncertainty of 2 % was assigned.

All three fit model uncertainties were rounded to 2 % for computational simplicity. One could evaluate one overall fit model uncertainty by applying Gaussian constraints on the fixed parameters (CB σ 's, combinatorial background for $B^0 \rightarrow D^- K^+$ and $\Lambda_b \rightarrow \Lambda_c \pi^-$ and $B_s^0 \rightarrow D_s^- \pi^+$ backgrounds to $B^0 \rightarrow D^- \pi^+$ and $B^0 \rightarrow D^- K^+$) — taking into account all correlations between these effects automatically. This could significantly reduce the computing time, especially in presence of a larger number of source of uncertainty. The current procedure provides more information and has the advantage of showing and quantifying the individual effects.

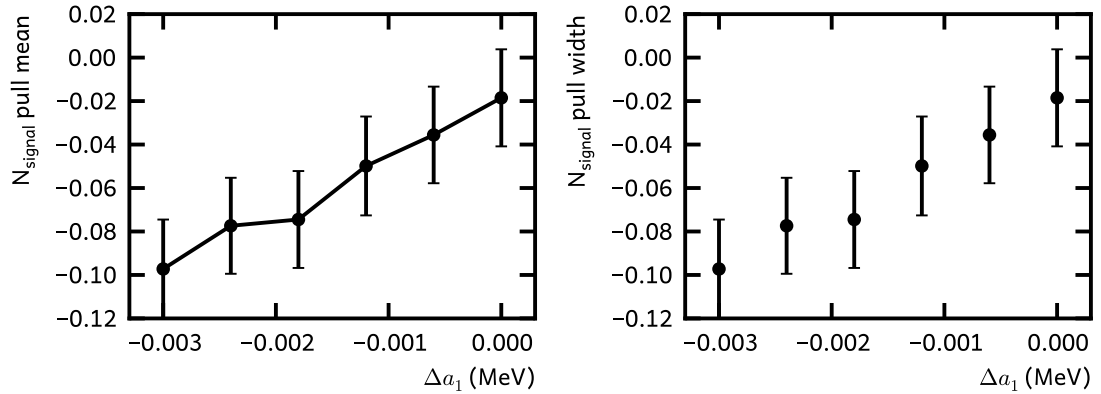


Figure 3.23: Change of $B^0 \rightarrow D^- K^+$ signal yield pull μ and σ , when changing the generated combinatorial background shape

As a cross-check of the analysis method the sensitivity of the results on the selection value for the multivariate selection was studied. The change in the ratios was found to be much negligible compared to the statistical uncertainty.

