Department of Physics and Astronomy
University of Heidelberg

Bachelor Thesis in Physics
submitted by

Nicola Alain Ackermann

born in Ottweiler (Germany)

2019

Identification of non-prompt $D^+ \to K^-\pi^+\pi^+$ decays and application to $\bar{B}^0 \to D^+ D_s^-$ at the LHCb experiment

This Bachelor Thesis has been carried out by Nicola Alain Ackermann at the
Physikalisches Institut in Heidelberg
under the supervision of
Prof. Dr. Ulrich Uwer

Abstract

This thesis presents the development of a generic multivariate selection of detached $D^+ \to K^-\pi^+\pi^+$ decays from B meson decays. The tool that was used for this selection are Boosted Decision Trees (BDTs) and the selection was conducted and calibrated on the $\bar{B}^0 \to D^+\pi^-$ channel. Different approaches to the classification of signal and background for the training of the BDTs were tested and a comparison between the BDT selection and a selection using PID variables was conducted on the decay $\bar{B}^0 \to D^+D_s^-$. The work presented here uses data taken at the LHCb experiment in the year 2018 corresponding to an integrated luminosity of 2.19 fb^{-1}.
It was found that the performance of the BDT selection is superior to the one using PID variables and a specific classification of signal and background in the BDT training, where several misidentified decays were classified as sgnal along with the combinatorial background, was found to give the best results.

Kurzfassung

Diese Arbeit präsentiert die Entwicklung einer generischen multivariaten Selektion von $D^+ \to K^-\pi^+\pi^+$ Zerfällen beim Zerfall von B-Mesonen. Hierfür wurden Boosted Decision Trees (BDTs) verwendet und die Selektion wurde mit dem Zerfallskanal $\bar{B}^0 \to D^+\pi^-$ kalibriert. Verschiedene Herangehensweisen für die Unterteilung der verschiedenen Komponenten in dessen Massenspektrum in Signal und Untergrund wurden für das Training der BDTs untersucht und ein Vergleich zwischen einer Selektion mit BDT cuts und einer Selektion mit PID-Variablen wurde für den Zerfall $\bar{B}^0 \to D^+D_s^-$ getestet. Die hier präsentierte Arbeit verwendet Daten, die im Jahr 2018 bei einer integrierten Luminosität von 2,19 fb^{-1} aufgenommen wurden.
Es wird gezeigt, dass die Selektion mit BDT cuts bessere Ergebnisse liefert als die Selektion mit PID-Variablen und dass eine Unterteilung der Komponenten in Signal und Untergrund, bei der verschiedene misidentifizierte Zerfälle als Untergrund klassifiziert wurden, sich als die effizienteste herausstellt.

Acknowledgements

This part is deticated to the people who have helped me and given me support during the completion of this thesis in the last 5 months.
First I would like to thank my two supervisors Dr. Marian Stahl and Dr. Nicole Skidmore without who this thesis could have never been done. Their constant support and direction has made this task possible and I was able to learn a lot from them.
Secondly would like to thank Dr. Sebastian Neubert, the head of the spectroscopy department of the LHCb group in Heidelberg, and the whole spectroscopy group for welcoming me into the group and also giving advice, when needed. The small Tuesday morning meetings were always a pleasure and have given me valuable insight into the work of phycisists in the field of high energy particle physiscs.
I would also like to thank Dr. Ulrich Uwer for giving me the opportunity to study in his working group and to complete this thesis under his guidance.
Lastly I would like to thank my two fellow students Nils Hoyer and Maurice Pierre Morgenthaler with whom I shared an office for the last 5 months and who have helped make this task more enjoyable for myself.

Contents

1 Introduction

This goal of this thesis is the development of a generic multivariate selection of $D^+ \to K^-\pi^+\pi^+$ decays in the context of the LHCb experiment. The tool that was chosen for this selection are Boosted Decision Trees (BDTs). These boosted decision trees are multivariate analysis tools for distinguishing between events that belong to a particular decay of interest (**Signal**) and events that are background data. In particular a D-from-B BDT is established for the D^+ meson, i.e. a BDT that selects D^+ meson candidates from the decay of a B meson. This BDT will then be applied to the $\bar{B}^0 \to D^+D_s^-$ channel and the results will be compared to a standard signal selection using PID cuts. Such BDTs have been established for $\Lambda_c^+ \to pK^-\pi^+$, $D^0 \to K^-\pi^+$ and $D_s^+ \to K^+K^-\pi^+$ [1], and this thesis will present the development of an $D^+ \to K^-\pi^+\pi^+$ BDT, where especially the mass spectrum of the calibration channel will be analysed with more scrutiny.
The first step is the detailed study of the decay $\bar{B}^0 \to D^+\pi^-$ which will be used as the calibration channel for the D-from-B BDT. Here an exact understanding of the $\bar{B}^0 \to D^+\pi^-$ is of interest since this will lead to more stable fits for the efficiency correction and will also allow for different approaches to the BDT to be tested. After that the BDTs are trained on this channel and then applied to the $\bar{B}^0 \to D^+D_s^-$ channel were the results will be obtained.
In total this thesis consists of 9 sections. Section 2 gives a brief overview of the Standard Model of particle physics as a theoretical background for the work that is carried out. Section 2 then explains the data taking at the LHCb experiment and the different components of the detector and also focuses on the trigger mechanism used in the experiment. The data that passes this trigger then goes on to the stripping selection, which is the topic of section 4. Section 5 explains the two methods of offline signal selection that are compared in this thesis, boosted decision trees and PID variables. The aforementioned analysis of the $\bar{B}^0 \to D^+\pi^-$ mass spectrum is subject of section 6 and section 7 details the training of the different BDTs on this channel. Section 8 then presents the results, first a comparison between the different approaches to the D^+ BDTs and then a comparison between BDT and PID cuts on the channel $\bar{B}^0 \to D^+D_s^-$. Lastly section 9 gives a conclusion and an outlook on future tasks to be completed on this topic.

Motivation for the use and study of Boosted Decision Trees

Figure 1 illustrates the different cross sections and production rates of different physical processes for the LHC at 14 TeV. As can be seen here the production of $b\bar{b}$ and $c\bar{c}$ quark pairs , which are the processes of interest for the LHCb experiment, happen at a rate of several MHz, which has to be handled accordingly.
The current strategy to reduce this rate and thus make the amount of data managable is compromised of the two stages of the **trigger** used in the experiment, which is detailed in section 3.3. However for Run III the amount of data produced will increase to allow for more statistics to be used in spectroscopy or precision measurements [4]. To be able to handle this data a more efficient trigger will also have to be employed that selects the events of interest and rejects all the unnecessary background data and, for the decays were the amount of signal

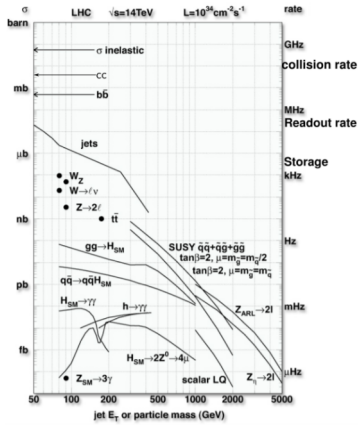


Figure 1: Different cross sections and production rates of different physical processes for the LHC at 14 TeV [3]

events is abundant, is able to differentiate between the more interesting parts of the signal and the ones that are not of particular interest. In order to do this the trigger needs good selection variables. The problem that one faces here is that this selection of events of interest (signal events) is not trivial and all the data that is rejected by the trigger is lost and can no longer be used for analysis. Hence the trigger has to meet two requirements: low bandwidth, due to a limited amount of disk space and time, and a well understood selection, to ensure that the output is calibrated to extract physical quantities of interest.

To overcome these challenges, in particular when it comes to classification problems in the selection stage of the trigger, machine learning algorithms have proven to be a possible solution. These machine learning algorithms have the power to summarise the information of potentially very complex physical processes into a single selection variable, which can be treated as a high level physics object. An example for the application of such an object would be as a seed particle for semi-inclusive selections, where only part of the event is written to disk, known as selective persistence. [5].

The machine learning algorithms currently under study are Boosted Decision Trees (BDTs), including D-from-B BDTs which will be the topic of this thesis. These BDTs aim to select a $D^+ \to K^-\pi^+\pi^+$ decay coming from the decay of a B meson.

As an important step towards the inclusion of these BDTs in the update trigger, they have already been included in the current incremental restripping campaign for the 2018 LHCb data (cf section 4), a collaboration wide pre-selection of the 2018 data for user analyses. For this the updated D-from-B BDTs partially presented in this thesis were included in a double-D inclusive line, which aims to select D mesons by requiring a certain BDT response and building a vertex of at least two of these mesons. The analyst can then build any decay which involves two D mesons that originate from a common vertex offline.

This is an important step to establish the D-from-B BDTs as an effective tool for the whole collaboration.

2 The Standard Model

The following section will be a brief summary on the Standard Model (SM).
The Standard Model of particle physics describes three of the four known fundamental forces (the electromagnetic, weak, and strong interactions, and not including the gravitational force) in the universe, as well as classifying all known elementary particles [6]. The elemantary particles consist of the fermions, which have spin $\frac{1}{2}$ and make up the matter in the universe and the bosons, which have integer spin and mediate the interaction between the fermions, i.e. the fundamental forces. The fermions are further divided into 6 leptons and 6 quarks and the 6 different types of leptons and quarks are called flavours. These flavours can be divided into three generations with each generation including a charged lepton ($q = -e$), a neutral neutrino, a positively charged quark ($q = \frac{2}{3}$e) and a negatively charged quark ($q = -\frac{1}{3}$e). Additionally there is an anti-particle for each of these fermions, which has the exact opposite charges.
While leptons can be observed on their own in an experiment, this is not true for quarks. This is due to the fact that quarks (and gluons) can only exist in colour-neutral bound states. This is called colour confinement [7]. The most common of these bound states are baryons, combinations of three quarks, including the proton (uud) and the neutron (udd), and mesons, which are the combination a quark and an anti-quark. Table 1 gives an owerview of the 12 fermions of the Standard Model.

Table 1: Masses and charges of the 12 fermions

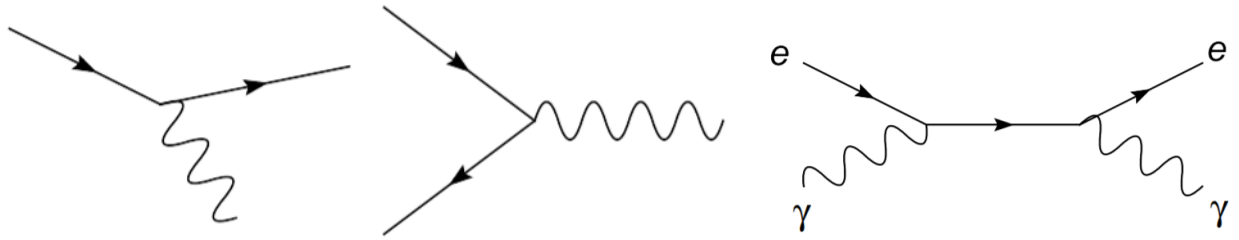
	First generation			Second generation			Third generation		
	flavour	$\frac{q}{e}$	m [GeV]	flavour	$\frac{q}{e}$	m [GeV]	flavour	$\frac{q}{e}$	m [GeV]
Quarks	up	$\frac{2}{3}$	0.005	charm	$\frac{2}{3}$	1.3	top	$\frac{2}{3}$	174
Quarks	down	$-\frac{1}{3}$	0.003	strange	$-\frac{1}{3}$	0.1	bottom	$-\frac{1}{3}$	4.5
Leptons	electron	-1	0.0005	muon	-1	0.106	tau	-1	1.78
Leptons	electron neutrino	0	$< 10^{-9}$	muon neutrino	0	$< 10^{-9}$	tau neutrino	0	$< 10^{-9}$

The three fundamental forces described by the SM are mediated by spin 1 gauge bosons. Quantum Electro Dynamics (QED) describes the interaction between electrically charged particles via exchanging of photons and Quantum Chromo Dynamics (QCD) describes the interaction between particles with colour charge via exchanging of gluons. For the weak interaction there are two bosons, the neutral Z^0 and the charged $W^{\pm}$. In general a boson can only couple to a fermion if the fermion carries the necessary charge for that interaction. These charges are the electric charge for the electromagnetic interaction (photons), the colour charge for the strong interaction (gluons) and the weak isospin for the weak interaction (W- and Z-bosons).

