University of Zurich
Department of Physics

Measurement of $\sigma(pp \to b\bar{b}X)$ at 8 TeV in the Forward Region of the LHCb Experiment

Master Thesis of Dario Biasini

June 22, 2015

Supervisors:
Prof. Dr. Ueli Straumann
Dr. Christian Elsasser
Dr. Katharina Müller

Abstract

The $b\bar{b}$ production cross-section in proton-proton collisions $\sigma(pp \to b\bar{b}X)$ has been measured in the forward region of LHCb for a center-of-mass energy of 8 TeV. The decay channel $b \to D^0\ \mu^-\overline{\nu}_\mu X$ with $D^0 \to K^-\pi^+$ has been used. The measured cross-section in the forward direction of 77.4 ± 7.1 µb agrees well with the prediction of $87.2^{+55.4}_{-33.7} \pm 11.1$ µb[1]. Extrapolated to the entire phase space, the result amounts to 278.3 ± 25.2 µb. The high statistics in this channel allows for a differential cross-section measurement as a function of rapidity y and transverse momentum p_T of the B meson. The results agree with theory within their uncertainties.

Zusammenfassung

Der $b\bar{b}$ Wirkungsquerschnitt bei Proton-Proton Kollision $\sigma(pp \to b\bar{b}X)$ wurde in der Vorwärtsrichtung von LHCb bei einer Schwerpunktsenergie von 8 TeV gemessen. Der Zerfallskanal $b \to D^0\mu^-\overline{\nu}_\mu X$ mit $D^0 \to K^-\pi^+$ wurde dabei verwendet. Der gemessene Wirkungsquerschnitt in Vorwärtsrichtung 77.4 ± 7.1 µb stimmt gut mit der Vorhersage von $87.2^{+55.4}_{-33.7} \pm 11.1$ µb[2] überein. Auf den gesamten Phasenraum ausgeweitet ergibt sich 278.3 ± 25.2 µb. Die hohe Statistik im Zerfallskanal erlaubt eine differentielle Messung als Funktion von Rapidität y und transversalem Impuls p_T des B-Mesons. Die Resultate stimmen mit der Theorie innerhalb der Unsicherheiten überein.

[1] The first error comes from the parton-parton cross-section, the second from the parton distribution function (Sec. 2).

[2] Der erste Fehler stammt von der Parton-Parton Wirkungsquerschnitt Berechnung, der zweite Fehler von der Parton Verteilungsfunktion (Kap. 2).

Contents

1 Introduction

The goal of physics is nothing more than to explain the fundamental laws of nature. At present, we recognise four fundamental forces that govern all physics processes: electromagnetism, gravity, the weak- and the strong force. At LHC and many other particle physics experiments three of the four forces[1] are examined by colliding high energetic particles, examining the collision artifacts and putting them into context with theoretical predictions.

In this thesis, a measurement of the $b\bar{b}$ production cross-section in proton-proton collisions $\sigma(pp \to b\bar{b}X)$ at a center-of-mass energy of 8 TeV is presented. The measurement is performed on data from the LHCb detector. The decay channel $b \to D^0\mu^-\overline{\nu}_\mu X$ with $D^0 \to K^-\pi^+$ has been used where X is signifies any number of additional final state particles. This channel has a reasonably high branching fraction of about 6.8% [1] with a good signal-to-background ratio. Quantum field theory predicts $\sigma(pp \to b\bar{b}X)$ in next-to-leading order. However, the theoretical uncertainties are rather large. A cross-section measurement allows to improve the understanding of b quark production in pp collisions and to improve and calibrate theoretical predictions. The analysis follows the trail set by a similar measurement at 7 TeV, albeit with much higher statistics [2]. The thesis is outlined in the following paragraph.

Section 2 presents a brief overview of the theoretical framework. The focus lies in a qualitative description of the physics.
In Sec. 3 a description of the LHCb experiment follows. The different sub-detectors are introduced. Furthermore, the crucial role of selecting interesting events prior to data storage is highlighted.
The analysis is presented in Sec. 4. The measurement is carried out *globally*[2]. The differential measurement is featured separately (Sec. 4.9).
Finally, the thesis is concluded in Sec. 5 with a discussion of the results, where the measurements are compared to the theoretical predictions.

[1] all except gravity
[2] global is to be understood as *not binned* i.e. integrated over the phase space in the forward region.

2 Theory

Most phenomena observed in modern particle physics experiments are explained by the Standard Model (SM). The SM is based on a set of symmetries called *gauge* symmetries that lead to three of the four fundamental interactions. Although there are open questions and the search for physics not described by the SM is ongoing, the theory is very successful in explaining a broad range of phenomena. The SM does explain the strong-, the weak- and the electromagnetic interaction, but not gravitation. Shortcomings lie for example in the lack of an explanation for dark energy, dark matter or the abundance of matter over anti-matter [3].
The fundamental particles that can be encountered at particle colliders are presented in Sec. 2.1.1, including properties of important composite particles (Sec. 2.1.2).
The electroweak theory explains two of the four fundamental interactions in one swing and has succeeded in this task to a 0.1% accuracy [4]. Quantum chromodynamics is the theory behind the strong interaction. The fundamental SM interactions are described in Sec. 2.2.
Finally, the interactions of particles with the detector material are described briefly (Sec. 2.3).

2.1 Particle Zoo

2.1.1 Fundamental Particles

Particles that are assumed to be point-like (i.e. undividable) are called *fundamental* particles. We divide the fundamental particles into the two categories *fermions* and *bosons*. The particles are classified according to their interaction ability.

Fermions All matter consists of *fermions*. The characteristic property is a spin of 1/2. Each fermion has a corresponding anti-particle with opposite charge denoted by a bar (e.g. $\overline{b}$). There are two groups of fundamental fermions – quarks and leptons. Contrary to leptons, quarks are able to interact via strong interaction.

Leptons can be grouped into three isospin doublets or *generations* as shown in (1)[1]. The top row has an electrical charge of $Q = -1$, the bottom row of $Q = 0$ in units of electron charge. Of the charged leptons, only the electron is stable.

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix} \quad \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} \quad \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix} \tag{1}$$

Neutrinos are the only particles in the SM that *exclusively* interact weakly. Their interaction cross-section is very small[2]. At collider experiments, neutrinos are in practice undetectable.

Quarks are the building blocks for hadrons. They come in six flavours, that can be grouped to isospin doublets as shown in (2)[3]. The top row has an electrical charge of $Q = +2/3$ and the bottom row of $Q = -1/3$ in units of electron charge. Of these six quarks only the first two form stable particles.

[1] The anti-particles are not shown.
[2] An interesting way of stressing *how* small the interaction probability is, is to acknowledge that for a lethal dose of neutrino radiation one would have to be about an astronomical unit away from an active supernova [5].
[3] The anti-particles are not shown.

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix} \tag{2}$$

In addition to the electrical charge, quarks have color charge (red, green, blue and their anti-colors) that allows for strong interaction. Experimentally, quarks cannot be observed directly but only as a constituent of hadrons with zero net color charge[1]. The generations are ordered by mass from lightest to heaviest. The b quark is the second heaviest quark with a mass in the order of 4.5 GeV/c^2. It is the heaviest quark found in hadrons.

Bosons Fermions interact by exchanging force mediating particles called *bosons*. Bosons feature an integer spin. All bosons except the Higgs have spin 1. Table 1 lists the SM bosons including their mass.

boson	mass m [GeV/c^2]
$W^{\pm}$	80.385 ± 0.015
Z^0	91.1876 ± 0.0021
γ	0
g	0
H	125.7 ± 0.4

Table 1: List of SM bosons [1]. The eight gluons are summarised by g.

The most prominent gauge boson is the photon γ that mediates the electromagnetic force. The photon is known to be stable, massless and moving at the speed of light.
The isospin changing charged current is mediated by the bosons $W^{\pm}$. The neutral current is mediated by the Z boson. These bosons are massive and have a very small lifetime.
The strong force is mediated by the eight gluons g. The gluons are massless and carry color themselves. As a consequence, they can interact with each other.
Finally, the Higgs boson H has spin 0 and is responsible for the mass of fundamental particles.

2.1.2 Composite Particles

A hadron's behaviour is largely governed by the so-called *valence* quarks. The gluons that hold the hadron together constantly produce and annihilate $q\overline{q}$ pairs denoted by *sea* quarks. Hadrons with *three* quarks are called baryons, *two*-quark hadrons are called mesons consisting of a quark and anti-quark. Hadrons can be encountered in (albeit short-lived) excited states. Such particles are usually denoted by an asterisk (*), often including the mass in brackets.

Baryons Baryons consist of three quarks. The most common baryons are the proton p (valence quark content uud) and the neutron n (valence quark content udd) – building blocks for

[1] Except for the t quark that is the only quark that decays before hadronising [6].

the nuclei. Another example is the lambda baryon Λ_b^0 that contains an u, d and b quark. This baryon is unstable with a mean lifetime of $(1.451 \pm 0.013) \times 10^{-12}$ s. It is relatively heavy with $m = 5619.5 \pm 0.4$ MeV/c^2 [1].

Mesons A two-quark state is termed a meson. No stable mesons exist. The longest-lived meson is the charged pion with an average life-time in the order of tens of nanoseconds [1]. Some examples of mesons that are also used in the analysis are listed in Tab. 2. Each meson is assigned a corresponding anti-particle with charge-conjugated quark content and therefore opposite charge[1].

meson	quark content	mass m [MeV/c^2]	mean lifetime τ [s]
π^+	$u\overline{d}$	139.57018 ± 0.00035	(2.6033 ± 0.0005)×10^{-8}
K^+	$u\overline{s}$	493.677 ± 0.016	(1.2380 ± 0.0021)×10^{-8}
D^0	$\overline{u}c$	1864.84 ± 0.07	(410.1 ± 1.5)×10^{-15}
$D^*(2010)^+$	$\overline{d}c$	2010.26 ± 0.07	(7.89 ± 0.17)×10^{-21}
B^+	$u\overline{b}$	5279.26 ± 0.17	(1.638 ± 0.004)×10^{-12}
B^0	$\overline{d}b$	5279.58 ± 0.17	(1.519 ± 0.005)×10^{-12}

Table 2: Properties of mesons used in the analysis [1]. The corresponding anti-particles feature the same properties and are not shown in this table.

[1]The convention for the naming scheme of neutral mesons is that if the heaviest quark has negative charge, the meson is regarded an anti-particle. The opposite for the charge-conjugate.

2.2 SM Interactions

The SM has a long history of interplay between theory and experimental data [7]. In 1961, Sheldon Glashow succeeded in unifying the electromagnetic and the weak interaction in one theory. In 1967 Steven Weinberg and Abdus Salam added the electroweak symmetry breaking to the framework. The theory of the strong interaction was incorporated around 1973–1974.
The SM is based on the gauge group symmetries $SU(3)_C \times SU(2)_L \times U(1)_Y$. In the framework of quantum field theory, these symmetries define the structure of the SM Lagrangian $\mathcal{L}$.

2.2.1 Gauge Principle

A general Lagrangian for fermions is composed of fields and operators. For a fermionic field, a kinetic term containing covariant derivatives is written as

$$\mathcal{L}_{kin} = \overline{\Psi}(i\partial\!\!\!/ - m)\Psi \tag{3}$$

The postulate of an invariant Lagrangian after applying an infinitesimal symmetry group element results in the requirement of additional fields, called gauge fields. These fields compensate the left-over terms of the transformation, resulting in two additional players in the Lagrangian: the kinetic term of the gauge field and an interaction term, mixing fermionic and bosonic fields.
The strength with which an interaction enters the Lagrangian is determined by the *coupling strength* or *charge*. The space-time behaviour of the fields can be deduced by Euler's equations of motion. Noether's theorem predicts that for each symmetry, there exists a corresponding conserved current and charge. The internal symmetries $SU(3)_C \times SU(2)_L \times U(1)_Y$ lead to the conservation of color charge, electric charge and weak isospin, the external symmetries to the conservation of momentum, energy and angular momentum.

2.2.2 Electroweak Theory and the Higgs Mechanism

The $SU(2)_L \times SU(1)_Y$ groups generate the massless boson fields W^0, W^1, W^2 and B^0. However, these fields cannot be experimentally observed. The $SU(2)_L$ group is *defined* to act exclusively on particles with left-handed chirality, whereas the $SU(1)_B$ group is not chirality sensitive.
The Higgs field is introduced as a complex doublet. From the not enumerable set of possible representations of this field, nature decided to pick a particular one, where ν is the field in vacuum for $H(x^\mu) = 0$,

$$T_z: \quad \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad \rightarrow \quad \begin{pmatrix} 0 \\ \nu + H(x^\mu) \end{pmatrix} \tag{4}$$

i.e. a particular angle of rotation of $SU(2)_L$ around the z-axis is picked. Additionally, a representation of the $U(1)_Y$ symmetry is chosen, such that the vacuum representation ν of the Higgs field stays invariant under a rotation around the z-axis. This process is denoted by spontaneous symmetry breaking (SSB). However, there is a subgroup that is still invariant, called the electric group with conserved electric charge Q. The weak interaction *loses* its status as a symmetry of the vacuum.
The SSB rotation results in the fields A^0 and Z^0, W^+ and W^-:

$$\begin{pmatrix} A^0 \\ Z^0 \end{pmatrix} = \begin{bmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{bmatrix} \begin{pmatrix} B^0 \\ W^0 \end{pmatrix}, \tag{5}$$

where θ_W is called the *Weinberg angle* or *weak mixing angle*. A^0 corresponds to the field of the photon γ. Note, that the Z^0 boson mixes left-handed and right-handed interaction by an amount defined by θ_W.

$$\begin{pmatrix} W^+ \\ W^- \end{pmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix} \begin{pmatrix} W^1 \\ W^2 \end{pmatrix} \tag{6}$$

Most importantly, SSB explains the non-vanishing masses of the weak bosons. The masses are related by $M_Z = M_W/\cos\theta_W$ [8].

An example of a weak process is the decay $D^0 \to K^-\pi^+$ (Fig. 1). We call the decaying particle the *mother* and the decay products *daughters*.

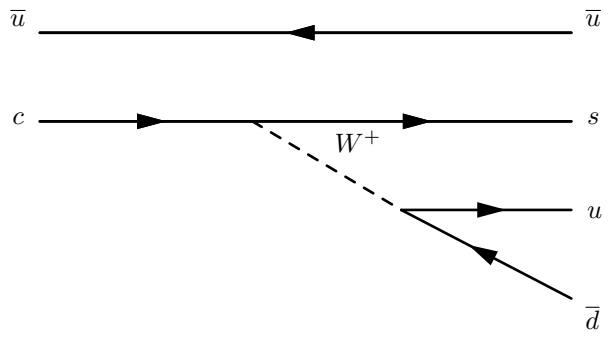


Figure 1: In this D^0 decay the mother (D^0 consisting of $c\overline{u}$) is on the left, the two daughters K^- ($s\overline{u}$) and π^+ ($u\overline{d}$) are on the right.

2.2.3 Quantum Chromo Dynamics

Quantum chromo dynamics (QCD) describes the interaction between quarks. Similar to the electroweak theory the $SU(3)_C$ symmetry is the foundation for the eight gluon fields. However, the third dimension adds some complexity. Since the generators behave in a non-abelian, way the gluons may interact among themselves, leading to complicated processes. The force behaves quite differently depending on the energy scale. There are two extreme behaviours:

- **Confinement** As quarks get separated, the energy of the gluon field increases up to the point where a $q\overline{q}$ pair is created. This leads to the non-existence of free quarks and gluons. Only colorless particles are observed.
- **Asymptotic Freedom** In interactions with collision energy, hadrons interact very weakly and can be approximated as interacting with a single constituent particle — denoted a *parton*.

QCD is not a purely predictive theory. Since the interaction can be strong, a perturbative expansion is not always feasible. It can only be done in the high-energy regime, although, there exist numerical and analytical methods to approximate QCD such as *lattice QCD* [9]. Below, important characteristics of the strong force are listed.

- **Running Coupling** In order to confront divergences in perturbative QCD, the coupling α_s is renormalised. The coupling *decreases* with increasing renormalisation scale μ_R, resulting in asymptotic freedom for hard processes[1].

- **Hadronization** Since a free quark cannot be observed, all quarks produced in hard scattering reactions must hadronise, i.e. become part of a baryon or meson. Perturbative QCD calculates processes on parton level, whereas the parton fragments to further particles. This showering is modeled separately by simulations.

- **Factorization** The factorization theorem allows the factorization of short- and long distance effects i.e. the separation of hard scattering calculations and parton distribution functions. The cross-section for a given final state f would then be

$$\sigma^f = \sum_{a,b} \int \int \mathrm{d}\xi_{a'} \mathrm{d}\xi_{b'} f_a(\xi_{a'}, \mu_F^2) f_b(\xi_{b'}, \mu_F^2) \sigma_{ab}^f(\xi_{a'}, \xi_{b'}, p, \mu_F^2, \mu_R^2), \tag{7}$$

  where the sum runs over all possible partons a and b and the integral over their momentum fraction. The parton-parton cross-section for a final state f is denoted by σ_{ab}^f. The parton distribution function (PDF) $f_i(\xi, \mu_f^2)$ is crucial for any theoretical calculation. It defines the probability of a parton i having a momentum fraction ξ of the hadron momentum p at a scale μ_f [10]. The PDF has to be determined by experiment.

2.2.4 CKM Matrix

In the quark sector, the *Cabbibo-Kobayashi-Maskawa* (CKM) matrix is responsible for the possibility of generation-changing weak decays. The mass eigenstates for quarks (production- or strong eigenstates) do not correspond to the eigenstates of the weak interaction. The rotation from the strong set to the weak set is described by the complex 3×3 unitary matrix V_{CKM}:

$$\begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix} = \begin{pmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{pmatrix}, \tag{8}$$

where the prime notation denotes the weak eigenstates [11]. Thus, the matrix element for the decay from initial state i to final state f is proportional to $V_{if} V_{if}^* = |V_{if}|^2$. The absolute value of the diagonal elements is close to one, whereas the off-diagonal elements are small.

If the CKM-matrix features complex phases, the CP (charge conjugation $\times$ parity) operator is no longer a symmetry of the weak interaction. CP violation has been experimentally observed [12] and is a prevailing topic in present-day research.

2.2.5 Lagrangian

The SM Lagrangian is a large and complicated construct that fills one page of paper in its raw form. Doing calculations with the SM Lagrangian can be very involved. The number of interactions (excluding Higgs interactions) can however be summarised in an efficient way using Feynman diagrams (Fig. 2).

[1] A *hard* process is an interaction with a high center-of-mass energy contrary to a low-energy *soft* process.