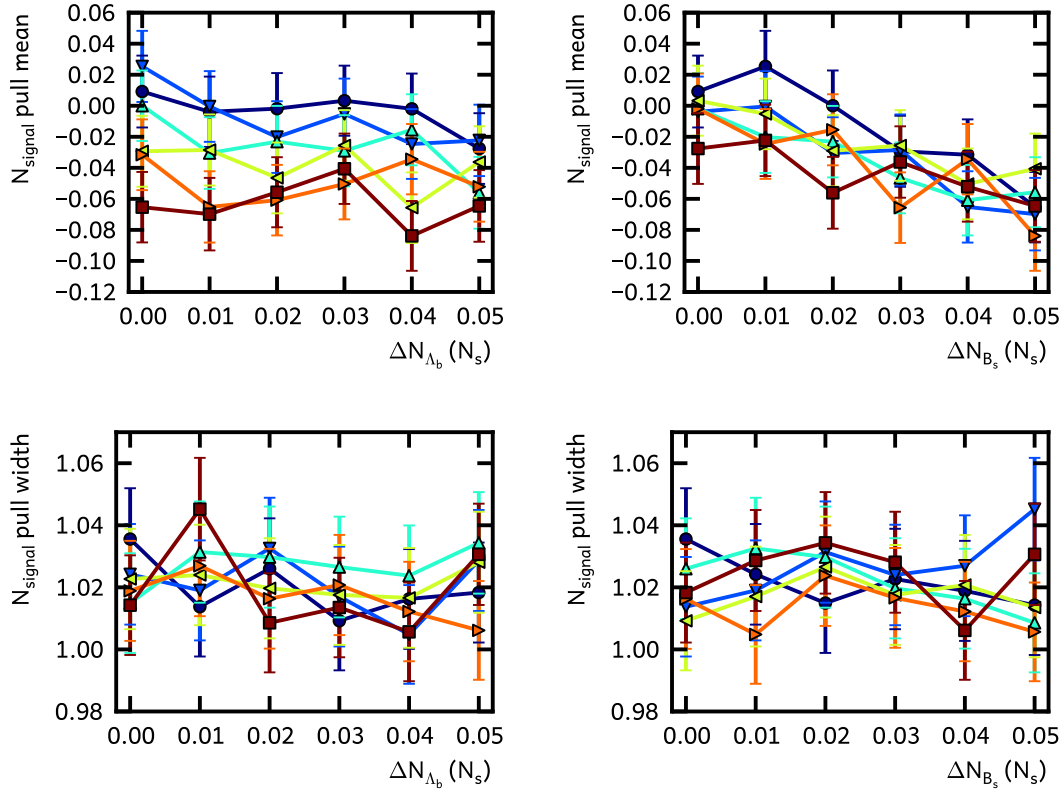


Figure 3.24: Change of the $B^0 \rightarrow D^- K^+$ signal yield pull μ and σ , when changing the generated amount of $B_s^0 \rightarrow D_s^- \pi^+$ and $\Lambda_b \rightarrow \Lambda_c \pi^-$ background

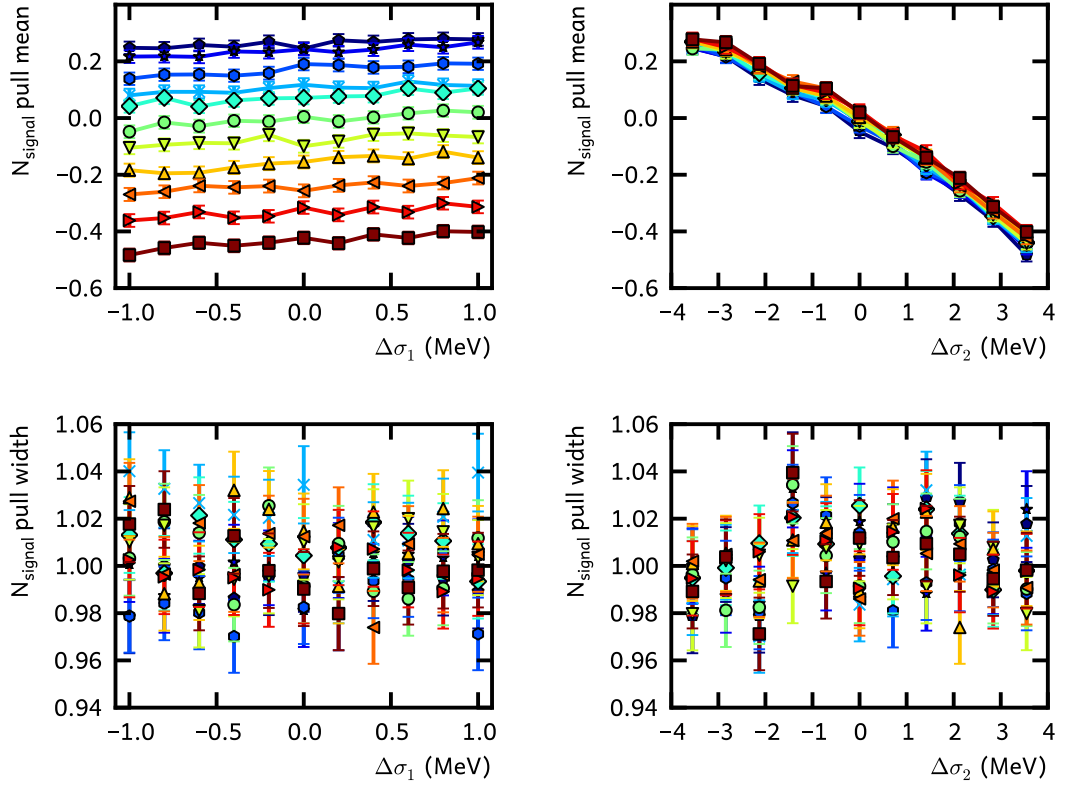


Figure 3.25: Change of the $B^0 \rightarrow D^- K^+$ signal yield μ and σ when changing the σ parameters of the signal CB function

Table 3.9: Correction factors for the relative branching ratio measurements.