2.1 The electromagnetic interaction

The electromagnetic interaction mediated by a photon has a long history of investigation and is now known to be described by quantum electrodynamics (QED). The range of interaction is

infinite since it decreases with $\frac{1}{r}$ and the coupling constant of the electromagnetic interaction, which describes the strength of the interaction, is $\alpha_{em} \approx \frac{1}{137}$. Typical interactions can be seen in Figure 2.



(a) electron emits a photon (b) electron-positron annihilation (c) Compton Scattering

Figure 2: Typical electromagnetic interactions

An example for a electromagnetic decay that is of importance for this thesis is the decay $D^{*+} \to D^+\gamma$ as part of the $\bar{B^0} \to (D^{*+} \to D^+\gamma)\pi^-$ decay. It is very similar to the decay shown in Figure 2a and is one of the partially reconstructed decays considered in section 6.2.

2.2 The strong interaction

The theory of QCD as a description of the strong interaction was developed by several phycisists in the 1950s to 1970s. It includes gluons as gauge bosons, which themselves have a colour charge. This makes the existence of bound states of gluons, the so-called glueballs possible.
The coupling constant of the strong interaction α_S varies with the squared momentum transfer Q^2 of the interacting particles. For small Q^2 α_S becomes large, which gives rise to colour confinement (see above). On the other hand α_S becomes small for large Q^2 as can be seen in deep inelastic scattering, where quarks and gluons behave as free particles. This is called assymptotic freedom.
The strong interaction has a very small effective range of about 10^{-15}m as opposed to the electromagnetic interaction.
A typical processes can be seen in Figure 3.

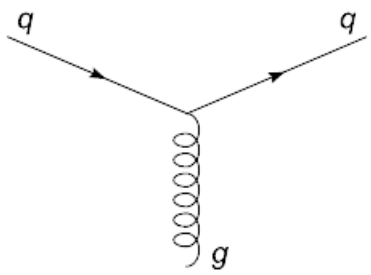


Figure 3: Quark emits a gluon

One of the strong decays that are of particular interest for this thesis is the decay $D^{*+} \to D^+\pi^0$ as it occurs in the decay $\bar{B}^0 \to (D^{*+} \to D^+\pi^0)\pi^-$. This is an important partially reconstructed decay in the $\bar{B}^0 \to D^+\pi^-$ mass spectrum and will again feature in section 6.2.

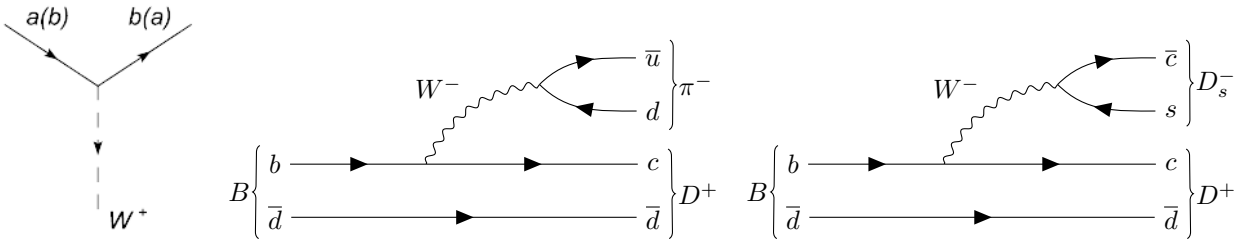
2.3 The weak interaction

It is easy to prove and experimentally verified that in QED and QCD parity is conserved, which means that all processes in QED and QCD behave the same, when the spatial coordinates are "reversed" ($\vec{x} \to -\vec{x}$). However this is not the case for the weak interaction, which was first suggested by Lee and Yang in 1956 [9] and then experimentally proven in the Wu experiment in 1957 [10]. This gives rise to the chiral structure of the weak interaction, which means that only left-handed chiral particles and right-handed chiral antiparticles participate in the charged-current weak interaction.
Additionally the weak interaction is the only interaction in which the participating fermions can change flavour, which is described by the the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix [11].

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

The entry $|V_{q_1 q_2}|^2$ describes the probability of the quark q_1 changing into q_2.
The weak interaction has a very small range of about 10^{-17}m, which is due to the fact that its gauge bosons Z^0 and $W^\pm$ have masses.
Figure 4 shows typical weak interactions.

Figure 4b and 4c show two decays that are of great importance to this thesis. 3b shows the decay $\bar{B}^0 \to D^+\pi^-$, which will be used as the calibration channel for the D^+ BDTs in section 6 and the understanding of its mass spectrum will be one of the main tasks presented here.
Figure 4c shows the decay $\bar{B}^0 \to D^+D_s^-$ which will be the decay on which the BDTs will be tested in section 8. As one can see clearly from the Feynman diagrams already the two decays show very similar topology, which allows the use of the $\bar{B}^0 \to D^+\pi^-$ channel as calibration



(a) quark emits a W boson (quark changes flavour)

(b) $\bar{B^0} \to D^+\pi^-$

(c) $\bar{B^0} \to D^+D_s^-$

Figure 4: Examples for weak interactions

channel.

3 The LHCb experiment

The Large Hadron Collider beauty experiment (LHCb) is one of the seven experiments currently being carried out at the Large Hadron Collider at the European Organization for Nuclear Research (CERN; from french: Conseil européen pour la recherche nucléaire) located at point 8 on the LHC tunnel close to Ferney-Voltaire, France. The LHCb collaboration involves 74 intitutes from 16 countries.
The goal of the experiment is the investigation of CP violation and rare decays through the analysis of decays involving b quarks [12]
The experiment uses a single-arm forward spectrometer (the LHCb detector) of a weight of 5600 tons and a volume of $21 \times 13 \times 10 m^3$ to study such decays. The results of these measurements can then be compared to the predictions of the Standard Model either confirm it or show potential for physics beyond the SM.

3.1 The Large Hadron Collider

The LHC is the largest particle collider in the world today at a circumference of approximately 27 km. Its primary focus is the investigation of pp collisions and PbPb collisions.
The collider was operated at a center-of-mass energy of 7 TeV in 2011, 8 Tev in 2012 (Run 1) and 13 TeV in 2015 - 2018 respectively (Run 2). Currently it is in a two-year shut down period, which started at the end of 2018, in order to perform several hardware and software upgrades around the machine. One of these updates will be an update on the trigger systems used in the LHCb experiment as mentioned in the introduction.

3.2 The LHCb detector

The following sections are a short summary of the hardware setup used at the LHCb experiment for precision measurements. The predominant production mechanism for heavy flavour hadrons at the LHC is gluon fusion [13], which boosts the produced system in either forward or backwards direction in the laboratory frame due to the assymetrical momenta of the participating partons. The spectrometer used by LHCb is taylored to this process.
A schematic view of the LHCb spectrometer is shown in Figure 5 [14].

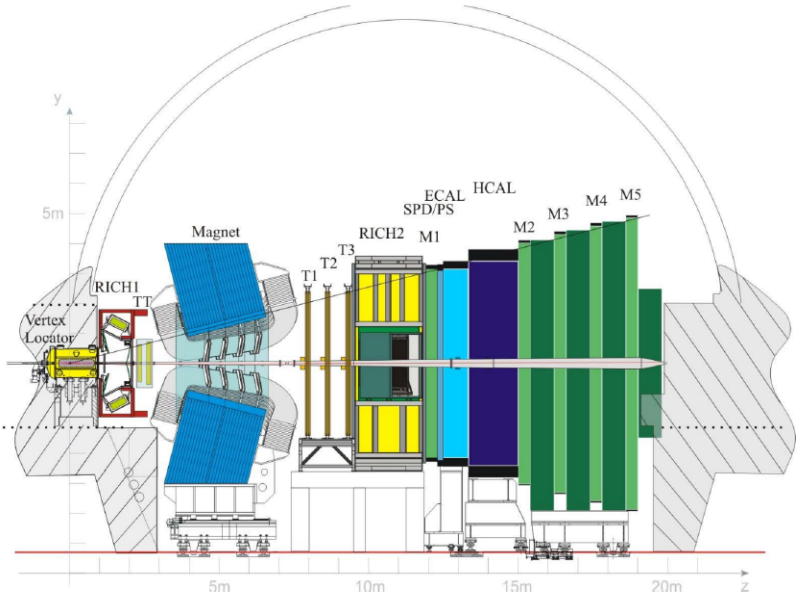


Figure 5: Schematic view of the LHCb spectrometer

3.2.1 Vertex locator

The vertex locator (VELO) is used to measure particle tracks near the point of collision and allows for a precise reproduction of the primary vertices from the pp collisions as well as the secondary vertices from the decay of b- and c-hadrons [15].
The VELO is made up of 42 silicon modules arranged along the beam as shown in Figure 6. The modules are seperated into r-modules and Φ-modules, that measure the radial distance r of a track to the beam and the azimuthal angle Φ of the track respectively. Using these coordinates and the position of the modules the 3D track of a particle can be accurately reconstructed.
To avoid unnecassary radiation damage the two halves of the VELO are movable. When the beam is not yet stabilized the two halves are opened with a seperation of 6 cm while in its closed state the closest module is at a distance of 8.2 mm to the beam.
The VELO can detect particles with pseudorapidity $1.6 < \eta < 4.9$ and within 10.6 cm range from the colliding point.

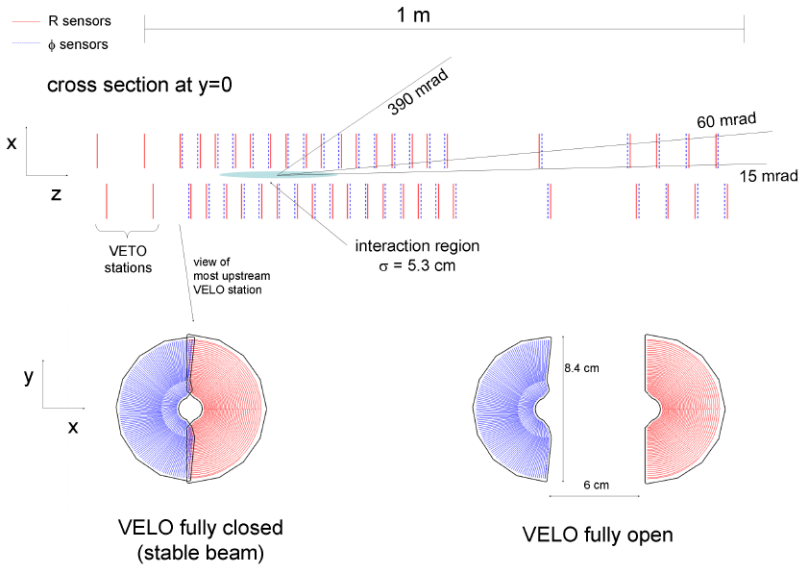


Figure 6: Arrangement of the modules and comparison between open and closed position [16]

3.2.2 Magnet

In order to measure the momenta of charged particles, that are produced, a dipole magnet of two coils is mounted 5 m from the colliding point. It provides an integrated magnetic field of about 4 Tm for tracks of 10 m [17] and this field deflects charged particles in the horizontal plane, due to the Lorentz force, which allows for a determination of the particle momentum.