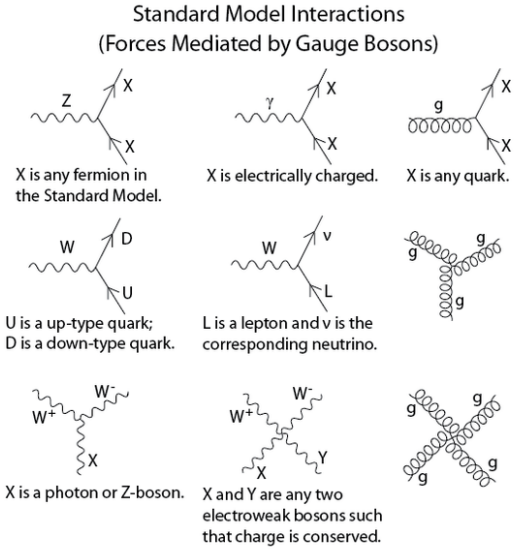


Figure 2: The nine SM interactions can be summarised by Feynman diagrams. The Higgs boson interactions are not included [13].

2.3 Interaction with Matter

The following list presents a short description of particle interactions with detector material.

Ionization and Excitation Scattering of charged particles off detector material leads to ionization and excitation of shell electrons. The energy loss per distance is described by the Bethe-Bloch formula [14] and is proportional to

$$-\frac{\mathrm{d}E}{\mathrm{d}x} \propto \frac{z^2}{m\beta^2}\left[\ln\left(\frac{\beta^2}{1-\beta^2}\frac{1}{I}\right) - \beta^2\right], \tag{9}$$

where z is the charge of the particle with mass m, β its speed in units of c and I the mean excitation energy of the target material. Note that $I/Z \approx const.$ for an atomic nucleus of charge $Z[e] > 20$. This energy deposit is most important for tracking detectors.

Coulomb Scattering Charged particles passing through matter scatter in the Coulomb field of the nucleus. The energy loss is negligible. However, the particle can get deflected. The angular distribution is smeared by a Gaussian with large tails. This effect worsens the resolution of the track reconstruction. Therefore, the amount of material used for a detector has to be kept small.

Bremsstrahlung For low mass charged particles (primarily $e^{\pm}$) Bremsstrahlung is the dominant process of energy loss. After one radiation length X_0 the electron loses all but $1/e$ of its energy. X_0 is roughly $\propto \frac{1}{Z^2}$, where Z is the atomic charge of the target material.

Photon Absorption Absorption of photons in matter is highly energy-dependent and follows an exponential governed by the radiation length X_0. X_0 is equal to 7/9 of the mean free path of a photon in material. Photon absorption involves the photoelectric effect, Compton scattering and pair production. Electron-positron pair production $\gamma \to e^+e^-$ and Bremsstrahlung lead to an electromagnetic showering.

Inelastic Scattering If at least one of the scattering particles is composite (e.g. a nucleus), the constituents can absorb collision energy. If the energy transfer is large enough, the particle can escape the hadron (and perhaps produce other particles).
The nuclear interaction length λ_i is the mean path length required to reduce the numbers of relativistic charged particles by the factor $1/e$ as they pass through matter. However, this quantity does include elastic processes that lead to diffraction and is generally longer than the mean free path of inelastic scattering.

3 The LHCb Detector

The *Large Hadron Collider beauty* (LHCb) detector is a forward spectrometer contrary to the other LHC experiments. It is specifically designed for precision measurements of CP violation as well as Rare Decays of b and c hadrons. At LHC energies, *beauty* quarks b are predominantly produced in the forward or backward direction (Fig. 3).

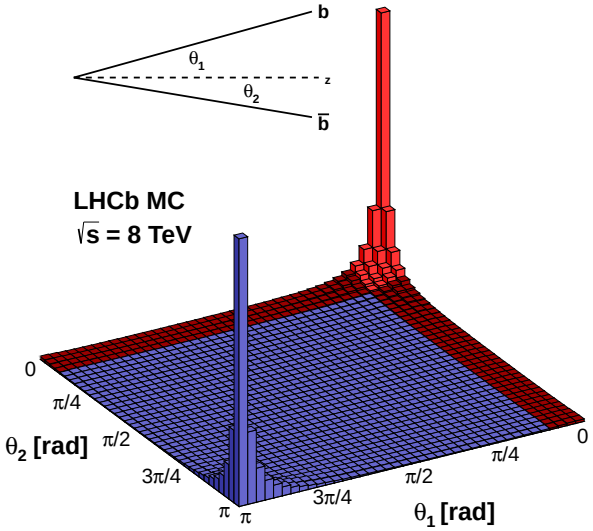


Figure 3: Production angles of a $b\bar{b}$ quark pair. The angle θ_1 is the angle of the b quark with respect to the beam axis, θ_2 the angle of the $\bar{b}$ quark with respect to the beam axis [15].

The pseudo-rapidity η is defined as $\eta = \frac{1}{2}\ln\left(\frac{|\boldsymbol{p}|+p_L}{|\boldsymbol{p}|-p_L}\right)$. It can be written as $\eta = -\ln\left[\tan(\theta/2)\right]$, where θ is the angle of the particle with respect to the beam axis (Fig. 4). The geometry of LHCb allows a coverage of $1.9 < \eta < 4.9$. In terms of θ, the angular acceptance translates to $10 \text{ mrad} < \theta < 250 \text{ mrad}$.

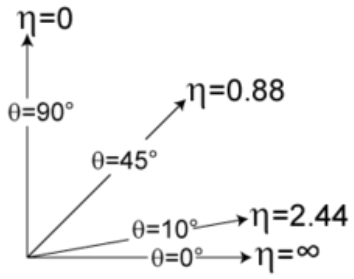


Figure 4: The pseudo-rapidity η diverges toward the z-axis (horizontal) and vanishes for high angles θ [16].

Figure 5 shows the detector layout. A right-handed coordinate system is used. The positive z-axis extends along the beam, the y-axis perpendicular to the ground floor. In cylindrical coordinates, θ is the angle with respect to the z-axis and ϕ the angle in the (x, y) plane.
A lot of effort has been made to reduce the amount of material used for detector and support structures. The types of material have been carefully chosen to reduce the number of interactions

in non-active material.
In the following an overview of the detector concept and data acquisition mechanism is given. A detailed description of the LHCb detector and its performance can be found in Ref. [17] and Ref. [18] where most of this section's content originates from.

3.1 Purpose

To account for the matter-antimatter asymmetry in the universe, the CP violation present in the CKM matrix is not sufficient in strength. Additional sources of CP violation may emerge from physics beyond the SM. The main goal of the LHCb detector is to search for indirect evidence of new physics in CP violation and Rare Decays of beauty and charm hadrons. The analysis focuses on effects of potential new particles. New physics might be seen in CP violation measured with hadrons containing b or c quarks or in decays of these quarks which are suppressed in the SM.

3.2 Detector Layout

In the following, a brief description of the detector components is given.

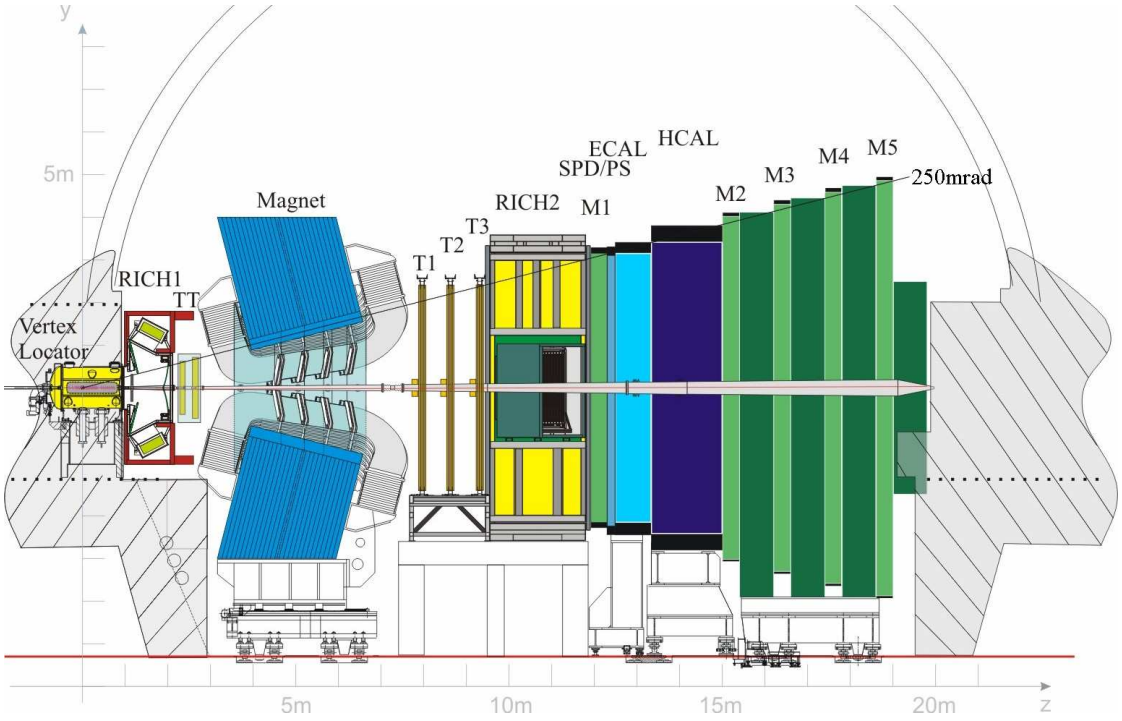


Figure 5: Detector layout of LHCb [17]. The detector components are introduced from left to right: VELO, RICH1, TT, magnet, tracking stations T1-T3, RICH2, the upstream muon station M1, SPD/PS, the calorimeters ECAL and HCAL and the muon stations M2 - M5. The maximum opening angle θ with respect to the z-axis is drawn as a thin blue line. The cavern where LHCb is situated in, is indicated as a grey semi-circle. Note the faint y- and z-axis.

The **Vertex Locator** (VELO) is located around the interaction region and kept in a secondary vacuum with respect to the beam vacuum. It contains 42 silicon sensor modules (Fig. 6) arranged along the beam. Each module provides a measurement of (r, ϕ). The VELO provides the distinction between primary- (PV) i.e. *pp* collision vertex and secondary vertices (SV) i.e. decay vertices of particles with a lifetime of 1 ps or larger. During beam injection, the modules are retracted for safety reasons until a stable beam is declared.

Figure 6: a photograph of LHCb's silicon modules [17]

Each **Tracking Station** features four detection layers with the middle layers rotated by $-5°/+5°$ around the z-axis to achieve a better position resolution. The Tracker Turicensis (TT – designed and built at UZH) is situated upstream, the three stations T1 to T3 downstream of the magnet. The TT uses silicon microstrip sensors with a pitch of 183 µm. T1 to T3 consist of an inner silicon strip detector called Inner Tracker (IT) covering the region close to the beam pipe. The outer region is covered by a gas drift-tube tracker called outer tracker (OT). The OT consists of about 300 straw-tube modules with drift-time readout.

The bending dipole **Magnet** is needed to measure the momentum of charged particles. It provides a magnetic field integral of about 4 Tm, deflecting particles in the (x, z) horizontal plane. Since the field affects the LHC beam in the pipe, three dipole magnets are used to compensate for the effect in the beam. The magnet can be operated with two polarities – *up* or *down*. Switching the magnet polarity regularly in between runs yields two data sets. The difference of the two sets may be used to examine charge-dependent efficiency effects. The bending power of the magnetic field increases the angular acceptance to 10 mrad $< \theta <$ 300 mrad in the bending plane.

For charged particles in the momentum range from 2 to 100 GeV/c, particle identification (PID) is achieved by two **Ring Imaging Cherenkov Detectors** (RICH). Depending on the radiator material, a different momentum range is covered. RICH1 located upstream of the dipole magnet uses Aerogel and C_4F_{10} as a radiator for a momentum range of about 2 to 60 GeV/c. RICH2 situated downstream of the dipole magnet uses CF_4 as a radiator covering 15 GeV/c to 100 GeV/c. Its main use is to distinguish charged kaons and pions.

The **Scintillation Pad Detector** (SPD) consists of two arrays of scintillator pads of high gran-

ularity enclosing a lead absorber/converter of 2.5 radiation lengths (X_0). It is used to separate photons from electrons by a binary threshold readout of the deposited energy. Additionally, it is used to identify events with a very high hit occupancy. Since these events lead to very long processing times during reconstruction, they are rejected at an early stage (see section 3.3 for the trigger system).

The **Pre-Shower Detector** (PS) is a similar scintillation detector designed to discriminate a γ from a π^0 decaying into two photons. Contrary to the electromagnetic calorimeter, the PS has a finer granularity but does not measure energy.

The **Electromagnetic Calorimeter** (ECAL) is a sampling scintillator/lead structure with a thickness of 25 X_0 that measures the energy and the position of the energy deposit. Like most calorimeters, this is a destructive measurement (the particles get destroyed or stopped in the process).

The **Hadronic Calorimeter** (HCAL) is a sampling calorimeter made from iron and scintillation tiles with a length of 5.6 nuclear interaction lengths (λ_i).

The **Muon Stations** provide muon identification and serve an important role in the trigger. There are five muon stations M1 - M5 equipped with multi-wire proportional chambers (MWPC), except for the inner part of M1 where gas electron multipliers (GEM) are used. M1 is placed in front of the calorimeters and to improve the p_T resolution for muons in the hardware trigger system[1].

3.3 Trigger System

If the LHC collides bunches under design conditions of 40 MHz, the rate of pp interactions producing at least two tracks within the LHCb acceptance is 10 MHz. Event data cannot be written to disk at this rate. Therefore, in the interest of the LHCb physics design, this rate is reduced by the trigger system. The trigger system consists of two levels.
The hardware trigger (level zero or L0) is designed to reduce the event rate from the LHC 40 MHz to a maximum of 1.1 MHz. It uses information from the calorimeters and the muon system. The trigger selects events containing particles with a high transverse energy E_T in the calorimeters or a high transverse momentum p_T measured in the muon chambers.
The events accepted by L0 are sent to the event filter farm to be processed for the high level trigger (HLT). The HLT is a simplified version of the offline event reconstruction and runs on software. The rate is reduced to about 2 kHz. The first stage (HLT1) partially reconstructs the event, reducing the rate by a factor of about 30 while the second stage (HLT2) performs a full event reconstruction [19].
There are multiple triggers that can lead to an event being accepted. If an event has been accepted, we say it has passed the *global* trigger decision, whereas it might not have passed a specific trigger.

[1] less scattering in detector material that smears p_T

3.4 Event Reconstruction

The tracking systems consists of the sub-detectors VELO, TT, T1 - T3 and the magnet. Hits of these stations get combined to form trajectories of charged particles called a *track*. The bending radius is used to determine the momentum. The identification of the different particle types uses information from the two RICH, the calorimeter and the muon system. The events passed from the trigger system are further selected and archived according to physics analysis interest. The resulting data structure is stored for the final analysis. The process of preselecting interesting events is called *stripping*, the sequence of this selection a *stripping line* [20]. Stripping lines that feature similar characteristics are grouped into *streams*.

3.5 Luminosity

In 2012 the LHC beam center-of-mass energy was increased from 7 TeV to 8 TeV. The instantaneous luminosity was 4×10^{32} cm^{-2}s^{-1}. This is about twice the design luminosity of LHCb.
There are two methods used to determine the luminosity at LHCb [21]. The *van der Meer scan* measures the interaction rate of well-understood processes as a function of beam separation to determine the luminosity. The *beam-gas imaging* method is unique to LHCb and is based on reconstructing vertices of interactions between beam particles and residual gas molecules in the beam vacuum. The integrated luminosity for 8 TeV recorded in 2012 is determined as

$$\begin{aligned} \mathcal{L} &= (990 \pm 12)\ \text{pb}^{-1} \quad \text{(magnet polarity down)} \\ \mathcal{L} &= (997 \pm 12)\ \text{pb}^{-1} \quad \text{(magnet polarity up)} \end{aligned}$$

4 Analysis

Notes The discussion in this section includes the charge-conjugate mode. Additionally, a color scheme has been used for consistency. *Red* stands for signal, *blue* for background and *gray* for data. The data is usually shown with error bars, whereas signal and background templates are shown as histograms without error bars.
For simplicity, the discussion is presented for data recorded with a downward magnet polarity only. The analysis of the data recorded with an upward magnet polarity gives a similar result. The two polarities are combined in Sec. 5.
The combination of a kaon candidate and a pion candidate to a D^0 is be denoted a 'D^0 candidate'. Likewise, the combination of D^0 candidate and muon candidate is denoted a 'B candidate'.
This measurement provides an *inclusive* measurement in contrast to an *exclusive* measurement. This means that part of the final state is left unmeasured i.e. there is no information used about about additional particles denoted by X in the final state of $b \to D^0 \mu^- \overline{\nu}_\mu X$.

4.1 Analysis Strategy

The cross-section σ is calculated as

$$\sigma = \frac{1}{2\mathcal{L}} \frac{N}{\prod_i \varepsilon_i \mathcal{B}}, \tag{10}$$

where N is the signal yield extracted from data, $\mathcal{L}$ is the (measured) integrated luminosity, $\mathcal{B}$ the branching fraction of the decay channel and $\prod_i \varepsilon_i$ the product of the various efficiencies. The factor $1/2$ arises from the fact that we are measuring 'b or $\bar{b}$' but require 'b *and* $\bar{b}$'.
The transverse momentum p_L is the momentum of a particle along the beam axis, whereas p_T is the momentum perpendicular to the beam axis. The rapidity y is defined as $y = \frac{1}{2}\ln\left(\frac{E+p_L}{E-p_L}\right)$, where E is the particle's energy. In high energy physics, one often uses y and p_T to parametrise the kinematical dependence of the cross-section. For relativistic particles, the pseudo-rapidity converges to the rapidity. The differential cross-section in terms of p_T and y of the B candidate is calculated as

$$\frac{\partial^2 \sigma}{\partial \eta \partial p_T} = \frac{1}{2\mathcal{L}} \frac{N}{\prod_i \varepsilon_i \mathcal{B} \Delta p_T \Delta y}, \tag{11}$$

where Δp_T and Δy is the bin area in units of [GeV/c]. The signal yield and efficiency calculation is carried out for each bin.
To obtain the (differential) cross-section, a sequence of steps is carried out with a dedicated section for each part. An event is called *signal* (short sgn), if it originates from a b quark that decays through the chosen decay channel. On the contrary, an event is called *background* (short bkg) if it originates from other particles or other decay channels.

- **Decay Channel** The decay channel is presented, including branching ratios and signal composition. Furthermore, the different sources of background are specified (Sec. 4.2).