Source	$\varepsilon(D\pi)/\varepsilon(DK)$
Bachelor reconstruction efficiency ratio	1.025 ± 0.003
PID cuts	1.183 ± 0.006
$p < 100$ GeV efficiency for $B^0 \rightarrow D^- K^+$	1.007 ± 0.002
Trigger	1.00 ± 0.02
Overall correction factor	1.221 ± 0.021

3.7 Determination of the $B^0 \rightarrow D^- K^+ / B^0 \rightarrow D^- \pi^+$ relative branching ratio

For the interpretation of the fitted $B^0 \rightarrow D^- K^+$ and $B^0 \rightarrow D^- \pi^+$ signal yields in terms of the relative branching ratio of these decay channels, the slight biases in the selection procedure must be corrected for, and systematic uncertainties must be assigned to account for uncertainties on other input or inherent to the applied procedure.

The correction factors are due to the different reconstruction and selection efficiencies, as discussed before. When the efficiencies measured on simulated events are equal within error and there is no physical reason for a difference, no correction factor is applied. The correction factors are summarised in table 3.9.

The number of sources of systematic uncertainties is limited because the $B^0 \rightarrow D^- K^+$ and $B^0 \rightarrow D^- \pi^+$ are measured within the same B flavour. The major contribution is due to the PID efficiencies used to reweight the $B^0 \rightarrow D^- \pi^+$ crossfeed mass shape. This enters the error budget in two ways: firstly as an uncertainty on the correction factors quoted in the previous section, and secondly as part of the “fit model” systematic uncertainty, since the fit shape for the combinatorial backgrounds relies on the accuracy of the momentum reweighting procedure.

Additionally to the fit model uncertainties discussed before, some large correlations were present in some fits. The absence of strong nonlinearities was checked, as these could, if present, result in a wrong estimate for the statistical uncertainty.

The overall systematic uncertainty budget is listed in table 3.10. Note that there is a 2 % uncertainty associated to the L0 trigger efficiency, which is 2 % higher in simulated events for $B^0 \rightarrow D^- K^+$ than for $B^0 \rightarrow D^- \pi^+$. Although the overall trigger efficiency correction factor is measured to be 1.00 ± 0.02 from simulated events, the kaon modes show a higher L0 efficiency. Considering a known problem in simulating the L0 calorimeter energy response for 2010 data, an additional systematic uncertainty, conservatively taken as the total discrepancy size, is assigned.

When all these correction factors are applied to the raw yield ratio, the measured relative branching ratio becomes

$$\frac{B(B^0 \rightarrow D^- K^+)}{B(B^0 \rightarrow D^- \pi^+)} = 0.0750 \pm 0.0066 \pm 0.0035, \quad (3.11)$$

Table 3.10: Systematics for the relative branching ratio measurements.

Source	Size for B^0
PID calibration	2.5%
Fit model	2.8%
L0 trigger efficiency	2.0%
Correction factors	1.7%
Total	4.6%

Table 3.11: Correction factors for the f_s/f_d measurements.

Source	$\varepsilon(D K)/\varepsilon(D_s \pi)$	$\varepsilon(D \pi)/\varepsilon(D_s \pi)$
$K^\mp/\pi^\mp$ reco.	1.014 ± 0.004	1.039 ± 0.003
PID cuts	0.959 ± 0.006	1.135 ± 0.005
$P < 100\text{GeV}$ cut	0.993 ± 0.001	N/A
Trigger eff. corr.	0.95 ± 0.020	0.95 ± 0.020
Overall correction	0.917 ± 0.020	1.120 ± 0.025

where the first error is statistical and the second systematic.

The PDG world average value for the $B^0 \rightarrow D^- \pi^+$ branching ratio results in a value of

$$B(B^0 \rightarrow D^- K^+) = (2.01 \pm 0.18 \pm 0.14) \times 10^{-4}, \quad (3.12)$$

where the 4.9 % uncertainty on the $B^0 \rightarrow D^- \pi^+$ branching ratio is included in the second, systematic, uncertainty.

3.8 Interpretation in terms of f_s/f_d

For the f_s/f_d measurement, several systematic uncertainties enter with the use of different B flavour decay modes: the trigger efficiency difference between B_s^0 and B^0 modes, as well as the knowledge of the branching ratios $D_s^- \rightarrow K^- K^+ \pi^-$, as opposed to $D_s^- \rightarrow \pi^- K^+ \pi^-$.