3.2.3 Tracking stations

In total there are four tracking stations used in the LHCb detector: the trigger tracker (TT) and the trackers T1, T2 and T3. These are used to reconstruct the tracks of charged particles similarly to the VELO, which allows a measurement of the momentum and the charge of these particles.
For these measurements the TT and parts of T1 - T3, the inner parts close to the beampipe,(the inner tracker IT) use silicon microstrip detectors [18], while the rest parts of T1 - T3 (the outer tracker OT) use a drift-tube gas detector [19]. The OT is in fact the largest part of T1, T2 and T3.

3.2.4 Ring Imaging Cherenkov Detectors

For particle identification of charged hadrons the LHCb experiment primarily uses two ring imaging Cherenkov detectors (RICH): RICH I ,which is located between the VELO and the TT, and RICH II, which is located behind the T3. Both of these detectors use aerogel (only in run I) and C^4F^{10} gas radiators. RICH I is used for particle identification (PID) of charged particles with low momenta ($p \approx 1 - 60$ GeV) while RICH II identifies particles with higher momenta ($p \approx 15 - 100$ GeV) [20].
PID in the RICH detector is based on the fact that particles that are travelling in a dielectric medium with refractive index n at a speed larger than the speed of light in the medium ($v > \frac{c}{n}$) emit photons [21]. This radiation is called Cherenkov radiation and is emitted at a fixed angle Θ_C to the velocity vector of the charged particle. This angle is directly related to the velocity of the particle:

$$cos(\Theta_C) = \frac{1}{n\beta} \tag{1}$$

where β is the velocity in units of c ($\beta = \frac{v}{c}$). The emitted Cherenkov radiation is registered by Hybrid Photon detectors (HPDs), which allows for a reconstruction of the angle Θ_C, which in turn gives the velocity of the particle. Together with the momentum measurement from the tracking stations the mass of the particle can now be calculated, which means that the particle can be identified.

RICH reconstruction

The information from the RICH detector, combined with the momentum measurement from the tracking stations is used to calculate particle identification (PID) variables for a particle, also known as delta-log-likelihood (DLL). They are the response of a log-likelihood algorithm that aims to minimise a negative overall event log-likelihood by first assuming that all particles in the event are pions and then step-by-step changing a particle hypothesis until no further improvement can be reached. The result of this algorithm is the DLL, which is defined as:

$$PIDX = \Delta log\mathcal{L}_{X-\pi} = log\frac{\mathcal{L}_X}{\mathcal{L}_\pi} \tag{2}$$

where $\mathcal{L}$ is the log-likelihood for a given particle hypothesis. This means that a positive PIDk value for example indicates that the particle candidate is more likely to be a Kaon than a Pion.
These variables are of great importance and will later also be used in the stripping, detailed in section 4.
Using additional information from the muon stations, the tracking detectors and the calorimeters an artificial neural network has been developed, whose response for a particle type X is referred to as ProbNNX [22]. These variables can later be used in an offline selection as will be done in section 8.
To illustrate the use of these ProbNN variables one can look at the variable $pi\_ProbNNk$ in the decay $\bar{B}^0 \to D^+\pi^-$, which will be of importance later on. This variable deals with the bachelor pion candidate from the stripping and gives information about how likely it is that this bachelor Pion candidate is in fact a Kaon (k). It holds that

$$0 < pi\_ProbNNk < 1 \tag{3}$$

except for tracks where the neural network lacks information. Placing the cut $pi\_ProbNNk > 0.9$ on the $\bar{B}^0 \to D^+\pi^-$ data for example would lead to only those events being selected for which the neural network has termed the Pion candidate to be very "kaon-like". In other words one would rougly select events of the decay $\bar{B}^0 \to D^+K^-$.

3.2.5 Calorimeters

The calorimeters form the basis for the hardware trigger L0 by selecting hadron, electrons and photons according to their deposited transverse energy, which will be detailed later in section 3.3. Furthermore it measures energy and position of particles which provides additional PID. The calorimeter system as seen in Figure 5 consists of a scintillator pad detector (SPD) plane in front of a preshower detector (PS), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL) [23]. The SPD is designed to identify charged particles, most of which are pions.

3.2.6 The muon system

Many important decay channels that are investigated in the LHCb experiment involve muons as final state particles. Those muons are not stopped by any of the calorimeters, which is why seperate muon systems are mounted furthest downstream, where they are the only particles expected to register a signal apart from fast charged tracks that can be easily distinguished from muons.
The muon system at LHCb consists of 5 muon stations M1 - M5. M1 is located in front of the calorimeters, while M2 - M5 are located behind those. To further reduce the signal and only select muons 80 cm thick iron plates are put between M2 - M5. Each of the stations is equipped with multi-wire proportional chambers (MWPC), except in the highest rate region of M1, where triple gas electron multipliers (GEM) are used [24].
Together with the calorimeters the muon stations make up the hardware trigger.

3.3 Trigger system

The following section will detail the trigger system used at the LHCb experiment. Since the work in this thesis will focus on data taken in 2018 the trigger conditions for this year will be given.

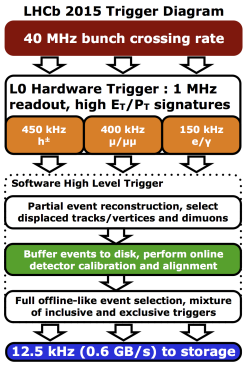


Figure 7: Schematic overview of the Trigger system for Run 2 (2015) [25]

The LHCb detector produces a very large amount of data. At an average luminosity of $2 \times 10^{32} cm^{-2} s^{-1}$ the detector registers events at a rate of 40 MHz [26], which needs to be reduced in order to be managable as many of the registered events are of no interest for further analysis and the amount of disk spave available is limited. For this purpose two levels of triggers are implemented in the LHCb detector: the Level 0 trigger (L0) and the High Level trigger (HLT). Together they reduce the rate to about 12.5 kHz (0.6 GB/s). Figure 7 shows an overview of the trigger setup for Run 2.
The L0 trigger is a hardware trigger and works by selecting hadrons, electrons and photons with high E_T deposits in the calorimeters as well as muons with high E_T deposited in the muon stations [26] [27].

The HLT is a software trigger and consists of two stages: HLT1 and HLT2. It uses a computing farm of 30000 cores (2012) to further reduce the number of events. In order to do so, HLT1 is based on dedicated algorithms, the so-called trigger lines, that use data from the VELO and the tracking stations. Due to limited computing resources the algorithms rely on a *partial* event reconstruction. The HLT2 decision then relies on several different selection algorithms, depending on which decay channel is of particular interest. These algorithms are based on many different selection criteria , such as invariant mass cuts or vertex reconstruction.

4 Stripping lines

After the data has arrived from the trigger system it is stored in so-called Tier 1 sites. From there the so-called *stripping* selection is applied to the data, which is a callaboration wide

pre-selection of data happening offline, where several selection algorithms from different LHCb working groups, called *stripping* lines are applied to the data. These stripping lines are a series of loose cuts to select candidate events for the different decays one wants to analyse. The cuts mainly focus on kinematics, topology and track reconstruction but also feature some very loose PID cuts. The strategy wll be explained in the following.
This thesis is intended to provide an analysis of the $\bar{B}^0 \to D^+\pi^-$ decay channel to be used as a calibration channel for the D^+ BDT, hence this decay will be chosen as an example for the stripping selection. Additionally a study of the BDT results for the $\bar{B}^0 \to D^+D_s^-$ channel will be provided, however this channel is very similar to $\bar{B}^0 \to D^+\pi^-$ and its stripping selection is based on the same module.

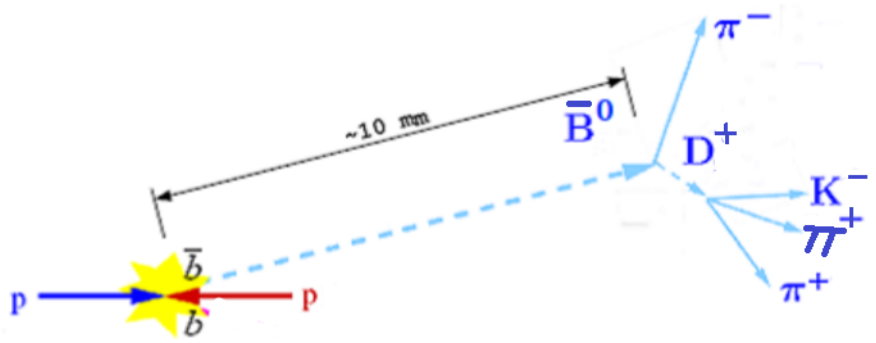


Figure 8: Topology of the decay $\bar{B}^0 \to D^+\pi^-$ [28]

Figure 8 gives an idea of the topology of this decay, which is an important signature to select such decays. After the initial pp collision (the primary vertex) the $\bar{B}^0$ flies a distance of 10 mm on average, corresponding to its life time $\tau_{\bar{B}^0} = (1.519 \pm 0.007) \times 10^{12}$s. After that it decays into a D^+ and a π^- at the secondary vertex. The D^+ then in turn decays into a negative Kaon and two positive Pions.
The stripping for this decay uses the line $StrippingB02DPiD2HHHBeauty2CharmLine$ [29] and starts with a global and track based selection. Tables 3 - 6 show each selection step used in this stripping line. The following sections will focus on explaining the important ones.