- **Selection Requirements** With knowledge of the signal and background topology, the signal-to-background ratio can be improved by imposing selection criteria on the data (Sec. 4.4). The efficiency ε is defined as the ratio of the number of *signal* events after the selection to number of signal events before the selection. The purity ρ is defined as the ratio of the number of *signal* events to the number of signal plus background events. The goal of the selection is to maximize the product of ε and ρ. The resulting position in the (ε, ρ)-plane is

called the *working point*. The signal yield extraction is much easier at this point as it allows for cleaner fits.

- **Determine the Number of D^0 Candidates** The quantization of the number of D^0 events is achieved by fitting the D^0 candidate mass peak (Sec. 4.5.1).
- **Determine the Number of B Candidates** A b hadron travels a significant distance from the production vertex before decaying. The impact parameter IP is defined as the shortest distance of the particle's path to the primary vertex (Fig. 7).

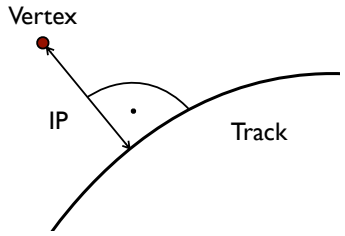


Figure 7: The IP is defined as the shortest distance of a particle's track to the primary vertex.

Thus, one can use the IP to distinguish D^0 produced at the PV and D^0 originating from a secondary decay vertex. By fitting the IP of the D^0 candidate one can extract the number of D^0 originating from a b hadron (Sec. 4.5.2).

- **Correct for the Missing Neutrino Energy** Since the neutrino leaves the detector unmeasured, one has to correct for the missing neutrino energy. This is particularly important for the differential measurement (Sec. 4.9.2), as the signal yields migrate in between bins.
- **Correct for Efficiencies** The signal yield is corrected for the various selection requirements, detection restrictions and acceptance effects by weighting the result with efficiencies ε_i (Sec. 4.6).
- **Calculate Cross-Section** Taking into account the integrated luminosity and the branching fraction of the decay channel, one can calculate the cross-section. The cross-section result is not shown until Sec. 5.
- **Calculate Uncertainties** Finally, the systematic- and statistical uncertainties are determined (Sec. 4.8).

These differential measurement is featured separately in Sec. 4.9.

4.2 Decay Channel

A b quark hadronises into different types of B mesons and b baryons. These particles have different probabilities to decay through the considered decay channel. An alternative point of view is to assume an admixture of b hadrons instead of looking at each type individually. This section presents the b hadron fractions and the decay channel branching ratio for this admixture.

b Hadronization To reproduce the energy and direction of the b quark, the assumption is made that the observed bottom hadron largely inherits the properties of the original b quark. In particular, the direction of motion at the production vertex is assumed to be unchanged. The production is assumed to be symmetric between b and $\overline{b}$ quarks.

b Hadron Production Fraction Light quarks (u,d,s) get produced with high multiplicity in the hadronization, whereas heavier quarks are found at a comparably rather low rate. The cross-section for charm production is about a factor of 10 smaller than the total interaction cross-section. For beauty production the it is even a factor of 200.
In Tab. 3 the b hadron production fractions measured at high energies are listed. These fractions[1] are assumed to hold at LHC energies. The fractions are provided by the particle data group (PDG) [1].

b hadron	fraction
$B^{\pm}$	$(40.2 \pm 0.7)\%$
$B^0/\overline{B}^0$	$(40.2 \pm 0.7)\%$
B_s^0	$(10.5 \pm 0.6)\%$
b baryon	$(9.2 \pm 1.5)\%$

Table 3: The measured b quark hadronization fractions above the Υ(4S) resonance energies measured at LEP are found in the PDG database [1].

Note the isospin symmetry between the $B^0/\overline{B}^0$ and $B^{\pm}$ production which are also the most abundant b hadrons.
b Decay Channel For this particular admixture of b hadrons, the probability for the chosen decay channel is given in Tab. 4. This fraction will be used for the cross-section determination. Because of lepton universality[2], usually l is written for both muons μ and electrons e.

decay mode	branching fraction $\mathcal{B}$
b hadron admixture $\rightarrow \overline{D}^0 l^+ \nu_l X$	$(6.84 \pm 0.35)\%$

Table 4: The branching fractions hold for the b hadron admixture of Tab. 3 [1].

The chosen decay channel's branching fraction is relatively large which is beneficial for the measurement since it leads to a large signal yield. Furthermore, the background contribution is smaller compared to non-leptonic channels.
Since the neutrino momentum cannot be measured, the sum of the D^0 and muon four-momentum will not amount to the b hadron mass. Instead, the mass will be smeared out and shifted towards lower masses.
The finite albeit small lifetime of the b hadrons implies that they travel a distance of the order of 1 cm before decaying. Therefore, the combined particle reconstructed from the D^0 and muon points to the secondary vertex (distorted by the missing neutrino momentum). The same holds for the charge-conjugates.

D^0 Decay Channel The relevant branching fractions for the D^0 decay are given in Tab. 5. The decay $D^0 \rightarrow K^-\pi^+$ is chosen to reconstruct D^0 candidates. The different fractions are explained by the CKM matrix.

[1] The values are combined from multiple measurements above the $\Upsilon(4S)$ resonance.
[2] Lepton universality means 'if the masses of the lepton flavours were equal, their branching fractions would be equal'. At LHC energies, the mass difference between electrons and muons is negligible. The tau is much heavier, thus the interaction capability is suppressed by its mass.

decay mode	branching fraction $\mathcal{B}$
$D^0 \to K^-\pi^+$	$(3.88 \pm 0.05)\%$
$D^0 \to K^+\pi^-$	$(1.38 \pm 0.03) \cdot 10^{-4}$
$D^0 \to K^+K^-$	$(3.96 \pm 0.08) \cdot 10^{-3}$
$D^0 \to \pi^+\pi^-$	$(1.40 \pm 0.03) \cdot 10^{-3}$

Table 5: These branching fractions of the D^0 decay feature a two-body final state composed of charged kaons and pions [1].

4.2.1 Signal Composition

Signal Composition Appendix A lists the branching ratios for semi-leptonic decays including a D^0. We denote D^0 originating from a b hadron as 'dfB' (D from B). The expected signal contribution in Tab. 6 can be calculated using these branching ratios and the production fractions of Tab. 3. Note, that this calculation does not include efficiencies differing between the types of b hadrons, that may have an effect.

b hadron	branching fraction $\mathcal{B}$	signal fraction
$B^\pm$	$(9.53 \pm 0.39)\%$	$(60.9 \pm 6.5)\%$
$B^0/\overline{B}^0$	$(5.70 \pm 0.43)\%$	$(37.6 \pm 6.4)\%$
B_s^0	$(0.86 \pm 0.21)\%$	$(1.4 \pm 0.4)\%$
b baryon	~ 0	~ 0

Table 6: The expected signal composition is calculated summing over the individual muonic decays and multiplying it with the admixture fractions of Tab. 3 [1].

Clearly, the signal is dominated by $B^\pm$ and $B^0/\overline{B}^0$ events. The B_s^0 contribution is fairly small as is the b baryon contribution. We will neglect B_s^0 and b baryon events throughout this analysis.

Most decay modes featured in appendix A include intermediate excited states denoted D^*-like. These states may decay to a D^0 releasing one or more soft pions in the process. These additional particles have a minor effect on the selection- and acceptance efficiency. One can therefore divide the signal into two types that we denote D^0-like and D^*-like. Efficiencies are calculated separately for these two cases. Table 7 shows the fractions calculated from PDG data.

type	signal fraction
D^0-like	$(14.2 \pm 1.3)\%$
D^*-like	$(85.8 \pm 1.3)\%$

Table 7: The expected signal composition in terms of D^0-like and D^*-like is calculated using branching fractions from PDG [1].

The two signal types are shown in Fig. 8 and Fig. 9[1].

[1]Note: These figures are not actual Feynman diagrams but rather graphical illustrations of a decay channel.

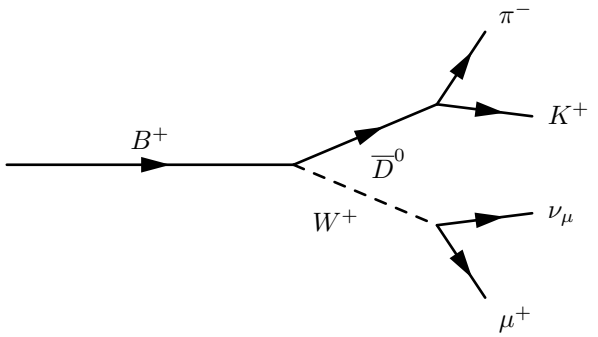


Figure 8: In this example of a D^0-like decay, the arrow lengths do not correspond to the FD of the particles. The W boson decays immediately after production.

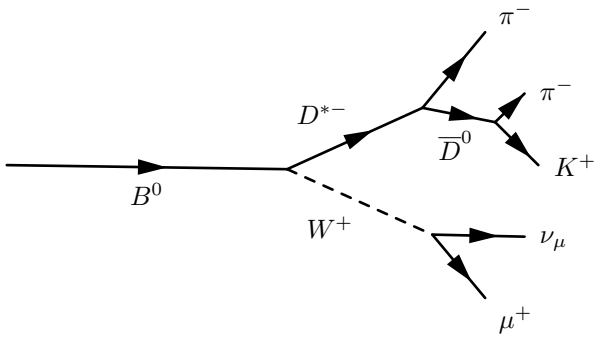


Figure 9: In this example of a D^*-like decay, the arrow lengths do not correspond to the FD of the particles. The W boson and the excited D^{*0} decay immediately after production.

4.2.2 Background Composition

At the defined working point of the analysis, several processes mimic the signal signature $b \to D^0\mu^-\overline{\nu}_\mu X$ with $D^0 \to K^-\pi^+$. These events are denoted as *background*. The following list describes the types of background encountered in this analysis.

- **D^0 Background Events** The sources of this background are numerous, such as a kaon and a pion that are wrongly combined to a D^0 candidate or involvement of additional particles in a D^0 decay such as in $D^0 \to K^-\pi^+\pi^0$. We denote this type of background as 'combinatorial'.

- **D^0 Misidentification (Mis-ID) Events** This background emerges from $b \to D^0\mu\overline{\nu}_\mu X$ but with a D^0 that decays through another channel. Paired with single or double particle mis-ID, these events mimic the signal signature e.g. $D^0 \to K^+K^-$ with K^+ misidentified as a π^+. Since there is a real B decaying to a D^0 with IP and $m_{K\pi}$ similar to the signal, the mis-ID background cannot be separated from the signal by fitting. Instead, we subtract the background separately as described in Sec. 4.5.3. Although compared to other D^0 channels, the analysis channel features a much higher statistics and the mis-ID of a kaon as a pion (and vice versa) is only 3% to 10% [22], the mis-ID background amounts to a relative contribution of $\sim 0.7\%$.

- **Prompt Events** These events contain true D^0 particles produced at the primary vertex. We denote this background as 'prompt'. It is noteworthy, that *most* D^0 get produced promptly.

- **μ Mis-ID Events** The muon-pion and muon-kaon mis-ID probability of $\sim 1\%$ to $\sim 3\%$ is smaller compared to the kaon–pion mis-ID probability [23]. This background will be even smaller than the D^0 mis-ID background and is neglected.

- **Events from $b \to D^0\tau^-(\to \mu^-\overline{\nu}_\mu\nu_\tau)X$** Appendix B lists the measured decay channels for $b \to D^0\tau^-\overline{\nu}_\tau X$. Table 8 lists the expected signal contribution. Clearly, the contribution to the signal fraction is small. Since τ particles have a small albeit non-zero lifetime, the vertex reconstruction will miss $D^0\tau$ vertices, reducing this contribution substantially. Thus, if one takes efficiencies into account, the signal contribution is dropping to about one percent, as will be shown later.

b hadron	signal fraction
$B^\pm$	$(1.4 \pm 0.2)\%$
$B^0/\overline{B}^0$	$(2.0 \pm 0.3)\%$
B_s^0	~ 0
b hadron	~ 0
total τ	$(3.4 \pm 0.3)\%$

Table 8: expected signal contribution $b \to D^0\tau^-\overline{\nu}_\tau X$ with $\tau \to \mu\overline{\nu}_\mu\nu_\tau$

4.3 Samples

Modern particle physics experiments would be impossible to perform without simulations. Such stochastic simulations are known as *Monte Carlo* (MC). Simulations are often used to determine efficiencies. These simulations have been compared extensively to detector data and adjusted. For this analysis, we have great trust in the accuracy of the simulation.
Two subsamples are distinguished. The first event type is composed of $\mu^+\overline{D}^0 X$ is called the *right-sign* (RS) sample, where the sign of the kaon is the same as the sign of the muon. $\mu^+ D^0 X$ is called the *wrong-sign* (WS) sample, where the sign of the kaon is opposite to the sign of the muon. The WS sample is strongly suppressed by the CKM matrix (also denoted *Cabbibo* suppressed). It will be used to check the mis-ID background estimates.
The events in simulation as well as in data are analysed and processed by LHCb's DaVinci package (version v36r3).
All simulation samples are generated with Pythia6 [24][25] as an event generator with a dedicated configuration [26]. The particle decays therein are simulated with EvtGen [27] while the final state radiation is described by the Photos package [28]. The interaction of the generated particles with the detector material is simulated by Geant4 [29][30].

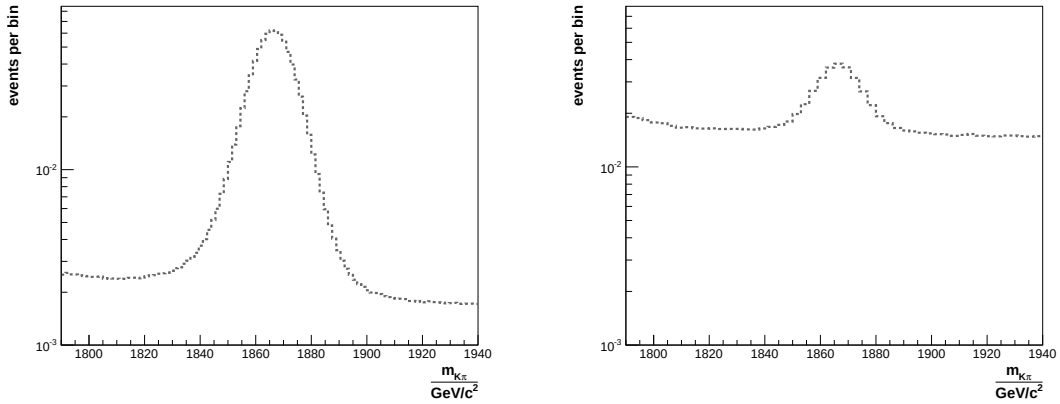
Data Sample The data sample has been selected by the stripping line *S20*[1] [31]. The high statistics in this stripping line is reduced by introducing a pre-scale of 0.5, randomly selecting every second event. This will be accounted for in the efficiency calculation. Specific triggers are required to reduce background[2]. The data file contains ~ 8.5 million events for each magnet polarity.
Signal Samples Two Monte Carlo signal samples have been generated. Ideally, there would be a signal file for each of the possible b mesons $B^{+/-}$, $B^0/\overline{B}^0$, B_s^0 and one for b baryons. Unfortunately, since there is limited amount of simulation power available, we have do serve with two signal templates. Because Tab. 6 predicts a mostly $B^{+/-}$ and $B^0/\overline{B}^0$ dominated signal, it is appropriate to use

- $B^{\pm}$ signal sample containing B^+ and B^- primary particles. It contains 0.5 million events generated in the LHCb acceptance.
- $B^0/\overline{B}^0$ signal sample containing B^0 and $\overline{B}^0$ primary particles. It contains 2 million events generated in the LHCb acceptance.

The number of events is chosen arbitrarily.
Prompt Sample Additionally to the simulated signal, a prompt D^0 background sample has been generated. Its D^0s are decaying as $D^0 \to K^-\pi^+$ without possessing a mother particle. It contains 2 million events generated in the LHCb acceptance, used for both RS and WS. Ideally a D^0 background *including* a mother would be in order. However, the contribution of this background is small at the working point and the effect of the missing mother will be accounted for in the uncertainty.
Combinatorial Sample The D^0 peak maximum resides at 1866.2 MeV/c^2. The background of

[1] more exact: the *b2D0MuXB2DMuNuX* stripping line
[2] The L0 trigger *L0Muon* requires $p_T(\mu) > 1.76$ GeV/c. The HLT1 *Hlt1TrackMuon* is based on a confirmed muon candidate from L0Muon and requires $p_T > 1$ GeV/c and $p > 3$ GeV/c. The HLT2 *Hlt2Topo* imposes topological requirements. It selects events based on a combination of two, three or four tracks which might originate from b hadrons by requiring a high IP and a detached vertex.

D^0 candidates is modelled by requiring that either 1810 MeV/c^2 $< m_{K\pi} <$ 1831.2 MeV/c^2 or 1901.2 MeV/c^2 $< m_{K\pi} <$ 1940 MeV/c^2 i.e. the sidebands 35 MeV/c^2 from the peak. The left portion of the background mass window is smaller due to a resonance appearance in the lower spectrum that is clearly visible in the WS, but is also present in the RS (Fig. 10).



(a) RS D^0 mass window

(b) WS D^0 mass window

Figure 10: The D^0 candidate mass window in data is divided into the central peak region and the sideband region 35 MeV/c^2 from the peak position signified by the two vertical lines. The data uncertainties are too small to be displayed.

4.4 Selection Requirements

This section describes the cuts defining the working point where the fits will be performed in. To minimise a bias in the fit of m_{D^0} or IP_{D^0}, cuts are performed as loosely as possible i.e. a cut in the signal region is avoided. The following list defines the variables, that have not yet been introduced.

- $\boldsymbol{\chi^2_{\mathrm{Track}}}$: The track reconstruction software looks for hit patterns that form a potential track. Due to the limited amount of measurement points and scattering in material, wrong combinations of hit points can be reconstructed as a track. To quantify the goodness of the track fit, the distance from the impact points to the track are summed up and divided by their experimental uncertainty $\chi^2_{\mathrm{Track}} = \sum \frac{(\boldsymbol{x}_i - \boldsymbol{x}_{\mathrm{Track}})^2}{\sigma_i^2}$. The χ^2 is divided by the number of degrees of freedom (d.o.f.).

- $\boldsymbol{\chi^2_{\mathrm{Vertex}}}$: The χ^2 of the daughter particles forming a common vertex. It is divided by d.o.f. The vertex fitter and χ^2 calculation are described in Ref. [32].

- $\boldsymbol{\chi^2_{\mathrm{IP}}}$: The χ^2 of the impact parameter can be interpreted the IP squared divided by its experimental uncertainty squared[1].