The correction factors are treated in the same way as for the $B^0 \rightarrow D^- K^+$ branching ratio, and summarised in table 3.11

Additionally to the systematic uncertainties considered for the relative branching ratio measurement, $B_s^0 \rightarrow D_s^- \pi^+$ fit model uncertainties and systematic effects in the modelling of the trigger response using simulated events arise. The former are studied similarly, using large numbers of toy fits, considering the sensitivity to the σ -parameters of the signal CB-functions. An uncertainty of 2.0 % was found. The overall systematic budget is listed in table 3.12.

Some differences between the two methods for the determination of f_s/f_d exist. The theo-

Table 3.12: Systematics for the f_s/f_d measurements.

Source	For $D_s^\pm \pi^\mp$ and $D^\pm K^\mp$	For $D_s^\pm \pi^\mp$ and $D^\pm \pi^\mp$
PID calibration	1.0%	2.5%
B^0 fit model	2%	2%
B_s^0 fit model	2%	2%
L0 Trigger efficiency	2%	2%
$D_s^\pm \rightarrow KK\pi$ B.R	4.9%	4.9%
$D^\pm \rightarrow K\pi\pi$ B.R	2.2%	2.2%
$\frac{\tau_{B_s^0}}{\tau_{B^0}}$	1.5%	1.5%
Correction factors	2.2%	2.2%
Total	7.0%	7.4%

retically cleaner extraction using $B^0 \rightarrow D^- K^+$ and $B_s^0 \rightarrow D_s^- \pi^+$ follows

$$\frac{f_s}{f_d} = 0.0743 \times \frac{\tau_{B^0}}{\tau_{B_s^0}} \times \left[\frac{1}{\mathcal{N}_a \mathcal{N}_F} \frac{\varepsilon_{DK}}{\varepsilon_{D_s \pi}} \frac{N_{D_s \pi}}{N_{DK}} \right]. \quad (3.13)$$

Using the world-average lifetime ratio for B_s^0 to B^0 [41]

$$\frac{\tau_{B^0}}{\tau_{B_s^0}} = 1.028 \pm 0.016,$$

and the combined factor accounting for non-factorisable corrections and form factors

$$\mathcal{N}_a \mathcal{N}_F = 1.24 \pm 0.084$$

and, finally, making use of $B(D^+ \rightarrow K^- \pi^+ \pi^+) = (9.14 \pm 0.20)\%$ and $B(D_s^+ \rightarrow K^- K^+ \pi^+) = (5.50 \pm 0.27)\%$, the value of f_s/f_d is found to be

$$\frac{f_s}{f_d} = 0.250 \pm 0.024^{\text{stat}} \pm 0.017^{\text{syst}} \pm 0.017^{\text{theor}}. \quad (3.14)$$

The theoretical uncertainty is dominated by form factor ratio.

The $B^0 \rightarrow D^- \pi^+$ and the $B_s^0 \rightarrow D_s^- \pi^+$ yields allow to extract f_s/f_d with a smaller statistical, but larger theoretical uncertainty, due to W exchange diagrams

$$\mathcal{N}_E = 0.966 \pm 0.075.$$

With the formula

$$\frac{f_s}{f_d} = 0.982 \times \frac{\tau_{B^0}}{\tau_{B_s^0}} \times \left[\frac{1}{\tilde{\mathcal{N}}_a \mathcal{N}_F \mathcal{N}_E} \frac{\varepsilon_{D\pi}}{\varepsilon_{D_s \pi}} \frac{N_{D_s \pi}}{N_{D\pi}} \right], \quad (3.15)$$

the value for f_s/f_d is found to be

$$\frac{f_s}{f_d} = 0.256 \pm 0.014^{\text{stat}} \pm 0.019^{\text{syst}} \pm 0.026^{\text{theor}}. \quad (3.16)$$