The first cut that the data that is selected by the topological trigger lines (cut 0 [27]) has to pass through are a cut to reject clone tracks (cut 1) and a cut to insure the quality of the track (cut 2). The latter is based on a good $\frac{\chi^2}{nDOF}$, which is calculated by the Kalman filter track fit [31]. The total number of long tracks in the event is limited to 500 (cut 3) and the fake-track rate is reduced by a cut on the track-ghost-probability (cut 4), which is a response of a neural network.

number	Cut	Applied to	Purpose
0	Hlt2Topo.*Decision	global	HLT2 selection
1	CloneDist > 5000	all tracks	clone track rejection
2	TRCHI2DOF < 3	all tracks	Good track quality
3	nLongTracks < 500	global	Remove outliers, limit cominatorics
4	TRGHP < 0.4	all tracks	Fake track rejection
5	P > 1000 MeV	all tracks	Remove low momentum tracks
6	PT > 100 MeV	all tracks	Remove low transverse momentum tracks
7	MIPCHI2DV(PV) > 4.0	all tracks	Reject tracks from primary vertex

Table 2: Global and track based selection [2]

Additionally a selection that removes tracks with low momentum (cut 5) and low transverse momentum (cut 6) is done. This is due to the fact that tracks with momentum lower than the given values, though they might be able to produce the desired particles, are simply surrounded by too much background data. This last selection step can be considered as loose as possible since multi-body beauty and charm decays tend to produce at least one soft particle.
The last cut at this stage removes tracks from the primary vertices (PV) by considering the minimal impact parameter (IP) [32], which is the closest approach by a track to a vertex. This is due to the fact that all tracks in this decay come from displaced vertices as can be seen in Figure 8.

number	Cut	Applied to	Purpose
8	ASUM(PT) > 1800 MeV	c daughters	Fast soft background rejection
9	**TRCHI2DOF < 2.5** **PT > 500 MeV** **P > 5000 MeV**	at least one c daughter	Require leading track
10	ACUTDOCA < 0.5 mm	c daughters	Save CPU time
11	PIDk > -10	K from D^+	Reduce combinatorics
12	PIDk < 20	π from D^+	Reduce combinatorics
13	VCHI2/VDOF < 10	D^+	c daughters from same point in space
14	BPVVDCHI2 > 36	D^+	Reject prompt charm hadrons
15	BPVDIRA > 0	D^+	Decay vertex downstream of PV

Table 3: Charm hadron selection [2]

Following the track based cuts the data goes through a series of cuts that intend to build a charm hadron candidate, in this case a D^+ (Table 3). The first cut listed here is a loose cut on the sum of the transvere momenta of the D^+ daughters to ensure they have enough momentum

to form the D^+ (cut 8). Furthermore tighter cuts on P, PT and TRCHI2DOF are placed to reject background data by selecting a leading track (cut 9). Cuts 11 and 12 are meant to further reduce the background by placing a very loose cut on the PID variables presented in section 3.2.4 .
After cut 12 the daughter particles form a common vertex of the D^+ candidate which is calculated via a vertex fit and the last three cuts listed here concern this vertex. VCHI2/VDOF is a loose cut on vertex quality, BPVVDCHI2 rejects tracks that come from the primary vertex (*prompt* charm hadrons) and BPVDIRA requires the vertex to be downstream from the primary vertex.

number	Cut	Applied to	Purpose
16	ASUM(PT) > 5000 MeV	b daughters	Fast soft background rejection
17	**TRCHI2DOF < 2.5** **PT > 1700 MeV** **P > 10000 MeV** **MIPDV(PV) > 0.1 mm**	at least one b daughter	Require leading track
18	**TRCHI2DOF < 2.5** **PT > 500 MeV** **P > 5000 MeV**	bachelor track	loose kinematic seletion
19	(4750 < AM < 7000) MeV	$\bar{B}^0$ combination	fast invariant mass cut
20	VCHI2/VDOF < 10	X_b	b daughters from same point in space
21	BPVVDCHI2 < 36	B^0	Candidate produced in PV
22	BPVDIRA > 0.999	B^0	Candidate points back to best PV
23	BPVLTIME > 0.2 ps	X_b	Reject combinatorial background

Table 4: Beauty hadron selection [2]

The selection of the beauty hadron candidate ($\bar{B}^0$, Table 4) is very similar to that of the charm hadron and it is formed from the charm candidate from the previous step and a pion candidate, the bachelor track. This bachelor track receives similar cuts to the leading track for the charm hadron.
The cuts on the b vertex are also similar and it is required to be located downstream of the primary vertex. Additionally the decay time of the b hadron is required to be larger than 0.2 ps to further reject background.

This is the last step in the stripping and the data that leaves here will be used as the basis for the analysis presented in this thesis.
However the selection so far concerned mainly geometrical quantities that are well understood in simulation and only very loose PID cuts were placed on the data. The result of this is that the data that leaves the stripping selection features a lot of background data that will have to be reduced to achieve a clean signature. Therefore the goal is to produce a PID-like quantity for $D^+ \to K^-\pi^+\pi^+$ and in order to do this Boosted Decision Trees will be used.

5 Boosted Decision Trees as a tool for offline signal selection

The main focus of this thesis will be to establish a Boosted Decision Tree for the selection of $D^+ \to K^-\pi^+\pi^+$ decays. The following section will give a short summary of these BDTs and their function.
The problem at hand is the following: Suppose there is a sample of events that is supposed to be used for the analysis of a particular decay. However this sample does not only consist of events that belong to this decay (**signal** events), it also features **background** events. The question is now how to properly classify these events into signal and background. One of the many tools to achieve this classification are Boosted Decicion Trees.

5.1 Decision Trees

A decsion tree is simply a binary tree that classifies the sample based on one-dimensional cuts on its input variables.

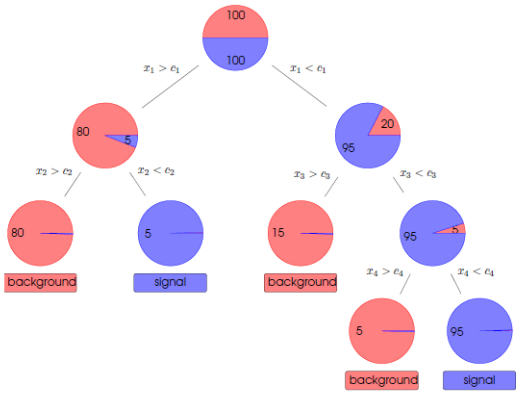


Figure 9: Example for a decision tree [33]

In order to "grow" the decision tree a **training sample** is needed in which for each event the class label is known. These class labels are **signal** and **background** in this case.
Starting from the top of the tree an even amount of background and signal events is given as input. Then a cut is placed on the data, in the case of Figure 9 $x_1 > c_1$, where x_1 and c_1 are values for one of the parameters of the training sample (e.g. decay angles, transverse momentum of the particles etc), and all the events with $x_1 > c_1$ follow the left branch, while all the events with $x_1 < c_1$ follow the right branch. More of these one-dimensional cuts are placed on the

data at each *node* (i.e. each step on the branch) and since the classification of background and signal events for the training sample is known, each node can be evaluated on the purity of the sample (i.e. if the sample consists of only signal/background). This is repeated until the tree arrives at an endpoint, a $leaf$, where only signal data or background data is left or until some stopping condition is met.
After the tree is grown in training it can then be applied to the actual decay of interest. However simple decision trees as shown in this example have several drawbacks. One of those is that they can be very non-robust, meaning that a small change in the training sample can lead to a large change in the tree and therefore in the final prediction. Another issue is that of *overtraining*. Suppose there is a particular type of event with a parameter value very close to one of the cuts. A statistical fluctuation in the training sample or in the real data might easily push these events into the wrong class, leading to the tree "learning" the statistical fluctuations of the sample.

These issues are overcome by using many decision trees and combining their outcome by *boosting*. This boosting algorithm also determines the exact cuts that are placed at each node as well as the number of nodes in each tree.

5.2 Boosting

The algorithm used for boosting the decision trees in this thesis is called **gradient boosting** [34], which is already implemented in the TMVA [35] software used here. It aims to minimise a *loss function* in training, whichis roughly speaking a function that penelises misclasification of events in training, where the classification of an event (signal or background) is known. For this function-minimisation problem the algorithm relies on the method of *graient-descent*, iteratively fitting *regression trees*, the so called weak learners, to the steepest local gradient to arrive at a global minimum. The final result of this boosting algorithm is a weighted sum of those regression trees, which behave like the above described decsion trees in a classification problem.

6 Analysis of the calibration channel $\bar{B}^0 \to D^+\pi^-$

As mentioned before a vital step in the application of the D-from-B BDTs, i.e the BDTs intended to select a D meson coming from the decay of a B meson, on the channel $\bar{B}^0 \to D^+D_s^-$ is the analysis of $\bar{B}^0 \to D^+\pi^-$. This channel will be used as a calibration channel for the BDT's which means that the BDT's will be trained on it.
For this training a fit of the full $\bar{B}^0 \to D^+\pi^-$ LHCb data sample taken in 2018 will have to be performed.

6.1 The choice of $\bar{B}^0 \to D^+\pi^-$ as calibration channel

In order for the BDT's to be trained on a particular channel the channel has to fulfill several requirements.
The first of these requirements is obviously that the desired particle, in this case D^+, is a decay product in the calibtration channel. Otherwise the BDT's would not be able to train on selecting this particle. This is fulfilled for $\bar{B}^0 \to D^+\pi^-$.
Another requirement is that the calibration channel provides high statistics, which means that there has to be a high amount of events in the data set. This can also be easily understood since more events mean that there is more data to train on for the BDT's which in turn leads to a better selection of the D^+. The 2018 $\bar{B}^0 \to D^+\pi^-$ data set also fulfills this requirement as it is comprised of roughly 6,000,000 events..
Of course this high amount of statistics also provides a challenge for the fit to the full data set, since the fit becomes harder for a higher amount of events. High statistics are sensitive to smaller contributions from other decays that can no longer be neglected and therefore have to be added in the fit.
Another requirment for the calibration channel is that it shows a clean signature, i.e. a rather clean signal in the data. This is the case for $\bar{B}^0 \to D^+\pi^-$ as can be seen in Figure 10.

In principal the BDT's could also be trained on Monte Carlo simulation samples. However Monte Carlo simulations do not reproduce some of the variables of the decay correctly, such as PID variables, and these variables are used to train the BDT's. Additionally the BDt draws much of its discrimination power from correlations between these variables which are also not reproduced by the simulation. In order to use Monte Carlo for the training one would have to reweight these variables according to the shapes of the variables in actual LHCb data, but with the number of variables used in the BDTs being very large this would take considerably longer.

6.2 The 2018 $\bar{B}^0 \to D^+\pi^-$ LHCb data set

Figure 10 shows a first look at the full 2018 data set for the decay $\bar{B}^0 \to D^+\pi^-$ as it comes out of the *stripping* selection (see Section 4).
As can be seen very clearly a peak at $m \approx 5280$ MeV is already visible. This peak is the main signal peak of the decay $\bar{B}^0 \to D^+\pi^-$ and it is located at the $\bar{B}^0$ mass of $m_{\bar{B}^0} = 5279.58 \pm 0.17$ MeV. This also validates the claim in section 6.1 that this channel provides a clean signature and is therefore suited for the use as a calibration channel for D-from-B BDTs.

6.2.1 Understanding the mass spectrum

The first important step in the analysis of the mass spectrum in Figure 10 is to understand all the major components that make up the data. It is evident that there are other decays which factor into this spectrum, otherwise a clear signal peak at m_{B^0} with only combinatorial background data would be visible.