- $\mathbf{PID}(\boldsymbol{X})$: PID is short for *particle identification*. Using RICH data as well as data from the calorimeters and the muon stations the particle hypothesis X can be quantified. It is defined as the difference in the log-likelihoods (DLL) of being particle X to being a (charged) π.

- $\boldsymbol{p}_{\mathrm{ghost}}$: A 'ghost track' is a term for a track, that was erroneously reconstructed from random hit points. A detailed description of the track ghost probability calculation is given in Ref. [33].

- $\boldsymbol{m_{K\pi}}$: The reconstructed mass $m_{K\pi}$ of the D^0 candidate is calculated by adding and squaring the momentum four-vectors of the K candidate and the π candidate.

- **DIRA**: DIRA is short for *direction angle* with respect to the candidate's primary vertex. It is the cosine of the angle between the particle momentum, reconstructed from its daughter particles and the line joining the end- and origin vertices.

- **FD**: The *flight distance* is the distance between the end- and origin vertex of a particle.

- $\boldsymbol{\chi^2_{\mathrm{FD}}}$: The χ^2 formed by dividing the FD by the experimental vertex position errors $\chi^2_{\mathrm{FD}} = \mathrm{FD}^2/(\sigma^2_{OV} + \sigma^2_{EV})$, where OV stands for origin vertex and EV for end vertex.

- **DOCA**: DOCA is short for *distance of closest approach* of a candidate's daughter particles.

- $\boldsymbol{N}_{\mathrm{Tracks}}$: The track multiplicity is defined as the number of tracks observed in the event.

The selection requirements are summarised in Tab. 9. For each particle, relevant variables are presented. For simplicity the simulated signal sample is composed as in Tab. 6, just without B^0_s. The background distribution is shown as a dashed line, signal as a solid line. Both simulation and data in the following figures have been stripped.
Since the angular acceptance of LHCb is limited, $1.9 < \eta < 4.9$ is imposed on all particles.

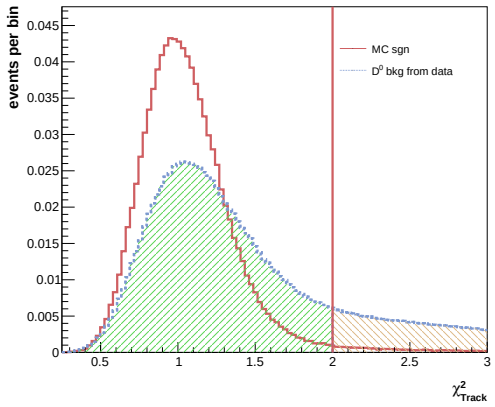
[1] At LHCb, the χ^2_{IP} is calculated a different way. It is calculated as the increase of the χ^2_{Vertex} of the PV fit when one adds the track into the vertex. It behaves almost like $\left[\mathrm{IP}/\sigma(\mathrm{IP})\right]^2$.

candidate(s)	minimal requirement	additional requirement
K, π, μ	$1.9 < \eta < 4.9$	-
K, π	$\chi^2_{\text{Track}} < 4$	$\chi^2_{\text{Track}} < 2$
K, π	$p_{\text{ghost}} < 0.5$	-
K, π	$p > 2$ GeV/c	-
K, π	$p_T > 0.3$ Gev/c	-
K, π	$p_T(K) + p_T(\pi) > 1.4$ GeV/c	-
K	PID(K) > 4	-
π	PID(K) < 4	-
K, π	$\chi^2_{\text{IP}} > 9$	
D^0	$\chi^2_{\text{Vertex}} < 6$	-
D^0	$\chi^2_{\text{FD}} > 100$	-
D^0	DIRA > 0.99	-
D^0	1790 MeV/c^2 $< m_{K\pi} <$1940 MeV/c^2	1810 MeV/c^2 $< m_{K\pi} <$1940 MeV/c^2
μ	$\chi^2_{\text{Track}} < 4$	-
μ	$p_{\text{ghost}} < 0.5$	-
μ	$p > 3$ GeV/c	-
μ	$p_T > 1.2$ GeV/c	-
μ	PID(μ)> 1.2	-
μ	$\chi^2_{\text{IP}} > 9$	-
B	$\chi^2_{\text{Vertex}} < 6$	-
B	DIRA > 0.99	-
B	2.5 GeV/c^2 $< m_{D^0\mu} < 6$ GeV/c^2	-

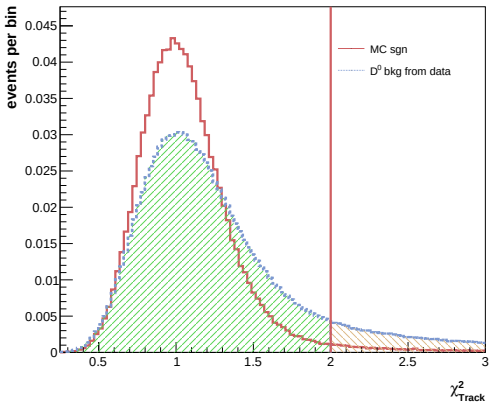
Table 9: The selection requirements are differentiated between *stripping* requirements and requirements imposed on top of the stripping.

4.4.1 Kaon and Pion Selection Requirements

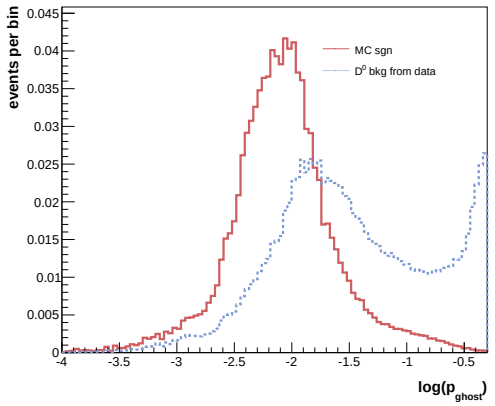
The kaon and pion play a similar role in the analysis. Most cuts can be performed equally on both particle candidates. First, a track quality threshold is set. They are constrained to $\chi^2_{\text{Track}} < 2$ (Fig. 11a and Fig. 11b) and a $p_{\text{ghost}} < 0.5$ (Fig. 11c and 11d). The distributions differ only slightly in the particle type.



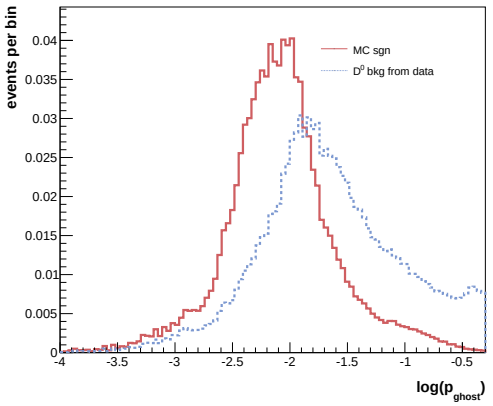
(a) K candidate χ^2_{Track}



(b) π candidate χ^2_{Track}



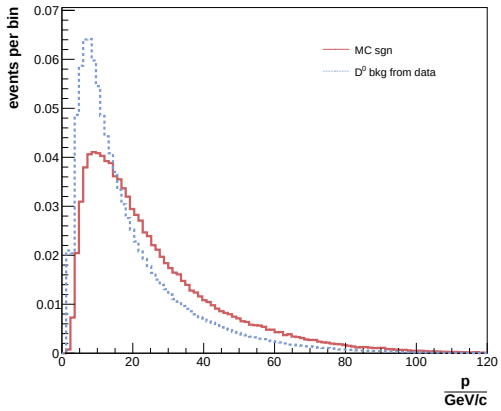
(c) K candidate p_{ghost}



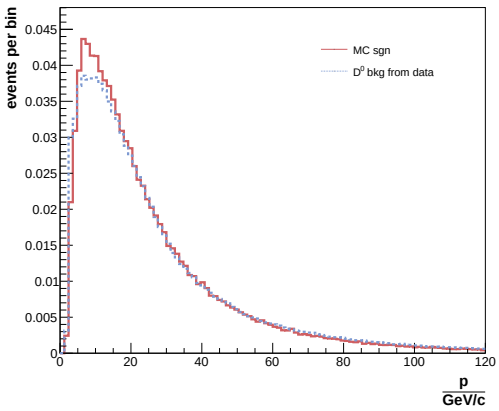
(d) π candidate p_{ghost}

Figure 11: The track quality of the kaon and pion candidate is expressed in χ^2_{Track} for the kaon (a) and pion (b) and in p_{ghost} for the kaon (c) and pion (d). The cut on top of the stripping is shown as a vertical red line. The area green signifies *accepted*, the area in red *rejected*. The combinatorial sideband sample acquired from data is shown in dashed light blue, the simulated signal sample in solid red. All histograms are normalised to unity.

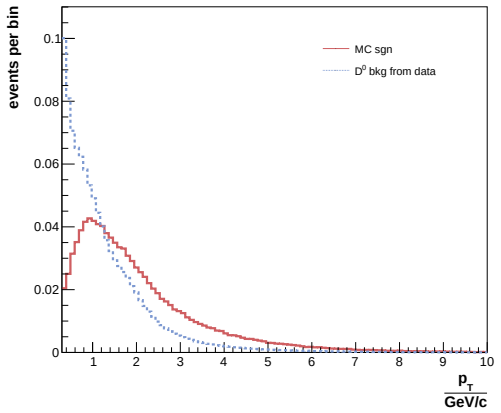
Second, we impose kinematical requirements on the kaon and pion. We require $p > 2$ GeV/c (Fig. 12a and Fig. 12b) and $p_T > 0.3$ GeV/c (Fig. 12c and Fig. 12d) for both candidates, as well as the transverse momentum sum $p_T(K) + p_T(\pi) > 1.4$ GeV/c.



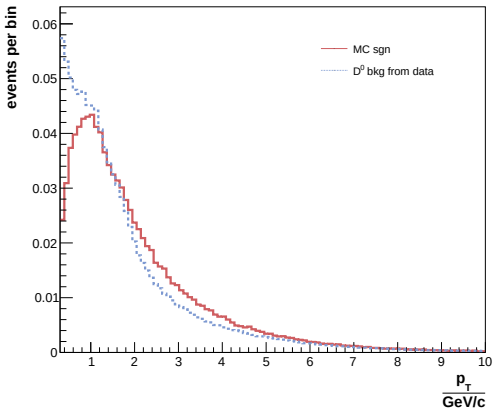
(a) K candidate p



(b) π candidate p



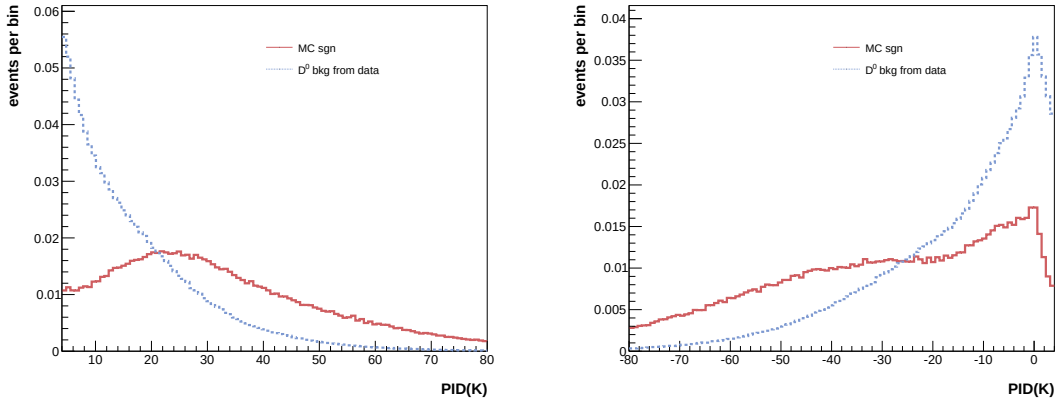
(c) K candidate p_T



(d) π candidate p_T

Figure 12: The kinematical constraints are imposed on the momentum p for the kaon (a) and pion (b) candidate and the transverse momentum p_T of the kaon (c) and pion (d) candidate. The combinatorial sideband sample acquired from data is shown in dashed light blue, the simulated signal sample in solid red. All histograms are normalised to unity.

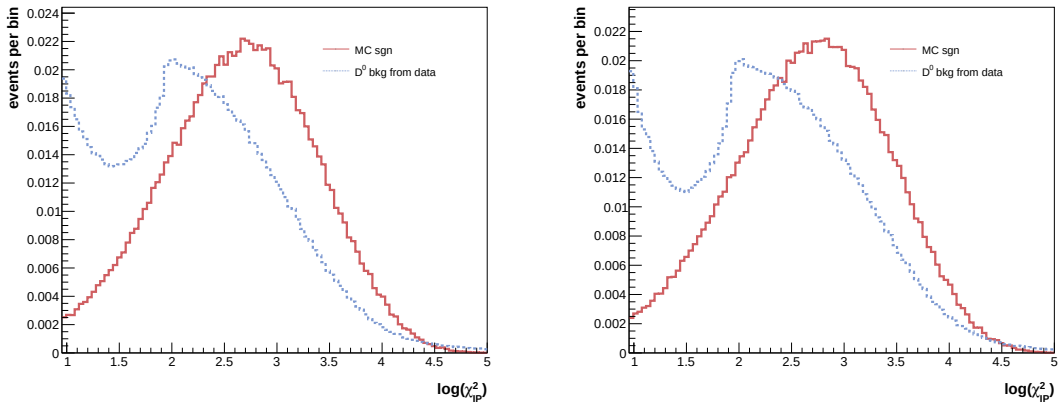
Third, the particle identification cuts are set. For the kaon candidate we require PID(K) > 4 (Fig. 13a), for the pion candidate PID(K) < 4 (Fig. 13b). Note the different orientation of signal and background for the kaon and pion (see definition of PID). The pion cut is less stringent than the kaon cut.



(a) K candidate PID(K) (b) π candidate PID(K)

Figure 13: The PID(K) cut of the kaon (a) is less stringent than for the pion (b). The combinatorial sideband sample acquired from data is shown in dashed light blue, the simulated signal sample in solid red.

Finally, we consider the IP which is a very good discriminator against background. We require that $\chi^2_{\mathrm{IP}} > 9$ for both (Fig. 14).

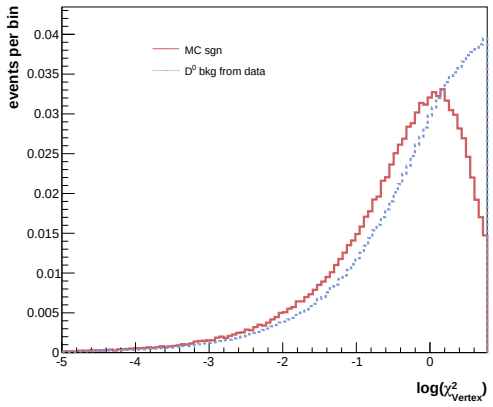


(a) K candidate χ^2_{IP} (b) π candidate χ^2_{IP}

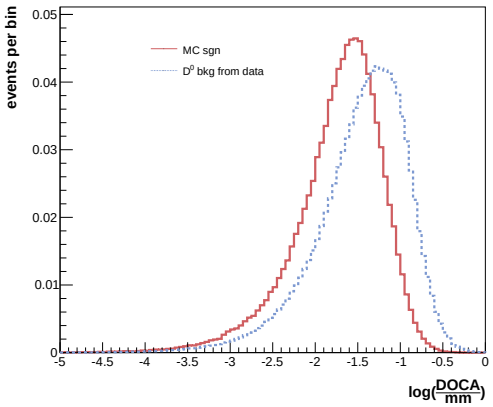
Figure 14: The χ^2_{IP} of the kaon (a) and pion is a good discriminator against prompt background. Note the steep rise of the background toward lower IP. The combinatorial sideband sample acquired from data is shown in dashed light blue, the simulated signal sample in solid red. All histograms are normalised to unity.

4.4.2 D^0 Selection Requirements

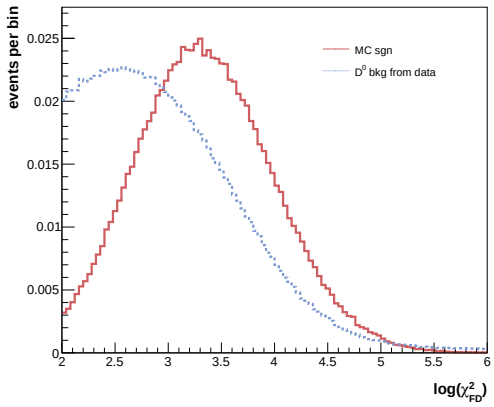
Now, that we have put a handle on the identification and reconstruction of high p_T kaons and pions, we examine the combined particle D^0. We require a well reconstructed kaon-pion vertex with $\chi^2_{\text{Vertex}} < 6$ (Fig. 15a). No requirement is set for the distance of closest approach (Fig. 15b). Additionally, a D^0 candidate should have $\chi^2_{\text{FD}} > 100$ (Fig. 15c) and a DIRA > 0.99 (Fig. 15d). Last but not least, we constrain the D^0 candidate's mass to the region $1810\ \text{MeV/c}^2 < m_{K\pi} < 1940\ \text{MeV/c}^2$ as shown in Fig. 10.



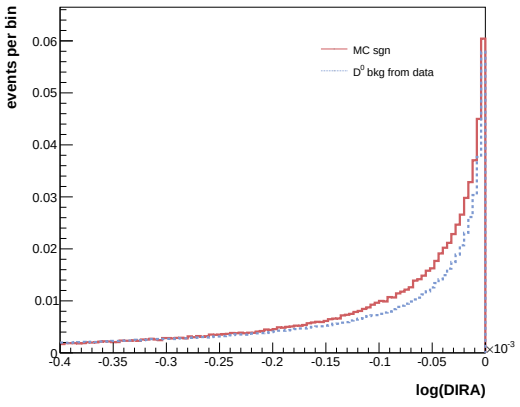
(a) D^0 candidate χ^2_{Vertex}



(b) D^0 candidate DOCA



(c) D^0 candidate χ^2_{FD}

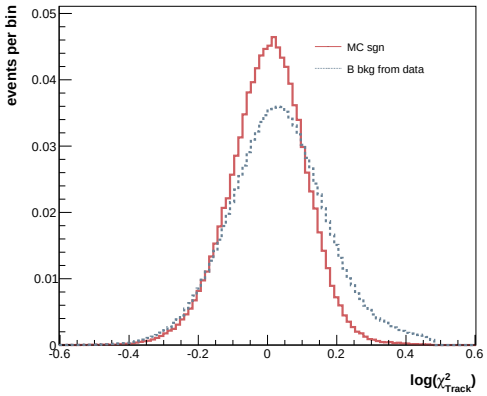


(d) D^0 candidate DIRA

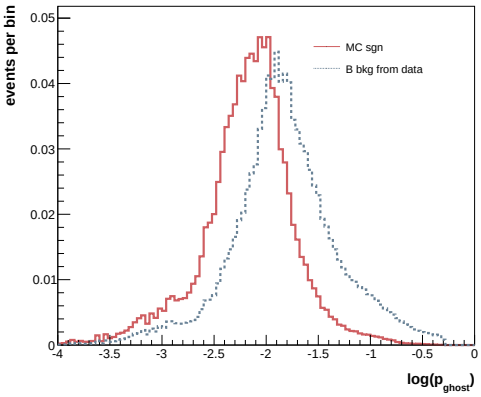
Figure 15: The selection requirements on the D^0 candidate are applied on the χ^2_{Vertex} (a), DOCA (b), χ^2_{FD} (c) and the direction angle DIRA (d). The combinatorial sideband sample acquired from data is shown in dashed light blue, the simulated signal sample in solid red. All histograms are normalised to unity.