Conclusion

With 35 pb^{-1} of data collected using the LHCb detector during the 2010 LHC running at a centre-of-mass energy of 7 TeV, the branching ratio of the Cabbibo-suppressed charmed B^0 decay mode $B^0 \rightarrow D^- K^+$ was measured as

$$B(B^0 \rightarrow D^- K^+) = (2.01 \pm 0.18 \pm 0.14) \times 10^{-4},$$

where the first error is statistical, and the second is systematic. Additionally, two measurements of the f_s/f_d production fraction were performed under different theoretical assumptions, from the relative yields of the B^0 modes $B^0 \rightarrow D^- K^+$ and $B^0 \rightarrow D^- \pi^+$ with respect to $B_s^0 \rightarrow D_s^- \pi^+$. The measured values of f_s/f_d are in good agreement with each other, being

$$\frac{f_s}{f_d} = 0.250 \pm 0.024^{\text{stat}} \pm 0.017^{\text{syst}} \pm 0.017^{\text{theor}},$$

and

$$\frac{f_s}{f_d} = 0.256 \pm 0.014^{\text{stat}} \pm 0.019^{\text{syst}} \pm 0.026^{\text{theor}}.$$

The two values for f_s/f_d can be combined into a single value, by taking all correlated uncertainties into account⁶. The following averaged value for f_s/f_d is obtained:

$$\frac{f_s}{f_d} = 0.253 \pm 0.017^{\text{stat}} \pm 0.017^{\text{syst}} \pm 0.020^{\text{theor}}, \quad (3.17)$$

where all contributions to the uncertainty are assumed to be Gaussian distributed.

⁶The correlation factors found are: 49% for statistics, 78% for the systematics and 66% for the theoretical error

Bibliography

- [1] R. Fleischer, N. Serra, and N. Tuning, *Tests of Factorization and $SU(3)$ Relations in B Decays into Heavy-Light Final States*, hep-ph/1012.2784.
- [2] A. A. Alves Jr. et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [3] The LHCb Collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003.
- [4] *LHCb: letter of intent*, Tech. Rep. CERN-LHCC-95-5. LHCC-I-8. CERN-LHCC-1995-005, CERN, Geneva, 1995.
- [5] S. Amato, A. Satta, B. Souza de Paula, and L. De Paula, *Hlt1 Muon Alley Description*, Tech. Rep. LHCb-2008-058. CERN-LHCb-2008-058, CERN, Geneva, Nov, 2008.
- [6] A. Pérez-Calero and H. Ruiz, *The muon+track alley of the LHCb High Level Trigger*, Tech. Rep. LHCb-2008-075. CERN-LHCb-2008-075, CERN, Geneva, Oct, 2008.
- [7] J. A. Hernando Morata, D. Martinez Santos, X. Cid Vidal, G. Guerrer, A. Gomes Dos Santos Neto, and J. Albrecht, *The Hadron Alley of the LHCb High Level Trigger*, Tech. Rep. LHCb-2009-034. CERN-LHCb-2009-034, CERN, Geneva, Apr, 2010.
- [8] K. Senderowska, M. Witek, and A. Zuranski, *HLT1 Electromagnetic Alley*, Tech. Rep. LHCb-PUB-2009-001. CERN-LHCb-PUB-2009-001, CERN, Geneva, Mar, 2010.
- [9] V. V. Gligorov, *A single track HLT1 trigger*, Tech. Rep. LHCb-PUB-2011-003, CERN, Geneva, Jan, 2011.
- [10] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214–215.
- [11] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Part I*, .
- [12] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Part II*, .
- [13] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (Aug, 1964) 321–323.
- [14] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (Oct, 1964) 508–509.
- [15] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*. Westview Press, 1995.
- [16] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