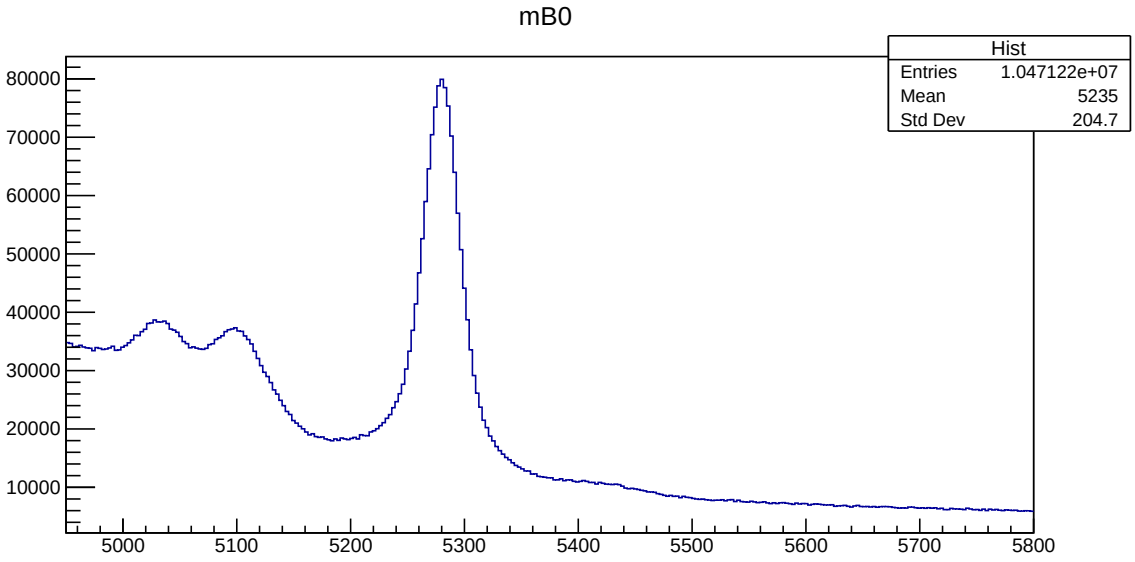


Figure 10: First look at the $\bar{B}^0 \to D^+\pi^-$ data

Understanding this background is of crucial importance for several reasons. First of all, knowing which other decays make up the data, allows for these decays to be included as seperate components in the fit. The inclusion of these components will provide for a more accurate fit that describes the data better. This also allows for systematic testing of different approaches to the training of the BDT which will be done in Section 7. For example it is not inherently clear which decays in the spectrum one should count as "Signal" and which as "Background" since the partially reconstructed decays, mentioned in the next section, for example also comes from a detached D^+ decay. These different ways of defining what the signal actually is lead to different BDT responses and performances. Secondly a correct understanding of the data is important for the correct determination of the signal yield, i.e. how many events belong to the actual $\bar{B}^0 \to D^+\pi^-$ decay. This will allow for a correct calculation of the BDT efficiencies since in order to properly determine them a large number of fits (≈ 1000) has to be performed. For such a large number of fits it is important that the fits are stable and robust.

In order to understand the spectrum a number of different tools could be used:

- **Full Monte Carlo Simulation**: Full MC simulation is of course perfect for understanding the different decays that might play a role in the spectrum, since it reproduces the decay with good accuracy. However this is not readily available for all decays
- **RapidSim**: This is a very fast simulation tool for heavy quark decays. However it does not reproduce every characteritic of the decay such as angular distributions and can therefore not always be used.
- **Kernal density PDFs**: This is a way to estimate probability distributions with gaussian kernel based on simulation data. The KDE PDFs are very easy to handle since all of

their parameters are given by simulation. However the simulation sometimes does not accurately reproduce parameters such as the mass position or the width of a certain decay.

- **Fitting the isolated component from the mass spectrum**: Using hard PID cuts one can positively select the desired decay from the mass spectrum in Figure 10 (see Section 6.3). These components can then be fitted and the results would be accurate since one would fit real data. However this is a multi-step procedure and first requires a fit to simulation data to find the general shape of the component. Additionally the PID cuts sometimes prove to be too hard or too loose resulting in mass spectrums that do not only consist of the desired decay.

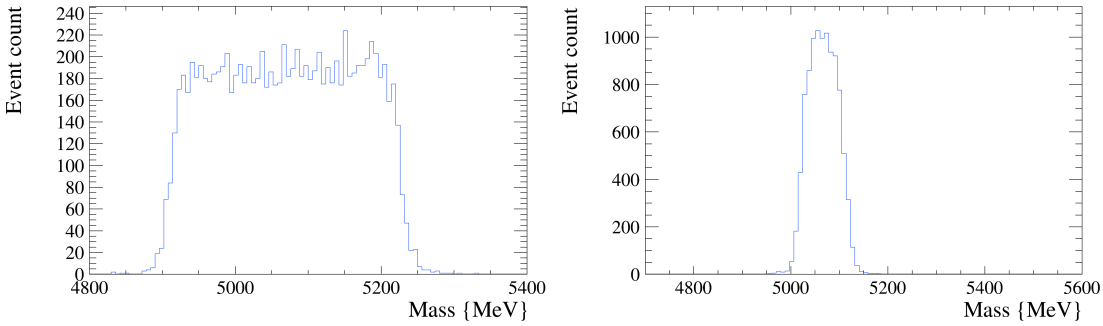
6.2.2 The different components in the data set

When looking at the data, the following decays have to be condidered:

Partially reconstructed decays

These are decays of the type $\bar{B}^0 \to D^+\pi^- X$, where X could possibly be a number of different particles but in particular a Photon, a Pion or two Pions in the mass range considered here. These decays are not rejected by the stripping lines and are therefore part of the spectrum in Figure 10.

They are located on the left end of the mass spectrum shown in Figure 10, which can be easisly understood. The stripping tries to reconstruct the $\bar{B}^0$ from the momentum and mass of its daughter particle, the Pion and the D^+, which in turn decays into two Kaons and a Pion. However if there are additional particles involved in the decay, that are not considered in the reconstruction, the mass and momentum of the daughter particles that are considered in the reconstruction is not enough to amount to the full mass of the $\bar{B}^0$, which is why these events are on the low end of the mass spectrum. However they still come from a true $\bar{B}^0$ decay.



(a) RapidSim Simulation of the decay $\bar{B}^0 \to (D^{*+} \to D^+\gamma)\pi^-$; Photon not reconstructed

(b) RapidSim Simulation of the decay $\bar{B}^0 \to (D^{*+} \to D^+\pi^0)\pi^-$; π^0 not reconstructed

Figure 11: Examples for partially reconstructed decays generated with RapidSim

Figure 11 shows two example of one of these decays namely the decay $\bar{B}^0 \to (D^{*+} \to D^+\gamma)\pi^-$, where the photon is not reconstructed and $\bar{B}^0 \to (D^{*+} \to D^+\pi^0)\pi^-$ where the π^0 is not reconstructed.
The data that can be seen in this Figure was produced with the Monte Carlo generator **RapidSim**, which is a very fast generator for simulated heavy-quark hadron decays [36].
However Figure 11 and especially Figure 11b illustrates the problem with the use of RapidSim. As will be discussed in Section 6.3 the decay $\bar{B}^0 \to (D^{*+} \to D^+\pi^0)\pi^-$ forms a characteristic double-peak structure which is not simulated by RapidSim but is vital to the mass spectrum. This is due to the fact that RapidSim does not take angular distributions into account and the structure is an effect of the spin structure of the D^{*+}. Due to conservation of angular momentum the missing π^0 can only go along or against the direction of the D^+ resulting in the Pion taking either very high or very low momentum from the D^+. This in turn results in the double horn structure. [37]

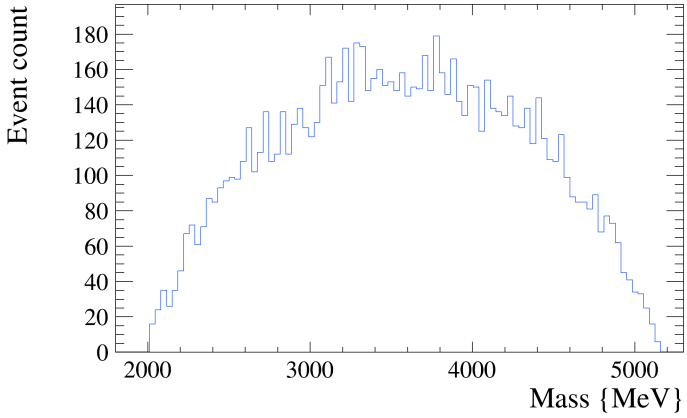


Figure 12: RapidSim Simulation of the decay $\bar{B}^0 \to D^+\pi^-\pi^0$

Another example, again simulated with RapidSim, is given in Figure 12. Shown here is a

simulation of the decay $\bar{B}^0 \to D^+\pi^-\pi^0$ and one can again see that the peak is located to the left of the $\bar{B}^0$-mass. In fact the shift between $m_{\bar{B}^0}$ and the right end of the peak in Figure 12 is roughly 130 MeV, which corresponds to the mass of the missing pion, that is not reconstructed in the decay ($m_\pi = 139.57018 \pm 0.00035$ MeV [38]). The peak does not start exactly a pion mass to the left of m_{B_0} because the pion was chosen to be smeared in RapidSim to account for the detector resolution.

For decays with two missing pions e.g $\bar{B}^0 \to D^+\pi^-\pi^0\pi^0$ one would see a very similar structure in the simulation, except for the fact that the peak would be shifted by two pion masses.

Misidentified decays

These are decays of a similar decay topology to $\bar{B}^0 \to D^+\pi^-$ that have been misidentified by the stripping due to the fact that there are no strong PID requirements in its selection as was detailed in Section 4. These misidentified decays mimic the topology of the $\bar{B}^0 \to D^+\pi^-$ decay, which means that they also feature a secondary b vertex made from an intermediate c hadron and a bachelor track and a tertiary c vertex made from three charged particles. Each of these tracks could be a proton, kaon, pion or muon which leads to a wide variety of possible decays.
The most prominent ones will be discussed in the following.

$\bar{B}^0 \to D^+ K^-$

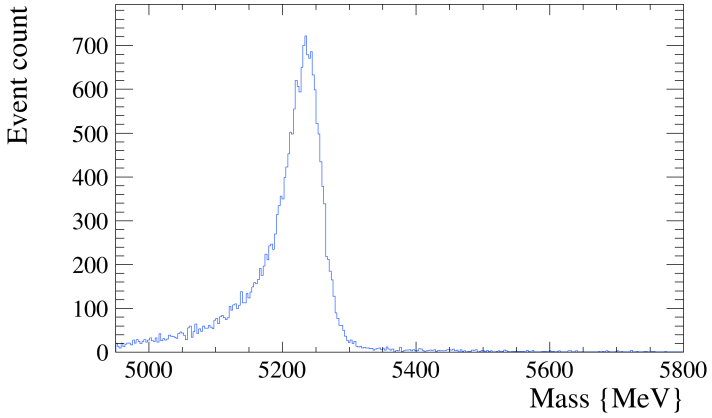


Figure 13: Full Monte Carlo Simulation of the decay $\bar{B}^0 \to D^+ K^-$

One of the misidentified decays that are prominent in the $\bar{B}^0 \to D^+ \pi^-$ data is the $\bar{B}^0 \to D^+ K^-$ decay, where a true K^- is given the pion mass. The mass peak is located to the left of the actual $\bar{B}^0 \to D^+ \pi^-$ peak and is responsible for the small rise in the spectrum at around 5200 MeV compared to the rest of the background. It also features a long tail towards lower masses that will have to be considered in the fit. The contribution of this decay to the mass spectrum is relatively small because this decay is Cabibbo suppressed.

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

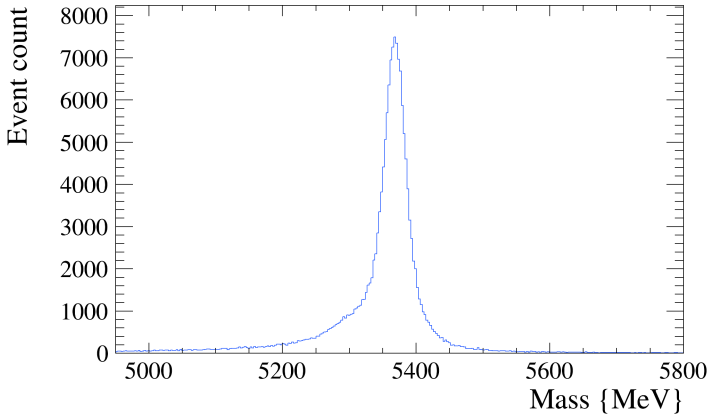


Figure 14: Full Monte Carlo Simulation of the decay $B_s^0 \to D_s^+ \pi^-$

Another one of these decays is $B_s^0 \to D_s^+ \pi^-$, where the D_s^+ is misidentified as an D^+. The $\bar{B}^0{}_s$ is actually heavier than the $\bar{B}^0$ ($m_{\bar{B}^0{}_s} = 5366.89 \pm 0.19$ MeV), which is why its peak

is located to the right of the $\bar{B}^0 \to D^+\pi^-$ peak. However it also features a tail towards lower masses, which means that some of the signal from this decay falls right under the signal peak of $\bar{B}^0 \to D^+\pi^-$ and therefore effects the signal yields.