4.4.3 Muon Selection Requirements

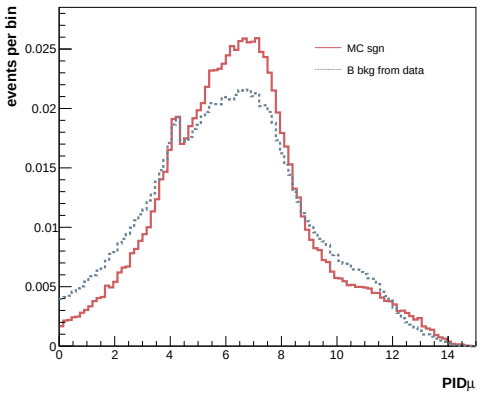
As stated above, there is no muon triggered prompt sample. Instead, a background sample is built that should be suitable for a qualitative prompt background characterization. The sample is generated from data with the requirement D^0 DIRA > 0.999 and D^0 $\chi^2_{\text{IP}} < 10$.
The muon candidate is constrained to $\chi^2_{\text{Track}} < 4$ (Fig. 16a), $p_{\text{ghost}} < 0.5$ (Fig. 16b) and PID(μ) > 1.2 (Fig. 16c). As with the D^0 decay products, the requirement $\chi^2_{\text{IP}} > 9$ (Fig. 16d) is important to reduce background.



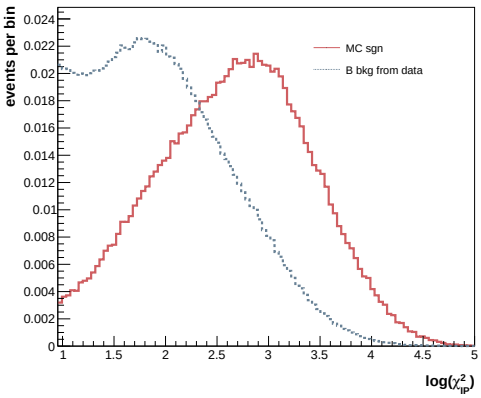
(a) μ candidate χ^2_{Track}



(b) μ candidate p_{ghost}



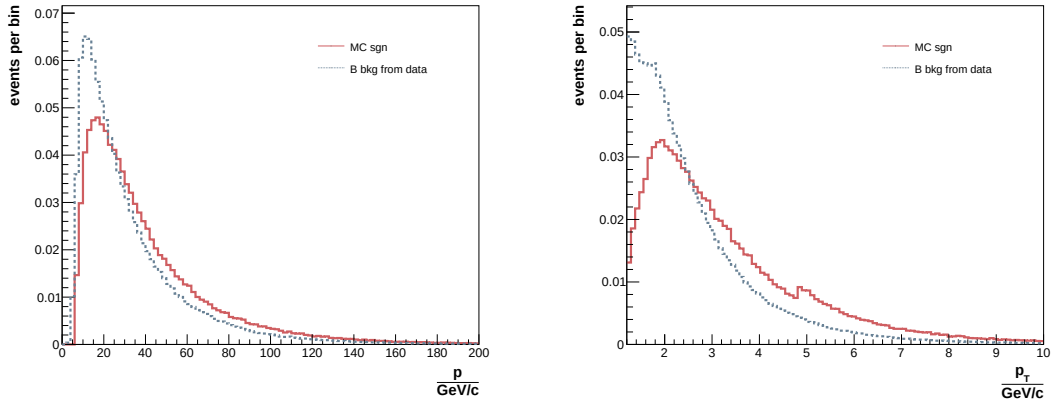
(c) μ candidate PID(μ)



(d) μ candidate χ^2_{IP}

Figure 16: The muon requirements on χ^2_{Track} (a), p_{ghost} (b), PID(μ) (c) and χ^2_{IP} are useful to reduce prompt background. Requirements are made on PID(μ) (c) and the IP χ^2 (d) for the muon. Note the large prompt background towards lower IP on the left of (d). The prompt simulated sample is shown in dashed dark blue, the simulated signal sample in solid red. All histograms are normalised to unity.

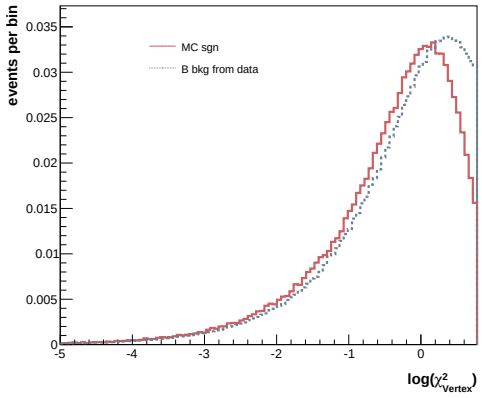
We impose the kinematical requirements $p > 3$ GeV/c and $p_T > 1.2$ GeV/c on the μ.



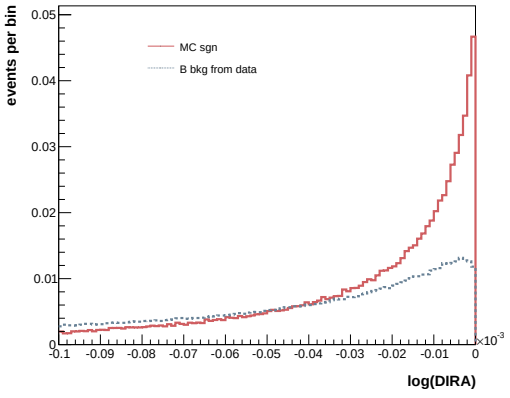
(a) μ candidate p (b) μ candidate p_T

Figure 17: The kinematical constraints for the muon are imposed on p (a) and p_T (b). The prompt simulated sample is shown in dashed dark blue, the simulated signal sample in solid red. All histograms are normalised to unity.

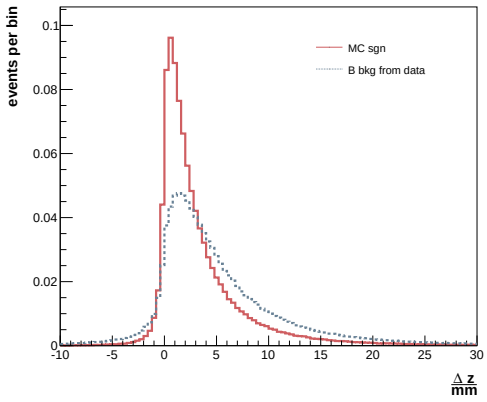
The B candidate is constrained to have $\chi^2_{\text{Vertex}} < 6$ (Fig. 18a) and DIRA> 0.999 (Fig. 18b). No cut is performed on $\Delta z = z(D) - z(B)$ (Fig. 18c), where z is the decay vertex z-coordinate. Since the muon neutrino cannot be detected, the invariant mass of the B candidate (Fig. 18d) is shifted towards lower values with respect to the the PDG mass. The mass is constraint to 2.5 GeV/c^2 $< m_{D^0\mu} <$ 6 GeV/c^2.



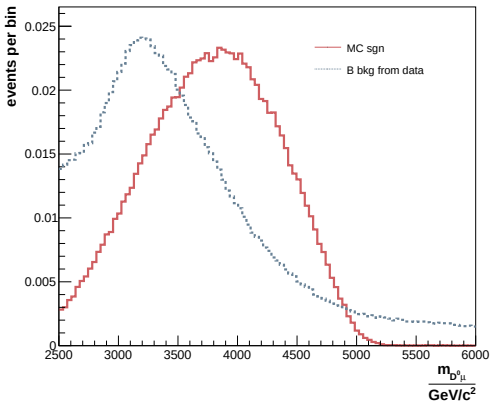
(a) B candidate χ^2_{Vertex}



(b) B candidate DIRA



(c) B candidate Δz



(d) B candidate mass $m_{D^0\mu}$

Figure 18: χ^2_{Vertex} (a), DIRA (b), Δz (c) and $m_{D^0\pi}$ are constrained in the stripping line. The prompt simulated sample is shown in dashed dark blue, the simulated signal sample in solid red. All histograms are normalised to unity.

4.5 Signal Yield Determination

4.5.1 Fit of the Invariant D^0 Candidate Mass

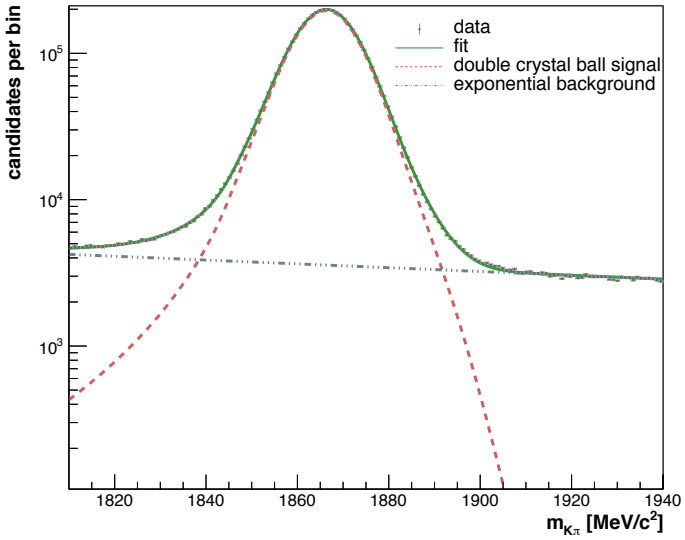
The mass peak is fitted using two Crystal Ball functions [34] as a signal. The advantage of a Crystal Ball over a pure Gaussian is the description of initial- or final state radiation. The Crystal Ball function is Gaussian in the center area and features a one-sided exponential tail. The Crystal Ball function is defined as

$$f(x;\alpha,n,\overline{x},\sigma) = N \cdot \begin{cases} \exp\big(-\frac{(x-\overline{x})^2}{2\sigma^2}\big) & \text{if } \frac{x-\overline{x}}{\sigma} > -\alpha \\ A \cdot \big(B - \frac{x-\overline{x}}{\sigma}\big)^{-n} & \text{if } \frac{x-\overline{x}}{\sigma} \leq -\alpha \end{cases}, \tag{12}$$

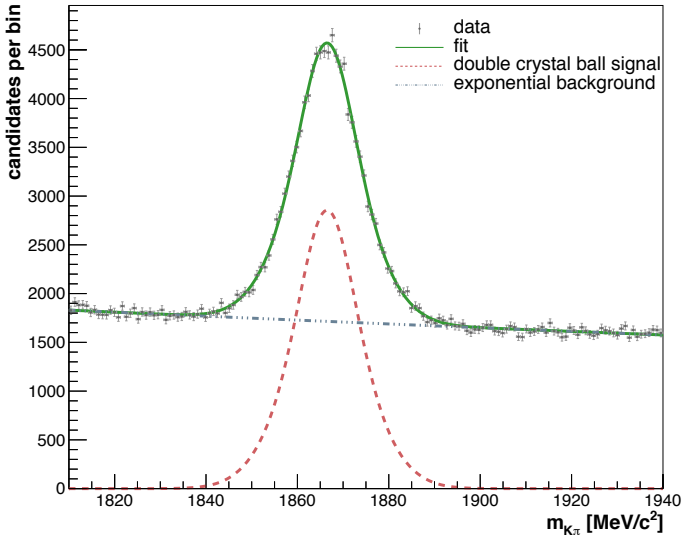
where N is a normalization factor. The parameters $\overline{x}$ and σ define the shape of the Gaussian kernel. The constant α defines the transition point from the Gaussian- to the exponential part. The parameter n dictates the slope of the tail. The constants A and B depend on α and n. For the two Crystal Ball functions, we take one α positive and one *negative* to ensure two-sided tails. The combinatorial background is described by an exponential

$$f(x,\beta) = N \cdot \exp\big(-\beta x\big), \tag{13}$$

where β is the exponent and N the normalization constant.
The extended maximum likelihood fit has been performed within the ROOT framework v5.34.21 [35] using RooFit v3.60. The fit converges nicely and has been done both for the WS and the RS data sample. The number of free parameters is 12: eight from the two crystal balls, one from the exponential and three composition parameters. The initial values have been adjusted with simulation. The resulting fit is shown in Fig. 19a and Fig. 19b. For the RS fit we get $\chi^2_{fit} \approx 1.44$, for the WS $\chi^2_{fit} \approx 0.97$. We get a signal yield of $N_{D^0} = (43851 \pm 3) \times 10^2$ events and a background of $N_{\text{comb}} = (5256 \pm 2) \times 10^2$ events. The signal-to-background ratio of $\frac{S}{B} \approx 8.3$ is beneficial for the measurement.



(a) RS D^0 candidate mass fit



(b) WS D^0 candidate mass fit

Figure 19: RS (a) and WS (b) D^0 candidate mass fit in the range 1810 MeV/c^2 $< m_{K\pi} <$ 1940 MeV/c^2. Since in the WS sample the mass peak is smaller compared to background, the WS fit is better shown on a linear scale. The data is shown as grey points with error bars, the simulated signal as a dashed red line, the exponential combinatorial background as a dashed-dotted light blue line and the sum of the two as a solid green line.

4.5.2 Fit of the D^0 Candidate Impact Parameter

The IP fitting range has been defined to range from 10^{-3} mm to 10 mm.
The histogrammed distributions of section 4.3 are used as probability density functions. Figure 20 shows the various IP fit templates for the RS, each normalised to unity. Note the position of the combinatorial D^0 background. The overlap of the combinatorial and dfB templates is explained by the shaping of the stripping line and the trigger.

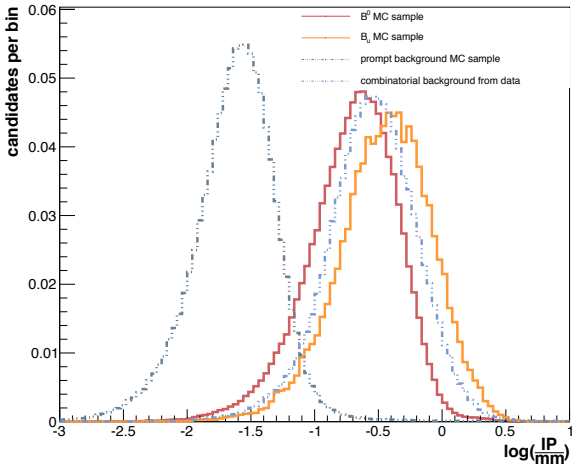
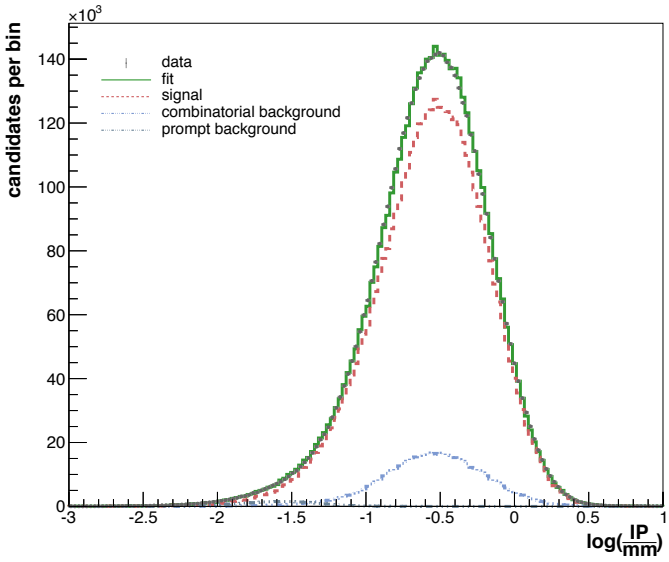


Figure 20: D^0 IP template shapes. The B^0 simulated signal is shown as a solid red line, B_u in solid orange, the prompt background in a dashed dark blue, and the combinatorial background in a dashed light blue. Each template is normalised to unity.

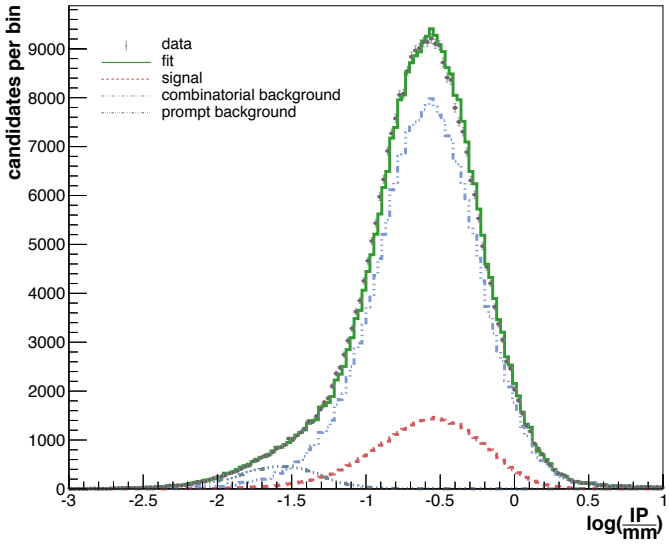
The mass fit background yield fixes the combinatorial D^0 background, the D^0 mass peak fit yield is used to fix the $dfB + prompt$ component in the IP fit. The number of free parameters in the IP fit is therefore just two: the fraction of dfB to prompt D^0 and the fraction of the $B^{\pm}$ to $B^0/\overline{B}^0$ templates. The fits are shown in Fig. 21a and Fig. 21b. Note, that for both RS and WS the same $D^0 \to K\pi$ background template is used. Additionally, the same RS signal template is used due to limited statistics in the WS sample. The sum of the templates is not a perfect description of the data since

- the composition of the signal is not purely $B^{\pm}$ and $B^0/\overline{B}^0$.
- the D^0 in the prompt background does not have a simulated mother particle. Selection requirements on the mother could change the kinematics of the D^0 prompt sample.