- [17] M. Kobayashi and T. Maskawa, *CP Violation in the renormalizable theory of weak interaction*, *Prog. Theor. Phys.* **49** (1973) 652.
- [18] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, *Phys. Rev. Lett.* **51** (Nov, 1983) 1945–1947.
- [19] A. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, *Soviet Physics Journal of Experimental and Theoretical Physics (JETP)* **5** (1967) 24.
- [20] R. Fleischer, *New strategies to obtain insights into CP violation through $B_s \rightarrow D_s^\pm K^\mp$, $D_s^{*\pm} K^\mp$, ... and $B_d \rightarrow D^\pm \pi^\mp$, $D^{*\pm} \pi^\mp$, ... decays*, *Nucl. Phys.* **B657** (2003) 459.
- [21] S. Cohen, M. Merk, and E. Rodrigues, *$\gamma + \varphi_s$ sensitivity studies from combined $B_s^0 \rightarrow D_s^\mp \pi^\pm$ and $B_s^0 \rightarrow D_s^\mp K^\pm$ samples at LHCb*, Tech. Rep. CERN-LHCb-2007-041, 2007.
- [22] V. V. Gligorov, *Measurement of the CKM angle γ and B meson lifetimes at the LHCb detector*.
PhD thesis, 2008.
CERN-THESIS-2008-44.
- [23] N. Uraltsev, *Heavy-quark expansion in beauty and its decays*, hep-ph/9804275.
- [24] M. Beneke, G. Buchalla, M. Neubert, and C. Sachrajda, *QCD factorization for exclusive, non-leptonic B meson decays: General arguments and the case of heavy-light final states*, *Nucl. Phys.* **B591** (2000) 313–418, [hep-ph/0006124].
- [25] T. Mannel, *Effective theory for heavy quarks*, in *Perturbative and Nonperturbative Aspects of Quantum Field Theory* (H. Latal and W. Schweiger, eds.), vol. 479 of *Lecture Notes in Physics*, pp. 387–428.
Springer Berlin / Heidelberg, 1997.
10.1007/BFb0104296.
- [26] M. Gronau, O. F. Hernandez, D. London, and J. L. Rosner, *Broken SU(3) symmetry in two-body B decays*, *Phys. Rev.* **D52** (1995) 6356–6373, [hep-ph/9504326].
- [27] D. Martínez, J. Hernando, and F. Teubert, *LHCb potential to measure/exclude the branching ratio of the decay $B_s \rightarrow \mu^+ \mu^-$* , Tech. Rep. LHCb-2007-033. CERN-LHCb-2007-033, CERN, Geneva, Apr, 2007.
- [28] **LHCb** Collaboration, R. Aaij et. al., *Search for the rare decays $B_s \rightarrow \mu\mu$ and $B_d \rightarrow \mu\mu$* , *Phys. Lett.* **B699** (2011) 330–340, [hep-ex/1103.2465].
- [29] R. Fleischer, N. Serra, and N. Tuning, *New Strategy for B_s Branching Ratio Measurements and the Search for New Physics in $B_s^0 \rightarrow \mu^+ \mu^-$* , *Phys. Rev.* **D82** (2010) 034038, [hep-ph/1004.3982].
- [30] P. Blasi, P. Colangelo, G. Nardulli, and N. Paver, *Phenomenology of B_s decays*, *Phys. Rev.* **D49** (1994) 238–246, [hep-ph/9307290].
- [31] S. Miglioranza and G. Corti, *Material Interaction Cross Section Studies*, Tech. Rep. LHCb-INT-2011-002. CERN-LHCb-INT-2011-002, CERN, Geneva, Jan, 2011.
- [32] H. Dijkstra, N. Tuning, and N. Brook, *Some Remarks on Systematic Effects of the Trigger and Event Generator Studies*, Tech. Rep. LHCb-2003-157, CERN, Geneva, Dec, 2003.

- [33] M. Williams, V. Gligorov, C. Thomas, H. Dijkstra, J. Nardulli, and P. Spradlin, *The HLT2 Topological Lines*, Tech. Rep. LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan, 2011.
- [34] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, [physics/0703039].
- [35] E. Baracchini and G. Isidori, *Electromagnetic corrections to non-leptonic two-body B and D decays*, *Physics Letters B* **633** (2006), no. 2-3 309 – 313.
- [36] M. J. Oreglia, “A study of the reactions ψ prime $\rightarrow$ gamma gamma ψ , ph.d. thesis, appendix d.” SLAC-R-236, 1980.
- [37] J. Gaiser, “Charmonium spectroscopy from radiative decays of the J/ψ and ψ' , Ph.D. thesis, Appendix F.” SLAC-R-255, 1982.
- [38] T. Skwarnicki, “A study of the radiative cascade transitions between the Υ and Υ' resonances, Ph.D. thesis, Appendix E.” DESY F31-86-02, 1986.
- [39] K. S. Cranmer, *Kernel estimation in high-energy physics*, *Comput. Phys. Commun.* **136** (2001) 198–207, [hep-ex/0011057].
- [40] G. Punzi, *Comments on likelihood fits with variable resolution*, physics/0401045.
- [41] **HFAG** Collaboration, D. Asner *et. al.*, *Averages of b-hadron and c-hadron Properties at the End of 2009*, hep-ex/1010.1589.