$\Lambda_b^0 \to \Lambda_c^+\pi^-$

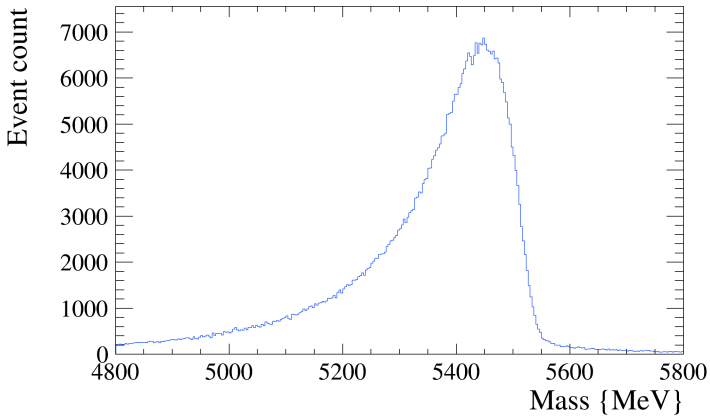


Figure 15: Full Monte Carlo Simulation of the decay $\Lambda_b^0 \to \Lambda_c^+\pi^-$

The last misidentified decay that should be mentioned here is the decay $\Lambda_b^0 \to \Lambda_c^+\pi^-$ in which the Λ_c^+ baryon is misidentified as a D^+. This is due to the fact that the two have similar masses ($m_{\Lambda_c^+} = 2286.46 \pm 0.14$ MeV, $m_{D^+} = 1869.65 \pm 0.05$ MeV) and the loose cuts in the stripping cannot differentiate between the two.
The peak of this decay is also visible in the data in Figure 10 as the small peak to the right of the signal peak at around 5400 MeV.

Combinatorial Background

The last component in the data that should be mentioned is the combinatorial background. This background is a typical by-product in any high energy particle experiment and consists of all the countless combinatorial combinations of other decay products that happen in the collision.

6.2.3 Which components should be included in the fit?

Of course the partially reconstructed and misidentified decays mentioned in section 6.6.2 are only a small fraction of all the possible decays that could be considered for the $\bar{B}^0 \to D^+\pi^-$ decay and for the fit to the data. Other examples for decays that were not mentioned are misidentified decays that occur at a lower rate such as $\Xi_b \to \Xi_c\pi$ or ones where only the tail of

the distribution feed into the signal, such as $\bar{B}^0 \to D^+ \mu \nu_\mu$. On top of that multiple misidentified particles could occur as well as a combination of misidentification and partial reconstruction. However the robustness of the fit has to be of concern as well. All of these components could be added to the final fit, but after a certain point the improvement of the fit would only be minimal and hardly worth the work.
This means the question is, which components are to be added for the fit and which componnets are unnecessary.

The strategy for finding the answer to this question was to simply add components until a satorfactory fit result was reached. First the components with visible effect in the data set in Figure 10 were added. These were as mentioned the three misidentified decays from section 6.2.2 and the double horn structure from the decay $\bar{B}^0 \to D^{*+} D^+ \pi^-$. The rest of the partially reconstructed decays were found to be sufficiently described by a single probability density funtion (PDF) and all further possible misidentified decays don't seem to play a major role in the data.
The strategy for adding the mentioned components to the final fit will be described in the following sections.

6.3 Finding and fitting the shapes of the various components

6.3.1 Strategy for finding the shapes

In this section the general strategy for fitting the components mentioned in section 6.2.3 will be discussed. These shapes are needed for inclusion in the final fit to the 2018 $\bar{B}^0 \to D^+ \pi^-$ data. The strategy consists of three steps:

Fitting simulated data of the decays
: First the simulated data of the decay in question will be fitted with appropriate probabilty density functions. The results from this fit will be taken as the basis for the next steps and certain parameters that dictate the shape of the distribution (the shape parameters) will be fixed or constrained to the values from the simulation.

Positively select the component from the full data set and fit with MC shapes
: In this step the values from the simulation fit will be taken to fit the actual component in the data. In order to do this, the individual component has to first be individually selected from the data using PID cuts. This is done to check the simulation results and to also get values for parameters that the simulation might not accurately represent, like the peak position or the peak width.

Include the component in the final fit
: After the previous values have been verified, the component can be added to the final fit.

6.3.2 The $\bar{B}^0 \to D^+K^-$ component

The first of these components that have to be analysed is the $\bar{B}^0 \to D^+K^-$ component.
The fits in the following sections are done using the generic b hadron decay fit software *beef* developed by Dr. Sebastian Neubert. [39].

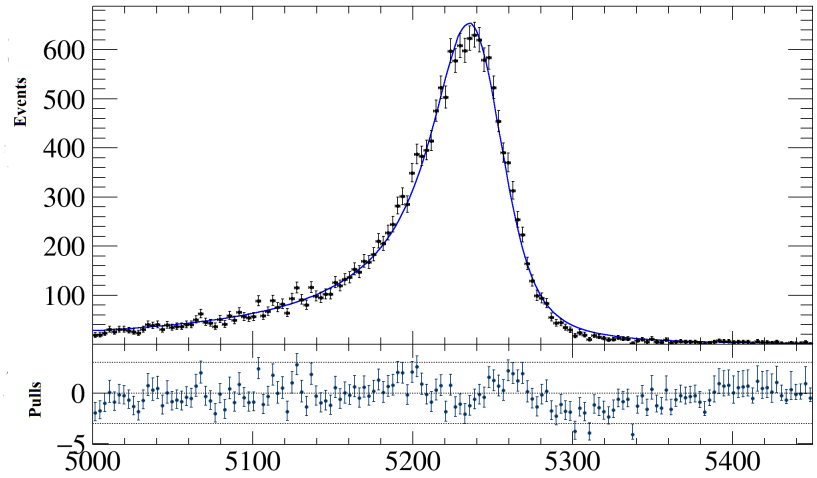


Figure 16: Fit to the MC data for $\bar{B}^0 \to D^+K^-$

Figure 16 shows the fit to the Monte Carlo simulation for this decay.
For this fit a convolution between a Johnson S_U distribution and a Gaussian was chosen:

$$PDF_{B2DK} = JohnsonS_U * Gaussian \tag{4}$$

where

$$JohnsonS_U = J(m, \mu, \sigma, \gamma, \delta) = \frac{\delta}{\sigma\sqrt{2\pi}\sqrt{z^2+1}} exp(-\frac{1}{2}(\gamma + \delta sinh^{-1} z)^2) \tag{5}$$

$$z = \frac{m-\mu}{\sigma} \tag{6}$$

and

$$Gaussian = G(m, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(\frac{(m-\mu)^2}{2\sigma^2}\right) \tag{7}$$

This Johnson S_U distribution is generally very well suited for describing distributions with long tails like in this case. In this distribution γ and δ are the parameters that account for the shape of the distribution, μ gives the position and σ the width. The convolution with the Gaussian distribution was added to account for the resolution of the detector.
As can be seen from the fit this funtion describes the data very well and the following values

for the shape parameters where taken to be fixed in the following steps:

Parameter	Value
γ	0.76432 ± 0.01482
δ	0.52776 ± 0.00790

Table 5: Shape parameters for the Johnson S_U distribution

The errors for these parameters are statistical errors found through the fit. It should be noted again that only these shape parameters but not the peak position and the peak width are taken from the MC simulation as the latter ones might differ from the actual data.

The next step is now to extract the $\bar{B}^0 \to D^+K-$ component from the actual $\bar{B}^0 \to D^+\pi-$ data. This is done using tight ProbNN cuts.

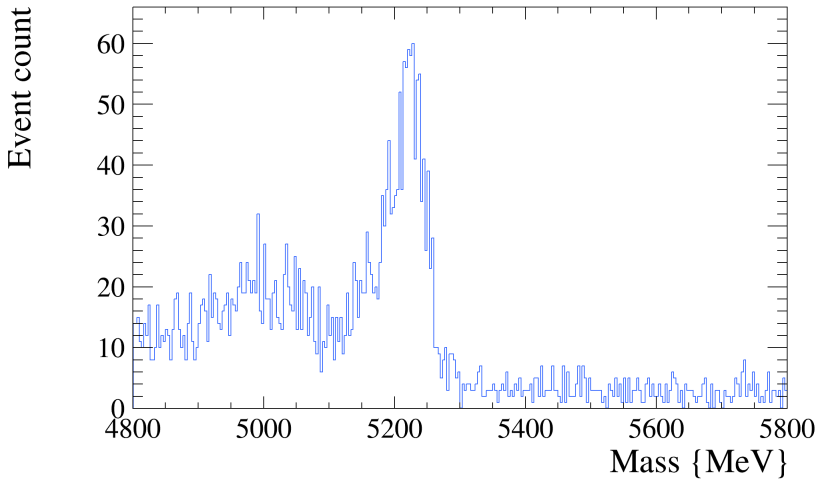


Figure 17: Positively selcted $\bar{B}^0 \to D^+K-$ component

The cuts that were used to produce this figure are

Cut	Purpose
$\lvert m_D - 1869MeV \rvert < 10$ MeV	Reduce non-D^+ background
pi_ProbNNk*(1 - pi_ProbNNpi) > 0.9	Select Kaon instead of Pion
D_K_ProbNNk > 0.4 D_pi1_ProbNNpi > 0.3	Select D^+

Table 6: PID cuts that were used to extract the $\bar{B}^0 \to D^+K^-$ component

Now the values from the MC simulation fit are used to fit the isolated data.

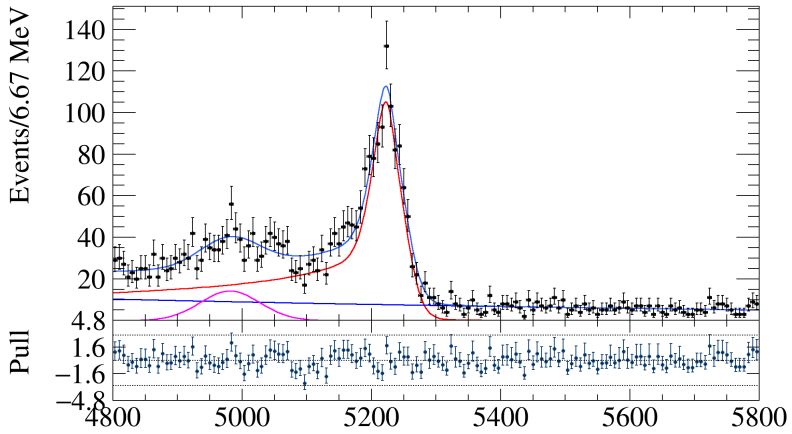
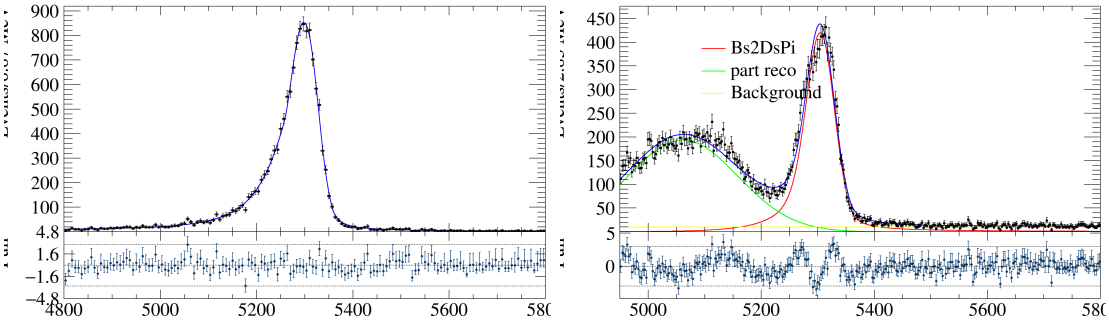


Figure 18: Fit to the isolated component for $\bar{B}^0 \to D^+K^-$

The $\bar{B}^0 \to D^+K^-$ peak is fitted in red in this plot. Additionally an exponential function was added to account for the combinatorial background and a simple Gaussian was chosen to fit the broad peak to the left of the $\bar{B}^0 \to D^+K-$ peak. This broad peak mainly consists of decays of the type $\bar{B}^0 \to D^+K^-\pi$ but is of no concern for the actual final $\bar{B}^0 \to D^+\pi^-$ fit since its contribution is too small and the structure is too broad.
As can be seen clearly from the fit the shape parameters found in the simulation describe the actual data very well and can therefore be used in the final fit.