Although the fit is in principle sensitive to the ratio of $B^{\pm}$ and $B^0/\overline{B}^0$, the value cannot be trusted for these reasons. Furthermore, the χ^2_{fit} of 11.7 (RS) and 3.0 (WS) is unreasonably even though the fit *looks* good. Since the template do not describe the data to satisfaction, one can get huge χ^2_{fit} if a large amount of statistics is present. The χ^2_{fit} is regarded a unsatisfactory fit quality descriptor in this case. Instead, the uncertainty in the choice of template is considered in the uncertainty calculation.
The relative signal yield $N_{\mathrm{WS}}/N_{\mathrm{RS}}$ is $\approx$ 1.2% and is of the same order as expected from D^0 branching fractions and mis-ID probabilities $(1.4 \pm 0.3)\%$. Note, the WS result is only used as a cross-check and has no part in calculating the cross-section.



(a) RS D^0 candidate IP fit



(b) WS D^0 candidate IP fit

Figure 21: The IP fit is shown on a logarithmic x-scale. Note the much smaller signal contribution in the WS as expected. The data is shown as grey points with small error bars, the signal in dashed red, the prompt background in dashed dark blue and the combinatorial background in dashed-dotted light blue. The total fit is shown as a solid green line.

4.5.3 Additional Background Subtraction

Misidentification Background This section describes how the D^0 mis-ID background is subtracted. For each PID cut in a particle hypothesis, we can commit a type I error (rejection of a true null hypothesis) that results in a PID efficiency. Additionally, there is a probability of a type II error $p(X|Y)$ (erroneously accepting the false null hypothesis). Both quantities depend on both p_x, η_x and slightly on the number of tracks per event. Following Tab. 5 we can calculate the mis-ID background fraction to

$$\frac{N_{\text{mis-ID}}}{N_{\text{signal}}} = \frac{p(\pi|K)\cdot\mathcal{B}(K^+K^-) + p(K|\pi)\cdot\mathcal{B}(\pi^+\pi^-) + p(K|\pi)p(\pi|K)\cdot\mathcal{B}(K^+\pi^-)}{\mathcal{B}(K^-\pi^+)} \tag{14}$$

The branching ratios and their uncertainties are taken from PDG [1]. The type II error probabilities are retrieved from independent studies [22]. Obviously, the efficiencies and error probabilities depend on the PID cut under consideration. The pion cut is less stringent, resulting in a higher mis-ID probability. We get, averaged over the p, η and N_{Tracks} of the simulated signal distributions (Tab. 10):

particle hypothesis	probability
$p(\pi\|K)$	$(3.98 \pm 1.12)\%$
$p(K\|\pi)$	$(8.56 \pm 1.54)\%$

Table 10: The mis-ID probabilities of the kaon and pion are averaged over the simulated signal distribution of p, η and N_{Tracks}.

The type II error events are subtracted from the fit yield using simulation for the p_x, η_x and N_{Tracks} distribution. One averages over p_x, η_x and N_{Tracks} and uses the standard deviation as the uncertainty. We get to a mis-ID fraction $(7.2 \pm 0.9)\cdot 10^{-3}$ of the IP fit yield.

Background from $b \to D^0\tau^-(\to \mu^0\bar{\nu}_\mu\nu_\tau)X$ Tau leptons decay readily to muons in flight, resulting in a small contribution to the signal. To retrieve the relative tau contribution, we use simulation by imposing a tau as the muon mother. This background amounts to $(1.37 \pm 0.04)\%$ of the IP fit yield and is subtracted.

4.6 Efficiency

As introduced in section 4.2.1, we divide the signal events into D^0-like and D^*-like. Each efficiency is calculated separately for these two cases. They are combined by the fraction obtained from PDG shown in Tab. 7. The uncertainty of the D^0/D^*-like fraction is included in the uncertainty calculation. We identify five efficiencies that are explained below. The calculated efficiencies are listed in Tab. 11.

type	D^0-like [%]	D^*-like [%]	combined [%]
ε_{acc}	35.0 ± 0.3	35.0 ± 0.3	35.0 ± 0.3
ε_{cut}	94.6 ± 0.1	94.2 ± 0.1	94.2 ± 0.1
ε_{sel}	5.74 ± 0.02	5.18 ± 0.01	5.26 ± 0.01
$\varepsilon_{\text{corr}}$	102.2 ± 0.2	102.2 ± 0.2	102.2 ± 0.2
ε_{PID}	83.32 ± 0.02	84.40 ± 0.02	84.26. ± 0.02
$\varepsilon_{\text{trig}}$	70.0 ± 0.2	68.5 ± 0.1	68.7 ± 0.1
ε_{tot}	2.27 ± 0.01	2.02 ± 0.01	2.06 ± 0.01

Table 11: The efficiencies for D^0-like and D^*-like differ only slightly. Note, that the D^0-like efficiency is generally higher. The efficiencies and efficiency uncertainties are combined by the factors in Tab. 7. The total efficiency uncertainty is calculated using error propagation by adding the relative efficiencies in quadrature.

(i) **Acceptance Efficiency** The acceptance efficiency takes the finite angular acceptance of the LHCb detector into account. The acceptance efficiency ε_{acc} is determined as the ratio of the signal events with a kaon, pion and muon inside the LHCb acceptance of 10 mrad $< \theta <$ 250 mrad to the total number of signal events. This number is obtained from simulation, where a random b hadron is produced with $p_T > 0$. The same sample is used for D^*-like and $D^0$0-like. The efficiency is calculated as $\varepsilon_{\text{acc}} = \frac{N_{\text{acc}}}{N_{\text{tot}}}$. The uncertainties are calculated using binomial errors $\delta\varepsilon = \sqrt{\varepsilon(1-\varepsilon)/N_{\text{tot}}}$.

(ii) **Selection Efficiency** The selection efficiency is a product of detection efficiency (tracking efficiency) and stripping efficiency. It does not include the PID efficiency that is determined separately. We use simulation to determine ε_{sel} in the same manner as ε_{acc}. The events are generated within the LHCb region. This efficiency can be rather low, especially in the border regions of LHCb.

(iii) **Tracking Efficiency Correction**: The selection efficiency includes the tracking efficiency and has been determined by simulation. To take possible differences in the tracking efficiency between simulation and data into account, a data-driven correction is applied. This efficiency ratio has been determined in an independent study using a pure $J/\psi \to \mu^+\mu^-$ sample. A method called *tag-and-probe* is used. The probe muon is reconstructed using information from TT and the muon system. The efficiency is determined as the fraction of events, where this muon-TT track can be matched to a track reconstructed by the VELO and the T-stations. One then compares the number of successfully reconstructed J/ψ events to the true number of J/ψ obtained from a fit of the invariant tag-probe mass. The tracking efficiency is governed by two observables and is hence defined as a function of the probe track momentum p and the pseudo-rapidity η. It is mostly independent of the polar angle ϕ and the track multiplicity N_{Track}. The data is compared to the simulated simulated efficiency. The correction amounts to $\varepsilon_{\text{corr}} = \frac{\varepsilon_{\text{track, data}}}{\varepsilon_{\text{track, sim.}}}$. The method is illustrated in Fig. 22. Table 12 shows the tracking efficiency correction as a function of η and p [36].

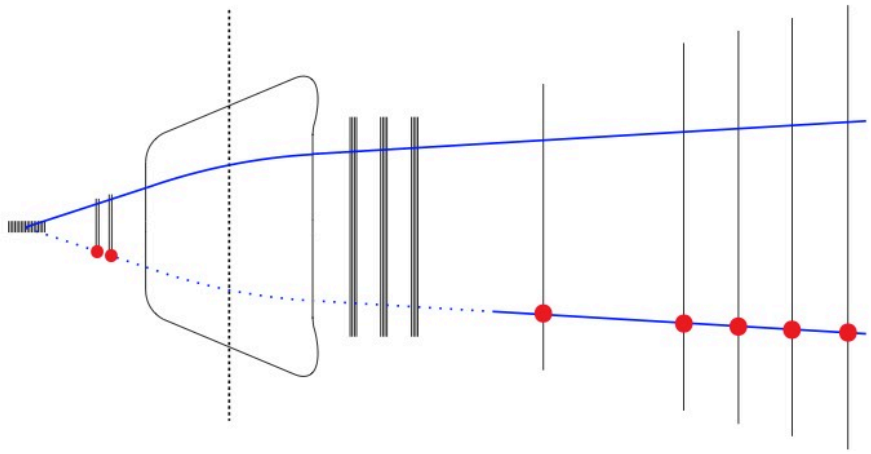


Figure 22: Illustration of the tag-and-probe method [36]. The detector parts are (from left to right): VELO, two TT stations, magnet coil, three T stations, five muon stations. The upper solid blue line represents the tag track and is fully reconstructed. The dotted bottom track is the probe. Required hits are shown as red dots.

η \ p[GeV/c]	0 – 10	10 – 20	20 – 40	40 – 100	100 – 200
3.2 – 5	1.073 ± 0.031	1.003 ± 0.004	0.991 ± 0.002	0.987 ± 0.001	0.995 ± 0.003
1.9 – 3.2	1.042 ± 0.004	1.011 ± 0.001	1.007 ± 0.001	1.008 ± 0.002	1.024 ± 0.016

Table 12: The tracking efficiency is parametrised as a function of p and η. p is presented in units of GeV/c on the horizontal, η on the vertical. Some values are greater than one.

The efficiency correction is calculated by averaging over the p, η and N_{Track} of the simulated signal distribution. The uncertainty is obtained by throwing N random efficiencies within a gaussian of mean and standard deviation as in Tab. 12, weight it with the simulated signal distribution (normalised to unity) and take the standard deviation. N has been set to 1000 throws.

(iv) **Trigger Efficiency**: The trigger efficiency for the triggers used in this analysis is calculated using simulation $\varepsilon_{\text{trig}} = \frac{N_{\text{trig}}}{N_{\text{tot}}}$. The uncertainty is taken to be binomial.

(v) **PID Efficiency**: The capability of particle identification is accounted for with this efficiency. The PID calibration is performed on a set of control channels with particle IDs reconstructed without any RICH information, using only kinematic selections of exclusive decays[1] [22]. The PID governed by the observables p, η and N_{Track}. Figure 23 shows the PID efficiency as a function of momentum for the kaon. The average values integrated over the simulated momentum and pseudo-rapidity distributions as well as N_{Track} and their uncertainties are shown in Tab. 13. The uncertainty is calculated by throwing random numbers as for the tracking efficiency correction.

[1]The mis-ID background probability is obtained in a similar way. P.ex. instead of requiring $DLL(K) > 0$ to get the efficiency we require $-\text{DLL}(K) < 0$ to get the mis-ID probability

particle	D^0-like[%]	D^*-like[%]
K	92.933 ± 0.009	93.617 ± 0.009
π	95.691 ± 0.006	95.759 ± 0.006
μ	94.210 ± 0.018	93.742 ± 0.017
combined	84.396 ± 0.018	83.421 ± 0.017

Table 13: The PID efficiency is determined for the kaon, pion and muon separately. The multiplication of these efficiencies is shown in *combined*. The efficiencies are weighted by the normalised distribution obtained from simulation for each bin and summed up.

(vi) **Cut Efficiency**: This efficiency summarises all cuts on top of the stripping. The major player is the tighter mass constraint. Simulation is used to calculate the efficiency. The efficiency is calculated as the ratio of events before and after the cut $\varepsilon_{cut} = \frac{N_{\text{cut}}}{N_{\text{tot}}}$. The uncertainty is taken to be binomial.

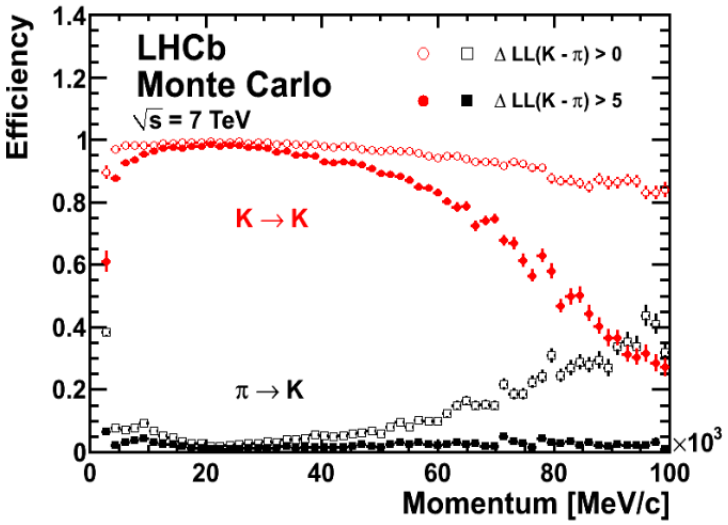


Figure 23: PID efficiency as a function of p [22]. Two different PID cuts (ΔLL) are presented. $K \to K$ signifies the efficiency, $\pi \to K$ the misidentification probability.

4.7 Cross-Section Calculation

The final cross-section result is calculated as a weighted mean of the two polarity results with Eqn. 10. In the scope of this calculation, the cross-section for each polarity i is denoted by μ^i. The systematical uncertainty common to both polarities i.e. luminosity and branching fraction uncertainty (Sec. 4.8) is denoted by σ^{glob}. For each polarity, the statistical uncertainty is denoted by σ_i^{stat}, the systematical uncertainty by σ_i^{sys}. The cross-section is then calculated as a weighted mean

$$\mu = \frac{\sum_i w_i \mu_i}{\sum_i w_i}, \tag{15}$$

with the weight $w_i = 1/\left[(\sigma_i^{\text{stat}})^2 + (\sigma_i^{\text{sys}})^2\right]$.

4.8 Uncertainty

Using the notation of the previous Sec. 4.7, the relative uncertainty for each polarity (Tab. 14) is calculated as

$$\left(\frac{\sigma_{\mathrm{i}}}{N}\right)^2 = (\sigma_i^{\text{stat, rel.}})^2 + (\sigma_i^{\text{sys, rel.}})^2 + (\sigma^{\text{glob, rel.}})^2, \tag{16}$$

where $(\sigma^{\text{sys}})^2 = \sum_j (\sigma^{\text{sys},j})^2$ and $(\sigma^{\text{glob}})^2 = \sum_j (\sigma^{\text{glob},j})^2$ for each uncertainty type j. Following the weighted average definition, we get a total uncertainty of

$$\sigma = \frac{\sum_i w_i \sigma_i}{\sum_i w_i} \tag{17}$$

The statistical error is calculated separately by

$$\left(\frac{1}{\sigma^{\text{stat}}}\right) = \sum_i \left(\frac{1}{\sigma_i^{\text{stat}}}\right) \tag{18}$$

uncertainty source	σ^{rel} magnet *down* [%]	σ^{rel} magnet *up* [%]	type
fit	0.01	0.01	σ^{stat}
signal template	1.3	1.3	σ^{sys}
prompt template	0.6	0.6	σ^{sys}
mis-ID	0.1	0.1	σ^{sys}
tau background	0.04	0.04	σ^{sys}
efficiency	0.6	0.6	σ^{sys}
signal composition	2.7	2.7	σ^{sys}
luminosity	1.2	1.2	σ^{glob}
B branching fraction	5.1	5.1	σ^{glob}
D^0 branching fraction	1.3	1.3	σ^{glob}
total	6.2	6.3	σ^{tot}

Table 14: Table of relative uncertainties. The relative uncertainties for each uncertainty type are added in quadrature. The fit uncertainty is treated separately as a statistical uncertainty.

The following list describes the statistical- and systematic uncertainty determination. These calculations are performed for each magnet polarity separately.

(i) **Fit Uncertainty** The only uncertainty denoted *statistical* is the uncertainty on the parameter values extracted from the fit. It corresponds to the parameter uncertainty obtained from the shape of the log likelihood function. Compared to the systematical uncertainties, this quantity is rather small.

(ii) **Signal Template Uncertainty** The two signal templates $B^{\pm}$ and $B^0/\overline{B}^0$ have a slightly different IP template shape (Fig. 20). The difference in the IP fit yield N, when only one of the two templates is used, is a good measure for this uncertainty (Eqn. 19, where $N_0 =$

$N[B^0/\overline{B}^0$ template] and $N_\pm = N[B^\pm$ template]). It is in the order of 1%.

$$\sigma^{\text{signal template}} = \left|N_\pm - N_0\right| \tag{19}$$

(i) **Prompt Template Uncertainty** As described in section 4.3, the D^0 prompt sample does not have a mother with a muon. The muon requirement might change the IP distribution, though. We define two extreme cases that feature a low IP or a high IP respectively. This can be achieved by selecting an appropriate (p_T, y) bin of D^0. The D^0 dependence on (p_T, y) is shown in Fig. 26. Similar to the B template uncertainty calculation, the difference in the extreme case of the IP fit yield is taken as the template uncertainty. The low IP sample is set to 8 GeV/c^2 $< p_T <$ 12 GeV/c^2, the high IP sample to $4 < y < 5$.

(ii) **Uncertainty on the Background Yield from Mis-ID** The uncertainty from mis-ID is calculated as described in Sec. 4.5.3.

(iii) **Uncertainty on the Background Yield from $b \to D^0\tau^-(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau X$** The uncertainty from the tau background is calculated using simulation. It is assumed to be binomial.

(iv) **Efficiency Uncertainty** The efficiency uncertainty is calculated by summing over the individual efficiency uncertainties in quadrature (Tab. 11). This is done separately for D^0-like and D^*-like events. The two uncertainties are weighted by the factor in Tab. 7 and summed (Eqn. 20).

$$\sigma^{\text{efficiency}} = f_{D^0}\sigma^{\text{efficiency}}_{D^0} + f_{D^*}\sigma^{\text{efficiency}}_{D^*} \tag{20}$$

(v) **Signal Composition Uncertainty** The efficiencies of D^0-like and D^*-like events do not differ significantly. The relative uncertainty is calculated by

$$\sigma^{\text{signal composition}} = \left|\frac{N}{\varepsilon_{D^0}} - \frac{N}{\varepsilon_{D^*}}\right|, \tag{21}$$

where N is the background corrected IP fit yield.

(vi) **Luminosity Uncertainty** The global luminosity uncertainty is determined experimentally and is described in Ref. [21].