6.3.3 The $B_s^0 \to D_s^+\pi^-$ component

The same strategy is now applied to the $B_s^0 \to D_s^+\pi^-$ component.



(a) Fit to the MC simulation

(b) Fit to the isolated component; taken from the data in Figure 10

Figure 19: The $B_s^0 \to D_s^+\pi^-$ component

For the fit to this component a sum of a Johnson S_U distribution and a Hypatia function was used.

$$PDF_{Bs2DsPi} = JohnsonS_U + Hypatia \tag{8}$$

This Hypatia function is a generalization of the standard Crystal Ball function for mass distributions with continuous per-event error distributions and was developed by D. Martinez Santos and F. Dupertuis of NIKHEF [40].
It was found that a sum of these two distributions as described above was able to best describe the data at hand.
The Hypatia distribution depends on 9 parameters

$$Hypatia = H(m, \mu, \sigma, \lambda, \zeta, \beta, a_1, a_2, n_1, n_2) \tag{9}$$

and Table 7 gives an owerview of their functions.

Parameter	Purpose
μ	peak position
σ	width of the distribution
λ ζ	Shape parameters
β	asymmetry parameter
a_1 a_2	transition points to the left/right tail of the distribution
n_1 n_2	power-law exponent for left/right tail

Table 7: The different parameters of the Hypatia distribution

The parameters that are of interest and that will be fixed to the values from the Monte Carlo simulation are $\lambda, \zeta, \beta, a_1, a_2, n_1$ and n_2 as well as γ and δ for the Johnson S_U distribution. The values from the simulation are

Parameter	Value
λ	-10
ζ	4.5246
β	0.01074 ± 0.00001
a_1	1.4918 ± 0.0007
a_2	2.3000 ± 0.0001
n_1	0.6858 ± 0.0008
n_2	1.0356 ± 0.0014
γ	1.8994 ± 0.0395
δ	1.3099 ± 0.0009

Table 8: Parameters taken from the MC simulation

It should be noted here that λ and ζ of the Hypatia distribution are very sensitive to small changes in the other parameters and generally behave very wild. Therefore it was chosen to already fix these two parameters in the MC fit and the two values that they were fixed to were found by attempting the fit in Figure 19a with different combinations of values for these parameters, out of which the best fit was chosen.

For the fit to the isolated component the following PID cuts were placed

Cut	Purpose
$\lvert M_{inv}(K^-\{\pi^+ \to K^+\}\pi^+) - 1968MeV\rvert < 10$ MeV	Reduce background and select D_s^+
$\lvert m_D - 1869MeV\rvert < 20$ MeV	Reduce background
pi_ProbNNpi*(1 - pi_ProbNNk) > 0.8	Select Pion
D_pi1_ProbNNk > 0.4 D_K_ProbNNk > 0.4	D_s^+ decays into two Kaons and one Pion

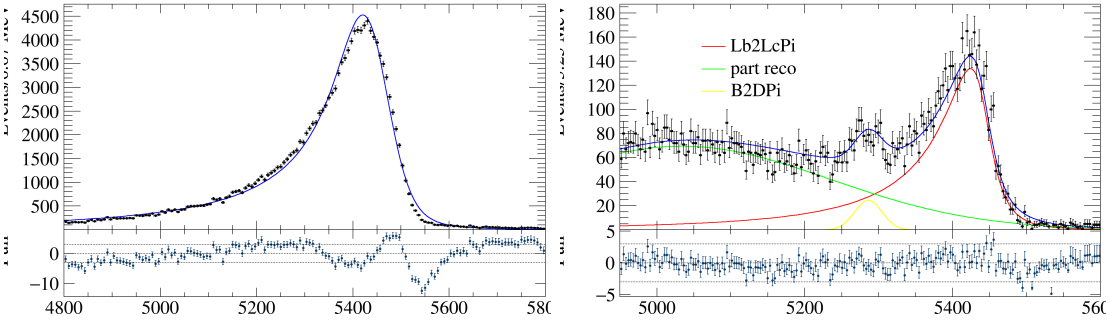
Table 9: PID cuts that were used to extract the $B_s^0 \to D_s^+\pi^-$ component

The cuts are similar to those used to extract the $\bar{B^0} \to D^+K^-$ component however the D_s^+ decays into two Kaons and one Pion. This is refelcted by the ProbNN cuts listed in Table 11 as well as the first cut listed here, which is an invariant mass cut on the decay products of the charm daughter particle, where instead of the pion mass hypothesis a kaon mass hypothesis was placed on one of the daughters.
The parameters taken from the simulation again describe the data well. The only point of concern might be the flanks of the main $B_s^0 \to D_s^+\pi^-$ peak, where the shape does not perfectly match the data, however the deviation is still within a 5 σ range and will not affect the final fit much since some of the parameters found here will not be completely fixed but rather constrained to the values found in the simulation (with a Gaussian constraint).
Additionally an exponential background and a Gaussian peak to the left of the main peak were

added again to effecticely describe partially reconstructed decays such as $B_s^0 \to D_s^+ \pi^- \gamma/\pi^0$.

6.3.4 The $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ component

Next this strategy is applied to the $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ component.



(a) Fit to the MC simulation

(b) Fit to the isolated component

Figure 20: The $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ component

As can be seen from the MC simulation data this decay again shows a similar shape to the $\bar{B}^0 \to D^+ \pi^-$ decay, which is why for the fit to this component a convolution of a Johnson S_U distribution and a Gaussian was used again.

$$PDF_{Lb2LcPi} = JohnsonS_U * Gaussian \tag{10}$$

The parameters that are taken from the simulation are

Parameter	Value
γ	1.0802 ± 0.0068
δ	0.6282 ± 0.0039

Table 10: Shape parameters taken from the simulation

It should be noted that compared to the simulation fits shown before the fit to the $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ simulation data does not describe the data as well. However this was the best fit that was found and the deviations from the simulation do not seem to affect the fit to the isolated component much, which is why this fit was chosen to be sufficient.

The PID cuts that were placed to get the data in Figure 20b are listed in Table 11.

Cut	Purpose
$\lvert M_{inv}(K^-\{\pi^+ \to p\}\pi^+) - 2286MeV\rvert < 30$ MeV	Reduce background and select Λ_c^+
$\lvert m_D - 1869MeV\rvert < 20$ MeV	Reduce background
pi_ProbNNpi*(1 - pi_ProbNNk) > 0.8	Select Pion
D_pi1_ProbNNp > 0.4 D_K_ProbNNk > 0.4	Λ_c^+ decays into a proton, a kaon and a pion

Table 11: PID cuts that were used to extract the $B_s^0 \to D_s^+\pi^-$ component

The cuts are again similar to the ones seen before, except that the Λ_c^+ decays into a proton, a kaon and a pion. This is refelcted by the ProbNN cuts listed in Table 11 as well as the first cut listed here, which is an invariant mass cut on the decay products of the charm daughter particle, where instead of the pion mass hypothesis a proton mass hypothesis was placed on one of the daughters.
The parameters taken from the simulation again describe the data very well and will therefore be added to the final fit.
What is interesting about the fit to this isolated component is that apart from the usual partially reconstructed decays to the left of the peak that are modelled with a Gaussian there is an additional peak at about 5280 MeV. This is a small residue of the $\bar{B}^0 \to D^+\pi^-$ peak that has not been completely rejected by the PID cuts.

6.3.5 Simulation data for the $\bar{B}^0 \to D^+\pi^-$ decay

The last simulation data that should be taken into account is the $\bar{B}^0 \to D^+\pi^-$ decay itself since an understanding of its shape is crucial to the fit of the full data set.

Figure 21 shows the fit to this simulation. For this fit a sum of a Hypatia distribution and a Gaussian was chosen

$$PDF_{B2DPi} = Hypatia + Gaussian \tag{11}$$

and the parameters that were taken from this fit are

Parameter	Value
λ	-1.8146 ± 0.0056
ζ	0
β	0
a_1, a_2	2.145 ± 0.030
n_1, n_2	2.4492 ± 0.0013

Table 12: Parameters taken from the MC simulation

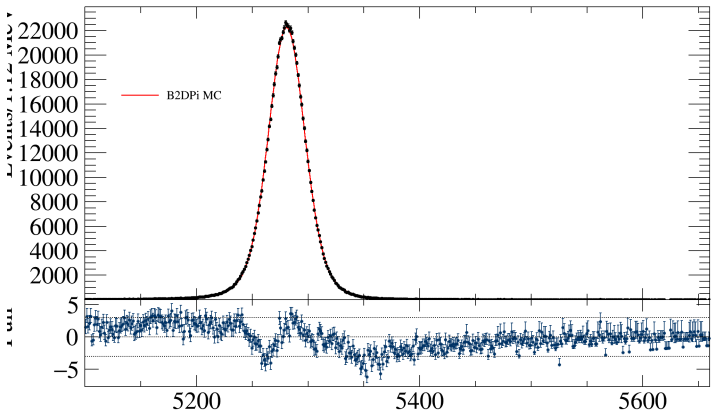


Figure 21: Fit to Monte Carlo Simulation of the decay $\bar{B}^0 \to D^+\pi^-$

What is different from this fit to the use of the Hypatia funtion for $\bar{B}^0{}_s \to D_s^+\pi^-$ is that the distribution is completely symmetrical for $\bar{B}^0 \to D^+\pi^-$. This is reflected by $\beta = 0$, as well as $a_1 = a_2$ and $n_1 = n_2$. The value for ζ was again fixed to a suitable value.

6.3.6 "Clean Signal" fit to extract shapes for partially reconstructed decays

The last shapes that have to be determined before the fit to the full $\bar{B}^0 \to D^+\pi^-$ data set can be performed are the shapes for the partially reconstructed background on the left end of the mass spectrum and the double horn structure of the decay $\bar{B}^0 \to (D^*+ \to D^+\pi^0)\pi^-$.
However for these two shapes there was no simulation data available and thus the shapes had to be determined from another source. In order to do this a "clean signal" was extracted from the data using the following PID cuts

Cut	Purpose
$\lvert m_D - 1869MeV\rvert < 10$ MeV	Reduce background
pi_ProbNNpi*(1 - pi_ProbNNk) > 0.9	Select Pion
D_K_ProbNNk > 0.4 D_pi1_ProbNNpi > 0.3	Select D^+

Table 13: PID cuts used to extract a clean $\bar{B}^0 \to D^+\pi^-$ signal

These cuts only select the decay $\bar{B}^0 \to D^+\pi^-$ but do not reject all the partially reconstructed decays including the Horn structure.