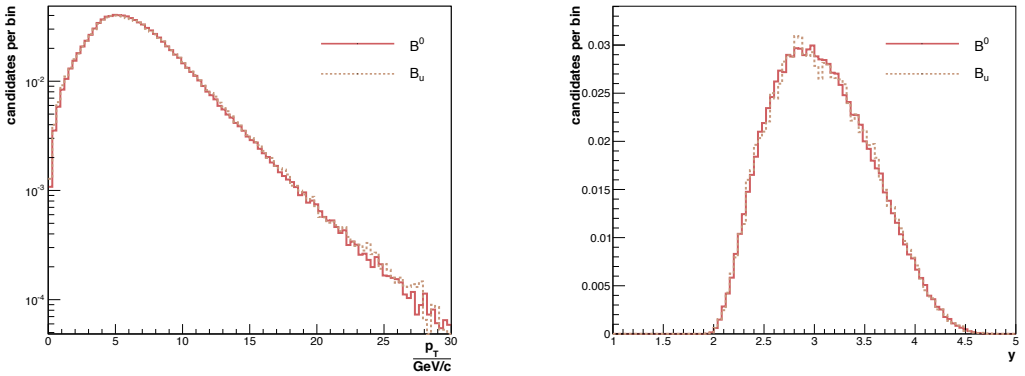
(vii) **Branching Fraction Uncertainty** and **Production Fraction Uncertainty** These global uncertainties are a major source of uncertainty and are retained from PDG [1].

4.9 Differential Analysis

The same procedure as for the global analysis can be applied to each bin individually. However, some additions have to be made as shown in this section.

4.9.1 Signal Yield Determination

The WS sample has fulfilled its role as a cross-check for the mis-ID background. We forego it entirely in the differential measurement.
The transverse momentum p_T and rapidity y of the B candidate fully describe the kinematical dependence of the cross-section. The p_T and y distribution of the B candidate are shown in Fig. 24.



(a) p_T distribution (b) y distribution

Figure 24: The B candidate p_T distribution (a) is presented on a logarithmic scale, the B candidate y distribution (b) on a linear scale. The samples are built from simulation and only include events that lie inside the LHCb acceptance.

The binning is chosen such that a large portion of the simulated signal is included. Thinly populated bins have been enlarged to ensure a reasonable fit convergence. The factor limiting the bin size is mostly the finite simulated signal template statistics. The bin edges have been chosen as follows:

$$p_T[\mathrm{GeV/c}^2] : 0, 2, 4, 6, 8, 12, 16, 26, \quad y : 1.5, 2.5, 3, 3.5, 4, 5 \tag{22}$$

The yield from the mass fit is shown in Fig. 15 and the yield from the IP fit is shown in Fig. 16. The mis-ID background and tau background is distributed more or less uniformly

p_T [GeV/c] \ y	1.5 – 2.5	2.5 – 3	3 – 3.5	3.5 – 4	4 – 5
0 – 2	7093 ± 121	39720 ± 79	47121 ± 32	30596 ± 126	8448 ± 574
2 – 4	46585 ± 542	250471 ± 83	299406 ± 651	181150 ± 96	45209 ± 875
4 – 6	107861 ± 1060	441097 ± 667	457490 ± 655	254274 ± 324	56523 ± 231
6 – 8	127297 ± 696	348655 ± 641	305778 ± 352	153050 ± 207	32193 ± 600
8 – 12	165960 ± 349	312605 ± 68	235989 ± 85	101399 ± 39	19320 ± 145
12 – 16	63985 ± 37	86517 ± 974	54281 ± 282	18735 ± 581	3393 ± 414
16 – 26	33248 ± 241	34337 ± 224	18238 ± 46	4895 ± 462	631 ± 23

Table 15: Differential mass fit yield in terms of p_T and y of the B candidate including the statistical fit uncertainty.

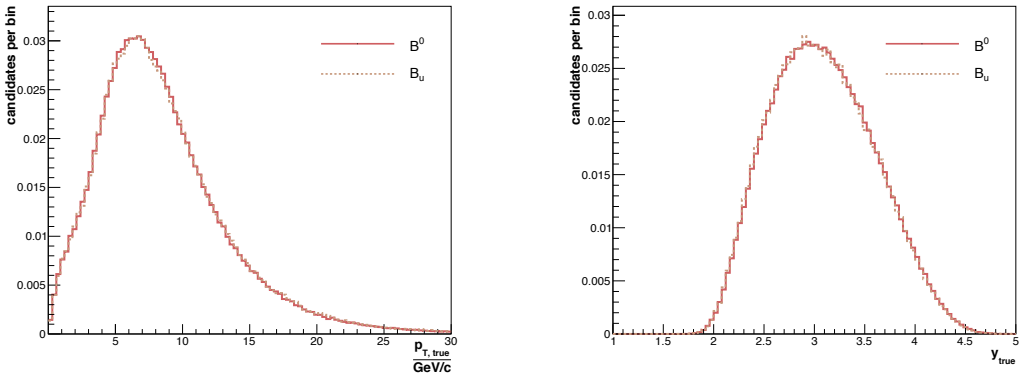
p_T [GeV/c] \ y	1.5 – 2.5	2.5 – 3	3 – 3.5	3.5 – 4	4 – 5
0 – 2	7007 ± 10	39390 ± 22	46164 ± 36	29747 ± 34	8244 ± 16
2 – 4	46294 ± 20	248629 ± 65	295182 ± 97	178744 ± 72	44612 ± 31
4 – 6	106919 ± 46	435397 ± 139	450087 ± 157	250913 ± 105	55859 ± 41
6 – 8	125724 ± 70	343402 ± 146	300325 ± 139	150981 ± 88	31727 ± 34
8 – 12	164222 ± 88	307925 ± 147	232501 ± 127	100120 ± 69	18993 ± 27
12 – 16	62824 ± 63	85110 ± 74	53602 ± 52	18426 ± 27	3331 ± 9
16 – 26	32626 ± 43	33610 ± 44	17775 ± 32	4724 ± 16	497 ± 12

Table 16: Differential IP fit yield in terms of p_T and y of the B candidate including the statistical fit uncertainty.

4.9.2 Migration

Cross-sections are usually computed in terms of $p_{T,\text{true}}$ and y_{true} of the B candidate which includes the momentum and energy of ν_μ. Since the measured (y, p_T) distribution does not include the missing neutrino energy, the obtained numbers have to be multiplied by the migration matrix $\mathcal{M}_{(\eta,p_T)\to(y_{\text{true}},p_{T,\text{true}})}$. In other words, we correct for the unmeasured neutrino energy to get the true kinematical B meson distribution. Since the binning region does not include the whole phase space, the migration matrix is not orthogonal.

We determine a migration matrix for the binning $(p_T, y) \to (p_{T,\text{true}}, y_{\text{true}})$ by simulation. We keep the same binning scheme as before. Figure 25 shows that the *true* variables have a similar distribution as in Fig. 24. The Poisson uncertainties $\sqrt{N}$ and are not used in the uncertainty calculation.



(a) $p_{T,\text{true}}$ distribution (b) y_{true} distribution

Figure 25: The B candidate $p_{T,\text{true}}$ distribution (a) and y_{true} distribution look similar to the distribution in Fig. 24. The samples are built from simulation and only include events that lie inside the LHCb acceptance.

The migration in y is rather small (in the order of a few percent), the migration in p_T is always to higher transverse momenta[1] and can be up to 40%. An example for a highly populated bin is shown in Fig. 17.

p_T [GeV/c] \ y	1.5 – 2.5	2.5 – 3	3 – 3.5	3.5 – 4	4 – 5
0 – 2	0	0	0	0	0
2 – 4	0	0.14 ± 0.37	1.83 ± 1.35	0.15 ± 0.39	0
4 – 6	0	3.44 ± 1.85	38.23 ± 6.18	3.43 ± 1.85	0
6 – 8	0	3.89 ± 1.97	31.10 ± 5.58	2.90 ± 1.70	0
8 – 12	0	1.50 ± 1.23	11.42 ± 3.38	1.27 ± 1.13	0
12 – 16	0	0.07 ± 0.27	0.57 ± 0.75	0.03 ± 0.18	0
16 – 26	0	0	0.01 ± 0.11	0.01 ± 0.08	0

Table 17: Example migration matrix for the bin $4 < \text{GeV/c}^2 < p_T < 6$ GeV/c^2, $3 < y < 3.5$ in [%].

4.9.3 Efficiency

The efficiencies are calculated the same way as in the global analysis. Since we exclusively use simulation to determine the efficiencies, we calculate them in terms of $(p_{T,\text{true}}, y_{\text{true}})$ and apply them *after* the migration took place. The overall efficiency is shown in Fig. 18. It is noteworthy that the efficiency increases towards higher transverse momentum and is highest in the center of the LHCb acceptance. The uncertainties are computed the same way as in the global analysis.

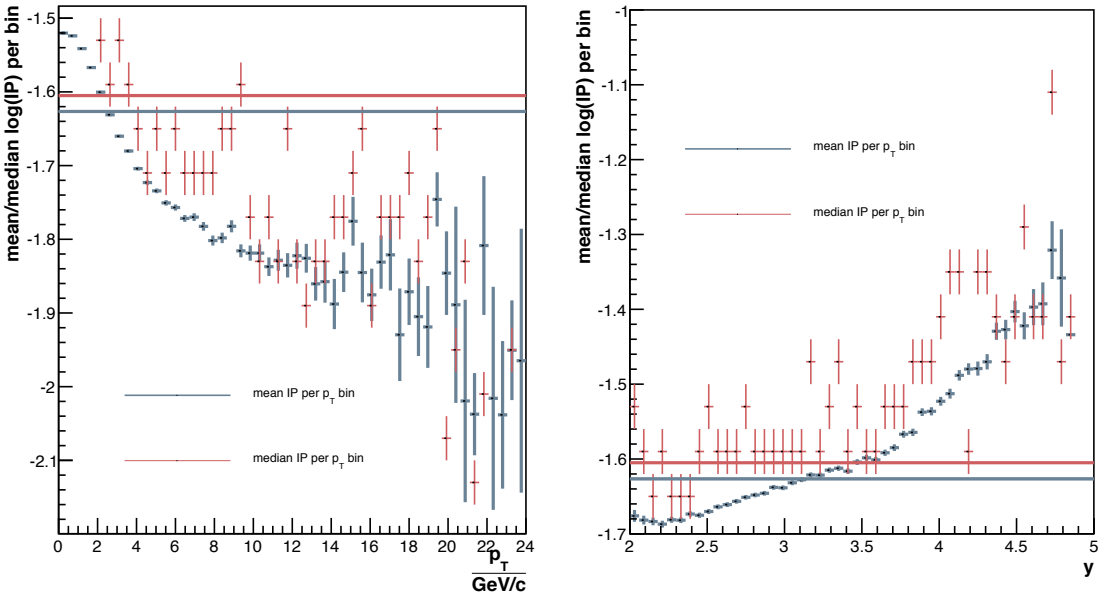
[1]The neutrino energy only increases the B meson energy. Migration to lower p_T can occur if the neutrino is produced in transverse direction.

p_T [GeV/c] \ y	1.5 – 2.5	2.5 – 3	3 – 3.5	3.5 – 4	4 – 5
0 – 2	0.031 ± 0.001	0.303 ± 0.013	0.421 ± 0.017	0.289 ± 0.015	0.043 ± 0.002
2 – 4	0.090 ± 0.002	0.907 ± 0.025	1.401 ± 0.028	1.080 ± 0.026	0.233 ± 0.009
4 – 6	0.308 ± 0.007	2.892 ± 0.052	3.933 ± 0.052	2.857 ± 0.071	0.602 ± 0.023
6 – 8	0.702 ± 0.017	5.887 ± 0.091	6.980 ± 0.128	5.156 ± 0.134	1.590 ± 0.053
8 – 12	1.591 ± 0.034	9.980 ± 0.124	11.022 ± 0.091	8.541 ± 0.114	2.602 ± 0.104
12 – 16	2.986 ± 0.090	13.958 ± 0.262	14.089 ± 0.099	11.585 ± 0.267	4.968 ± 0.110
16 – 26	4.386 ± 0.124	16.486 ± 0.261	15.583 ± 0.171	14.207 ± 0.240	6.251 ± 0.386

Table 18: Differential efficiency in terms of p_T and y of the B candidate including the systematical uncertainty in [%]. The D^0-like and D^*-like efficiencies have been combined. The uncertainty of the D^0-like fraction has not yet been considered.

4.9.4 Uncertainty

At thinly populated bins, the dominant uncertainty comes from the difference in the two magnet samples. The average magnet uncertainty is in the order of a few percent. Most bin uncertainties are dominated by the branching fraction uncertainty and the signal composition uncertainty. The statistical uncertainties are still very small compared to the systematical uncertainties.



(a) IP(p_T)

(b) IP(y)

Figure 26: p_T and y dependence of the D^0 background template. The solid lines represent the average.

5 Results

5.1 Prediction

The differential cross-section for $1.5 < y < 5$ has been calculated at leading order (LO) and next-to-leading order (NLO) in perturbation theory using MCFM [37]. The difference between leading order $\sigma_{LO} \approx 29$ µb and next-to-leading order $\sigma_{NLO} = 87.2^{+55.4}_{-33.7} \pm 11.1$ µb is substantial. Figure 27 shows the LO prediction and NLO prediction binned in p_T of the B candidate. The differences are significant. Therefore, the theoretical uncertainty is very large being in the order of 30% to 40%. This fact stresses the importance of a measurement to improve the sensitivity.
It is noteworthy, that not only is the integrated cross-section off, but there is also different p_T behaviour of the cross-section shape. The LO distribution looks as if it were rotated with lower cross-section at low p_T and higher at high p_T.
The uncertainty denoted 'theory' relates to $\sigma^f_{ab}(\xi_{a'}, \xi_{b'}, p, \mu_F^2, \mu_R^2)$, the PDF uncertainty relates to $f_a(\xi_a, \mu_F^2)$ (see Sec. 2.2.3). The uncertainties have been calculated by varying the renormalization- and factorization scale by a factor of two. The default scale is the b quark mass.

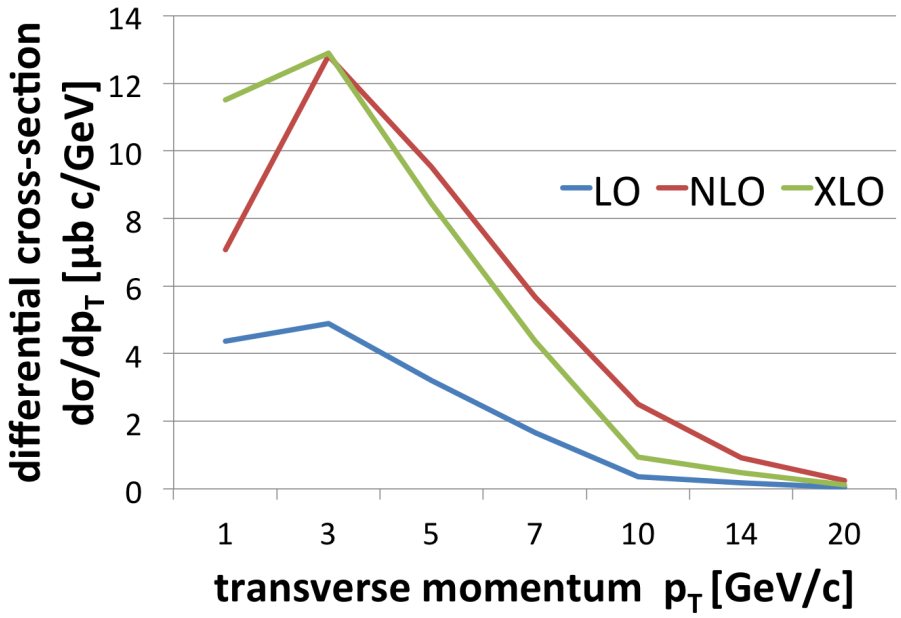


Figure 27: The LO prediction is shown in blue (bottom curve), the NLO prediction in red (top curve). XLO shows the LO prediction scaled to the total cross-section of NLO. The differential cross-section is shown in terms of p_T.

5.2 Global Analysis

The results for the two magnet polarities are shown in Tab. 19. We quote a cross-section of 77.39 ± 0.01 (stat.)± 7.06 (sys.)µb in the LHCb forward region. The result lies well within the error bars of the prediction $87.2^{+55.4}_{-33.7} \pm 11.1$µb.
Extrapolating the cross-section to the full phase space by scaling it by a factor 3.6 given by simulation, we get $278.33 \pm 0.04 \pm 25.24$µb.
Compared to the 7 TeV measurement of $74.9 \pm 5.3 \pm 12.9$µb [2] this corresponds to an expected increase. However, due to the different extrapolation factor of 3.77 used for the 7 TeV analysis,

they get a higher extrapolated cross-section of $282 \pm 20 \pm 49\mu b$. In the interest of time, this contradiction has not been examined.

magnet polarity	cross-section result $\sigma(pp \to b\bar{b}X)$
up	$78.29 \pm 0.01 \pm 4.88\mu b$
down	$77.39 \pm 0.01 \pm 4.93\mu b$

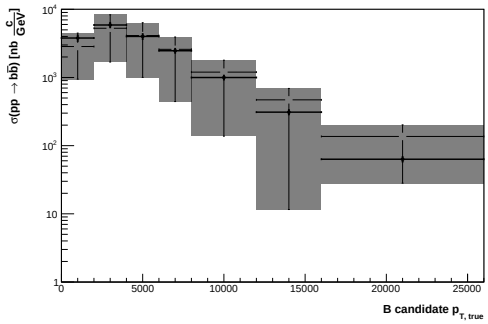
Table 19: Cross-section result for the two magnet polarities. The first uncertainty is statistical, while the second uncertainty is systematical.

5.3 Differential Analysis

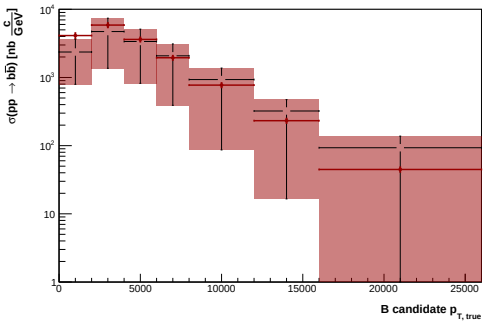
The result is binned and divided by its p_T and y bin size in GeV/c. The polarity combined cross-section in units of nb $\frac{c}{GeV}$ is given in Fig. 20. Comparing each y bin with the prediction, we get Fig. 28a to Fig. 28e. Theory and experiment agree within their uncertainties. Note, that the same tilt is visible as in the comparison LO to NLO.

p_T [GeV/c] \ y	1.5 – 2.5	2.5 – 3	3 – 3.5	3.5 – 4	4 – 5
0 – 2	3793.25	4124.36	3713.35	3362.99	3315.05
	±2.633 (0.07%)	±1.678 (0.04%)	±1.918 (0.05%)	±2.435 (0.07%)	±3.946 (0.12%)
	±379.708 (10%)	±382.375 (9.3%)	±369.644 (10%)	±425.922 (12.7%)	±488.849 (14.7%)
2 – 4	5891.09	5875.86	4688.03	3753.11	2473.74
	±1.794 (0.03%)	±1.33 (0.02%)	±1.233 (0.03%)	±1.241 (0.03%)	±1.252 (0.05%)
	±482.463 (8.2%)	±467.029 (7.9%)	±372.632 (7.9%)	±351.912 (9.4%)	±217.973 (8.8%)
4 – 6	4004.58	3617.54	3004.7	2450.47	1489.9
	±1.138 (0.03%)	±0.823 (0.02%)	±0.743 (0.02%)	±0.728 (0.03%)	±0.688 (0.05%)
	±357.18 (8.9%)	±281.268 (7.8%)	±234.813 (7.8%)	±197.098 (8%)	±125.771 (8.4%)
6 – 8	2469.01	1948.03	1635.58	1230.19	450.486
	±0.84 (0.03%)	±0.519 (0.03%)	±0.459 (0.03%)	±0.423 (0.03%)	±0.254 (0.06%)
	±181.943 (7.4%)	±148.525 (7.6%)	±130.51 (8%)	±99.264 (8.1%)	±41.439 (9.2%)
8 – 12	999.11	773.076	604.977	377.625	128.929
	±0.371 (0.04%)	±0.24 (0.03%)	±0.202 (0.03%)	±0.157 (0.04%)	±0.098 (0.08%)
	±73.268 (7.3%)	±61.014 (7.9%)	±46.581 (7.7%)	±30.847 (8.2%)	±11.71 (9.1%)
12 – 16	310.119	232.393	166.021	82.891	18.006
	±0.142 (0.05%)	±0.091 (0.04%)	±0.071 (0.04%)	±0.049 (0.06%)	±0.021 (0.12%)
	±24.548 (7.9%)	±18.437 (7.9%)	±12.733 (7.7%)	±6.688 (8.1%)	±1.425 (7.9%)
16 – 26	63.212	44.732	28.667	10.122	1.652
	±0.044 (0.07%)	±0.027 (0.06%)	±0.021 (0.07%)	±0.012 (0.12%)	±0.007 (0.39%)
	±5.204 (8.2%)	±3.692 (8.3%)	±2.355 (8.2%)	±1.106 (10.9%)	±0.42 (25.4%)

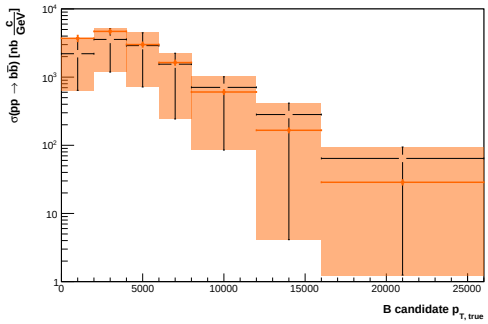
Table 20: Differential cross-section in terms of p_T and y of the B candidate including the statistical- (first) and systematical (second) error. The number in brackets is in [%] of the cross-section. The result is a combination of the two magnet polarity samples. Note the small fitting error.



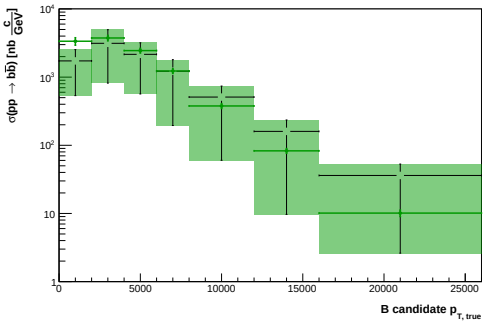
(a) $1.5 < y < 2.5$ [nb $\frac{c}{\mathrm{GeV}}$]



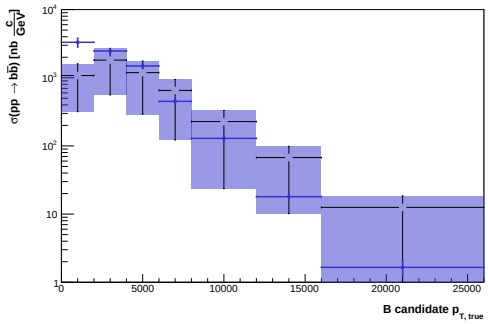
(b) $2.5 < y < 3.0$ [nb $\frac{c}{\mathrm{GeV}}$]



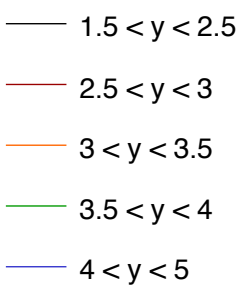
(c) $3.0 < y < 3.5$ [nb $\frac{c}{\mathrm{GeV}}$]



(d) $3.5 < y < 4.0$ [nb $\frac{c}{\mathrm{GeV}}$]



(e) $4.0 < y < 5.0$ [nb $\frac{c}{\mathrm{GeV}}$]



(f) legend

Figure 28: Differential cross-section result for each rapidity bin in [nb $\frac{c}{\mathrm{GeV}}$] (a) to (e). The boxes signify the prediction with uncertainty, the points with error bars the measurement. The uncertainties given are the systematic- and statistical uncertainties added in quadrature.

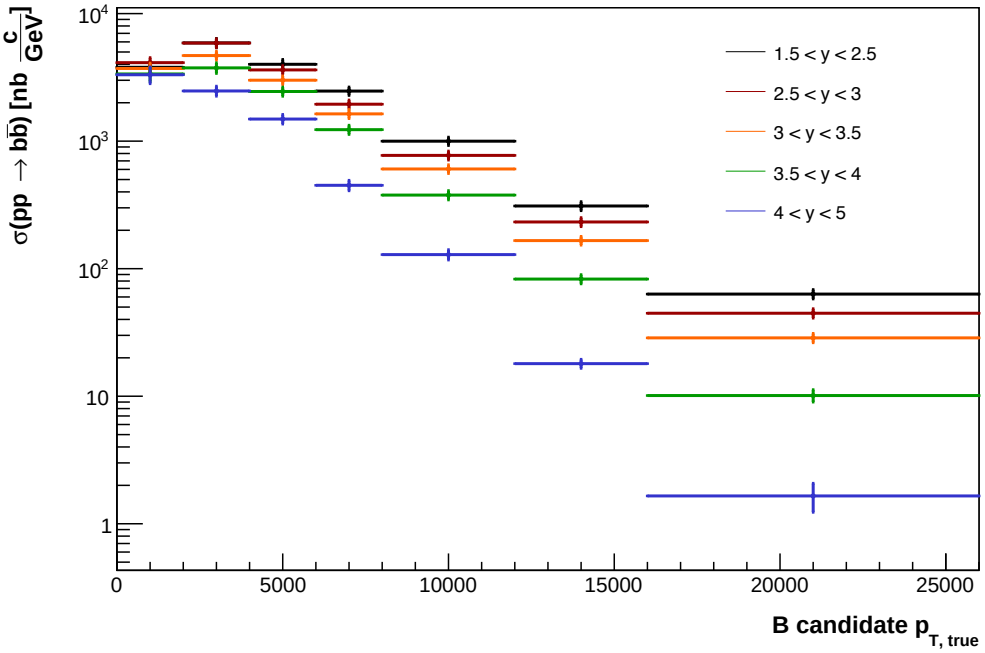


Figure 29: Differential cross-section result without the predictions in [nb $\frac{c}{GeV}$]. The color scheme is the same as in the legend of 28f.

6 Acknowledgements

I would like to thank Prof. Dr. Ueli Straumann for allowing me to contribute to the group's work and to write this thesis. My greatest gratitude goes to my supervisors Dr. Christian Elsasser and Dr. Katharina Müller who supported me throughout the analysis. They showed invaluable dedication and patience. Furthermore, I would like to thank the researchers in office Y36-K–94 for being outstandingly nice and calming me down when I have encountered a segmentation fault.

A Signal Composition Estimate

Because of lepton universality, we can write l instead of μ or e. Note, that modes with $n \cdot \pi$ have an unknown composition of D^* mesons. We assume half D^{*-} and half $\overline{D}^{*0}$ composition and introduce the difference as an additional uncertainty.

decay mode	branching fraction $\mathcal{B}$
$B^+ \to \overline{D}^0 l^+ \nu_l$	$(2.27 \pm 0.11)\%$
$B^+ \to \overline{D}^*(2007)^0 l^+ \nu_l$	$(5.69 \pm 0.19)\%$
$\longrightarrow \overline{D}^*(2007)^0 \to D^0\pi^0 or D^0\gamma$	$(100.0 \pm 4.1)\%$
$B^+ \to D^{(*)} n \cdot \pi l^+ \nu_l (n \geq 1)$	$(1.87 \pm 0.26)\%$
$\longrightarrow \overline{D}^*(2007)^0 \to \overline{D}^0\ \overline{\pi}^0 or \overline{D}^0\gamma$	$(100.0 \pm 4.1)\%$
$\longrightarrow D^{*-} \to \overline{D}^0\pi^-$	$(67.7 \pm 0.5)\%$
$B^+ \to \overline{D}^0 l^+ \nu_l X$	$(9.53 \pm 0.72)\%$

Table 21: B^+ decay modes according to PDG [1]. The charge-conjugate modes are likewise.

decay mode	branching fraction $\mathcal{B}$
$B^0 \to D^*(2010)^- l^+ \nu_l$	$(4.93 \pm 0.11)\%$
$\longrightarrow D^*(2010)^- \to \overline{D}^0\pi^-$	$(67.7 \pm 0.5)\%$
$B^0 \to \overline{D}^0\pi^- l^+ \nu_l$	$(4.3 \pm 0.6) \cdot 10^{-3}$
$B^0 \to D^* n \cdot \pi l^+ \nu_l (n \geq 1)$	$(2.3 \pm 0.5)\%$
$\longrightarrow \overline{D}^*(2007)^0 \to \overline{D}^0\ \overline{\pi}^0 or \overline{D}^0\gamma$	$(100.0 \pm 4.1)\%$
$\longrightarrow D^{*-} \to \overline{D}^0\pi^-$	$(67.7 \pm 0.5)\%$
$B^0 \to \overline{D}^0 l^+ \nu_l X$	$(5.70 \pm 0.86)\%$

Table 22: B^0 decay modes according to PDG [1]. The charge-conjugate modes are likewise.

decay mode	branching fraction $\mathcal{B}$
$B_s^0 \to D_{s1}(2536)^- \mu^+ \nu_\mu \to D^{*-} K_s^0 \mu^+ \nu_\mu$	$(2.5 \pm 0.7) \cdot 10^{-3}$
$\longrightarrow D^{*-} \to \overline{D}^0\pi-$	$(67.7 \pm 0.5)\%$
$B_s^0 \to D_{s1}(2536)^- \mu^+ \nu_\mu X \to \overline{D}^0 K^+ \mu^+ \nu_\mu X$	$(4.3 \pm 1.7) \cdot 10^{-3}$
$B_s^0 \to D_{s2}(2573)^- \mu^+ \nu_\mu X \to \overline{D}^0 K^+ \mu^+ \nu_\mu X$	$(2.6 \pm 1.2) \cdot 10^{-3}$
$B_s^0 \to \overline{D}^0 l^+ \nu_l X$	$(0.88 \pm 0.21)\%$

Table 23: B_s^0 decay modes according to PDG [1]. The charge-conjugate modes are likewise.

The contribution from b baryons seems to be negligible, since the decay modes almost exclusively involve other baryons and no D^0 decay is mentioned in the PDG book. [1].

B Tau to Muon Contribution

Table 24 shows the branching fraction of a $\tau \to \mu X$ decay. Table 25 shows the branching fraction

decay mode	branching fraction $\mathcal{B}$
$\tau^+ \to \mu^+ \nu_\mu \overline{\nu}_\tau$	$(17.41 \pm 0.04)\%$

Table 24: $\tau \to \mu X$ branching fraction according to PDG [1]. The charge-conjugate modes are likewise.

of the various $b \to D^0 \tau X$ decays for the analysis decay channel. For the D^* decays, see Tab. 21 to Tab. 23 in appendix A.

decay mode	branching fraction $\mathcal{B}$
$B^0 \to D^*(2010)\tau^+\nu_\tau$	$(1.84 \pm 0.22)\%$
$B^+ \to \overline{D}^*(2007)^0\tau^+\nu_\tau$	$(1.88 \pm 0.20)\%$
$B_s^0 \to D^0(^*)\tau X$	~ 0

Table 25: Branching fraction for $b \to D^0 \tau X$ according to PDG [1]. The charge-conjugate modes are likewise.

List of Figures

List of Tables

References

[1] **Particle Data Group** Collaboration, K. Olive *et al.*, "Review of Particle Physics," *Chin.Phys.* **C38** (2014) .

[2] **LHCb** Collaboration, R. Aaij *et al.*, "Measurement of $\sigma(b\bar{b})$ at $\sqrt{s} = 7$ TeV in the forward region," *Physics Letter B* **3** no. 694, (2010) 209–2016.

[3] M. Bustamante, L. Cieri, and J. Ellis, "Beyond the Standard Model for Montaneros," arXiv:0911.4409 [hep-ph].

[4] M. Herrero, "The Standard model," *NATO Sci.Ser.C* **534** (1999) 1–59, arXiv:hep-ph/9812242 [hep-ph].

[5] R. Monroe, *what if?* Houghton Mifflin Harcourt, Boston, 2014.

[6] **D0** Collaboration, V. Abazov *et al.*, "Evidence for production of single top quarks," *Phys.Rev.* **D78** (2008) 012005, arXiv:0803.0739 [hep-ex].

[7] PDG, "Modern View (Standard Model) timeline: 1964 - present," 15.05.2015. http://pdg.web.cern.ch/pdg/cpep/history/smt.html.

[8] G. Bhattacharyya, "A Pedagogical Review of Electroweak Symmetry Breaking Scenarios," *Rept.Prog.Phys.* **74** (2011) 026201, arXiv:0910.5095 [hep-ph].

[9] R. Gupta, "Introduction to lattice QCD: Course," arXiv:hep-lat/9807028 [hep-lat].

[10] R. Placakyte, "Parton Distribution Functions," arXiv:1111.5452 [hep-ph].

[11] C. Amsler, *Kern- und Teilchenphysik*. vdf Hochschulverlag AG an der ETH Zuerich, Zuerich, 2007.

[12] T. Nakada, "Review on CP violation," *AIP Conf.Proc.* **302** (1994) 425–463, arXiv:hep-ph/9312290 [hep-ph].

[13] Garyzx, "Standard Model Interactions," 08.04.2015. http://upload.wikimedia.org/wikipedia/commons/7/75/Standard_Model_Feynman_Diagram_Vertices.png.

[14] J. Ziegler, "The stopping of energetic light ions in elemental matter," *J. Appl Phys/Rev. Appl Phys.* **85** (1999) 1249–1272.

[15] LHCb, "b-quark production angle," 15.05.2015. https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles/07_rad_acc_scheme_right.pdf.

[16] Pamputt, "pseudo rapidity," 15.05.2015. http://upload.wikimedia.org/wikipedia/commons/thumb/3/30/Pseudorapidity2.png/220px-Pseudorapidity2.png.

[17] LHCb Collaboration, "The LHCb Detector at the LHC," *Journal of Instrumentation* **3** no. 08, (2008) S08005. http://stacks.iop.org/1748-0221/3/i=08/a=S08005.

[18] **LHCb** Collaboration, R. Aaij *et al.*, "LHCb Detector Performance," *Int.J.Mod.Phys.* **A30** (2015) 1530022, arXiv:1412.6352 [hep-ex].

[19] LHCb, "Trigger concept of LHCb," 13.04.2015. `http://lhcb-trig.web.cern.ch/lhcb-trig/`.

[20] LHCb, "Stripping Page of LHCb," 13.04.2015. `https://twiki.cern.ch/twiki/bin/view/LHCb/LHCbStripping`.

[21] **LHCb** Collaboration, R. Aaij *et al.*, "Precision luminosity measurements at LHCb," *JINST* **9** no. 12, (2014) P12005, `arXiv:1410.0149 [hep-ex]`.

[22] M. Adinolfi *et al.*, "Performance of the LHCb RICH detector at the LHC," *European Physical Journal C* **73** (May, 2013) 2431, `arXiv:1211.6759 [physics.ins-det]`.

[23] X. Cid Vidal, J. Hernando Morata, J. Lopes, D. Milanes, M. Palutan, E. Polycarpo, B. Sciascia, and F. Soomro, "Performance of the LHCb Muon Identification," *LHCb-INT* **016** (2012) .

[24] M. Clemencic *et al.*, "The simulation application, Gauss: Design, evolution and experience," *J. Phys. Conf. Ser.* **331** (2011) 032023.

[25] T. Sjöstrand, S. Mrenna, and P. Skands, "Pythia 6.4 physics and manual," *JHEP* **05** (2006) 026, `arXiv:hep-ph/0603175 [hep-ph]`.

[26] I. Belyaev *et al.*, "Handling of the generation of primary events in Gauss, the LHCb simulation framework," *J. Phys. Conf. Ser.* **331** (2011) 032047.

[27] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

[28] P. Golonka and Z. Was, "Photos Monte Carlo: a precision tool for QED corrections in Z and W decays," *Eur.Phys.J.* **C45** (2006) 97–107, `arXiv:hep-ph/0506026 [hep-ph]`.

[29] **Geant4 collaboration** Collaboration, S. Agostinelli *et al.*, "Geant4: a simulation toolkit," *Nucl. Instrum. Meth.* **A506** (2003) 250.

[30] **Geant4 collaboration** Collaboration, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Dubois, *et al.*, "Geant4 developments and applications," *IEEE Trans.Nucl.Sci.* **53** (2006) 270.

[31] LHCb Collaboration, "Strippingb2D0MuXB2DMuNuXLine," 09.04.2015. `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/stripping/config/stripping20/semileptonic/strippingb2d0muxb2dmunuxline.html`.

[32] Y. X. (LHCb), "OfflineVertexFitter," 16.05.2015. `https://twiki.cern.ch/twiki/bin/view/LHCb/VertexFitters#Vertex_Fitters`.

[33] N. M. (LHCb), "Identification of Ghost Tracks using a Likelihood Method," 16.05.2015. `http://cds.cern.ch/record/1107564/files/lhcb-2008-026.pdf`.

[34] T. Skwarnicki, "A STUDY OF THE RADIATIVE CASCADE TRANSITIONS BETWEEN THE UPSILON-PRIME AND UPSILON RESONANCES,". DESY-F31-86-02.

[35] C. ROOT, "Data Analysis Framework in HEP," 14.04.2015. `http://root.cern.ch`.

[36] **LHCb** Collaboration, R. Aaij *et al.*, "Measurement of the track reconstruction efficiency at LHCb," *JINST* **10** no. 02, (2015) P02007, arXiv:1408.1251 [hep-ex].

[37] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC," *Nucl.Phys.Proc.Suppl.* **205-206** (2010) 10–15, arXiv:1007.3492 [hep-ph].