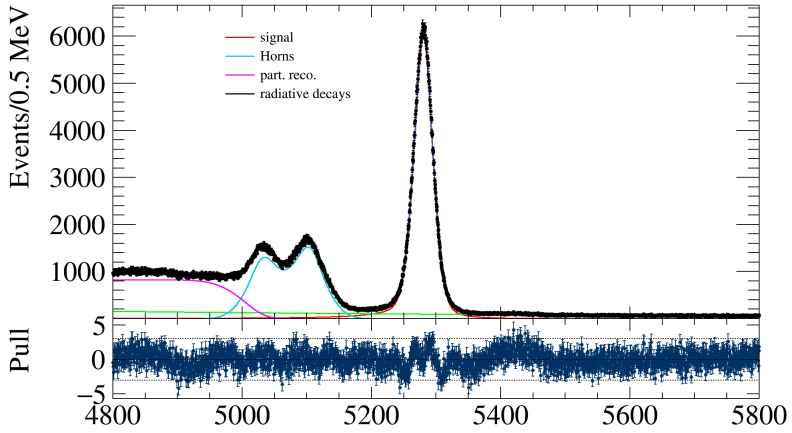


Figure 22: Fit to the clean $\bar{B}^0 \to D^+\pi^-$ signal

For the partially reconstructed decays one single PDF, called the **tilted Subbotin distribution**, was used. This distribution is a new custom build PDF [41] that is related to the generalized normal Subottin distribution [42] and was found to be very powerful in describing the plateau-like shape of the partially reconstructed decays. It has one additional parameter compared to the normal Subbotin distribution, the slope s that gives it its name, and depends on 4 parameters in total:

Parameter	Purpose
μ	Position
σ	Width
β	Shape parameter
s	Slope of the plateu

Table 14: Parameters of the Tilted Subbotin

It has the shape

$$TiltedSubbotin(m) \propto exp\left(\frac{s*m}{\sigma} - \left(|\frac{m-\mu}{\sigma}|\right)^{\beta}\right) \tag{12}$$

For the double horn structure of the $\bar{B}^0 \to (D^*+ \to D^+\pi^0)\pi^-$ a PDF, specifically developed for this structure was used. This PDF is referred to as $ROOHornsDini$ and describes the Horn structure as a Gaussian from which a parabola-shaped part has been cut out. [43]. It features 7 parameters as listed in Table 15.

Parameter	Purpose
a,b	kinematic endpoints of the decay
csi	relative height of the two peaks
$shift$	uniform shift of the whole shape across B mass
$sigma$	width of the peaks
$ratio\_sigma$	relative width of the two peaks
$fraction\_sigma$	fractional amount of the main gaussian in resolution double-gaussian

Table 15: Parameters of the RooHornsdini distribution

The values for these parameters that were found and then fixed for the final fit are

Parameter	Value
csi	0.7999 ± 0.0001
$shift$	18.508 ± 0.007
$ratio\_sigma$	2.413 ± 0.001
$fraction\_sigma$	0.1214 ± 0.0003
β	7.3748 ± 0.0015
s	0

Table 16: Parameters taken from the Clean Signal fit

Another advantage of this clean signal fit is that the shape parameters of the $\bar{B}^0 \to D^+\pi^-$ simulation fit can be tested. As can be seen from Figure 22 they generally describe the data well and just feature small deviations on the flanks of the distribution. However this is very likely an effect of the PID cuts that were applied in order to get this clean signal data and indicates that there are still other components in the region of the signal peak that have not been rejected by the cuts.

6.4 Fit to the full 2018 $\bar{B}^0 \to D^+\pi^-$ data set

Now that all the necessary componnets have been fitted and their shapes verified the final $\bar{B}^0 \to D^+\pi^-$ fit can be performed. For this all of the previously determined shapes are included.

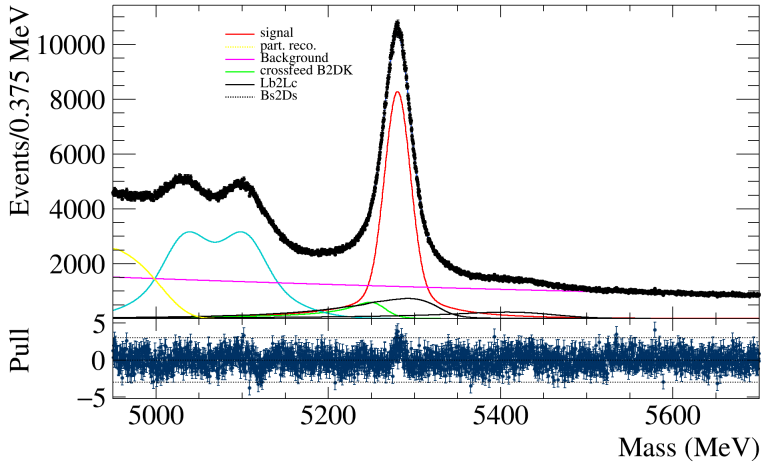


Figure 23: Fit to the full $\bar{B}^0 \to D^+\pi^-$ data (Scenario B)

As can be seen from Figure 23 the final result describes the data very well. For the signal yield

$$N_{Sig} = 954,520 \pm 6580 \tag{13}$$

was found.

6.4.1 Using the full $\bar{B}^0 \to D^+\pi^-$ fit for systematic studies of training the D^+ BDT

The exact understanding of the $\bar{B}^0 \to D^+\pi^-$ mass spectrum acquired in this section allows for different scenarios for the training of the D^+ BDT to now be tested. In particular it can be tested which components of the mass spectrum should be classified as signal and which as background to test which of these classifications gives the best BDT response. In total three scenarios for this classification were tested which means that this fit was performed three times. The classification of the components for the three scenarios is the following:

Scenario A

- **Signal**: $\bar{B}^0 \to D^+\pi^-$
- **Background**: combinatorial background
- **Rest**: $\bar{B}^0 \to D^+K^-$, partially reconstructed decays, $B_s^0 \to D_s^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\pi^-$

Scenario B

- **Signal**: $\bar{B}^0 \to D^+\pi^-$
- **Background**: $B_s^0 \to D_s^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\pi^-$, combinatorial background
- **Rest**: $\bar{B}^0 \to D^+K^-$, partially reconstructed decays

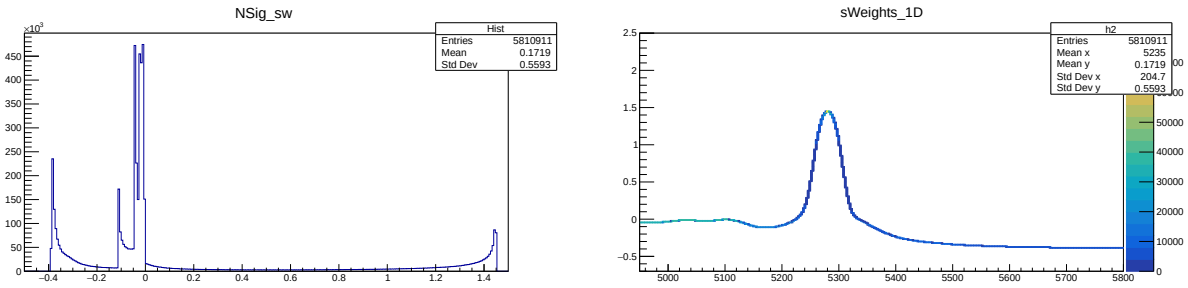
Scenario C

- **Signal**: $\bar{B}^0 \to D^+\pi^-$, , partially reconstructed decays, $\bar{B}^0 \to D^+K^-$
- **Background**: $B_s^0 \to D_s^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\pi^-$, combinatorial background

The three fits all are very similar to the one in Figure 23, for which scenario B was chosen, and are therefore not explicitely shown here.
The BDTs in the following sections were always trained against the category called background.

7 Training the D-from-B BDTs

The training of the BDTs follows the strategy described in section 5.1.To label the training data with the desired classification categories, signal and background, $_s$Weights are used [44]. These $_s$Weights can be used to project out different components of the data set, which is shown in Figure 24b and they are extracted from the $\bar{B}^0 \to D^+\pi^-$ fits in the previous sections. They are extracted for each of the three scenarios mentioned before and Figure 24 shows the $_s$Weights for scenario B as an example.



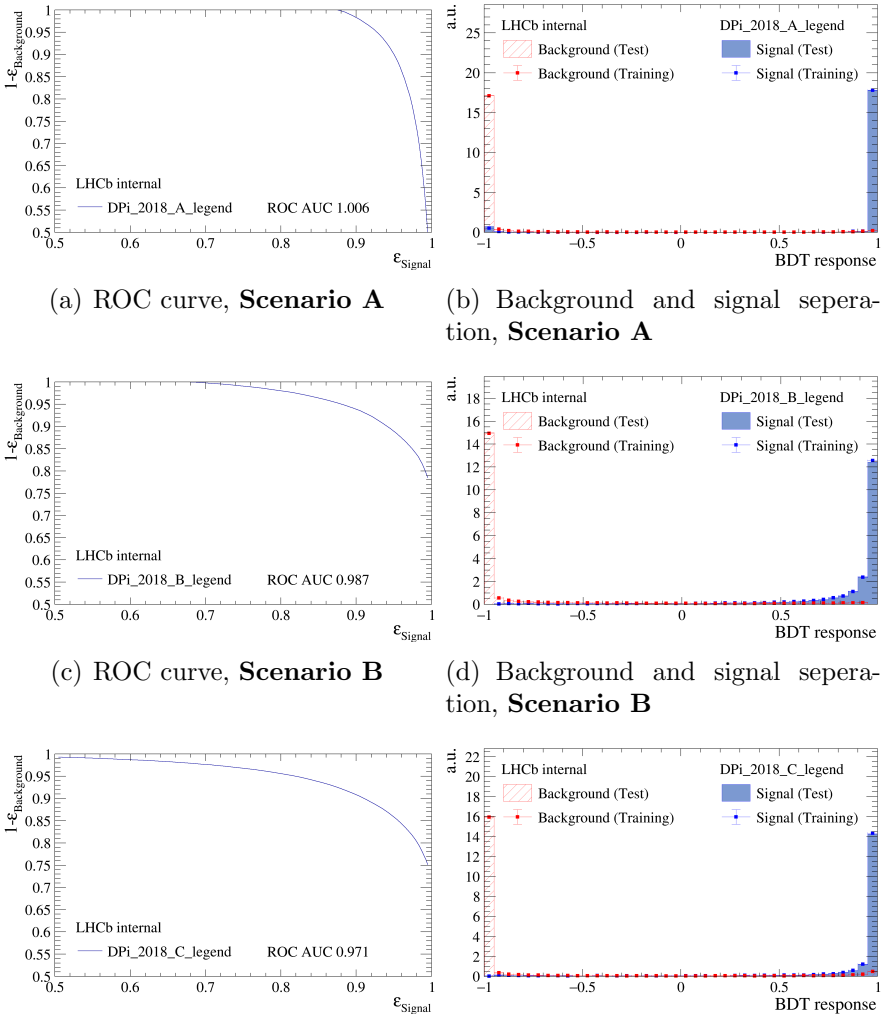
(a) Extracted sWeights from the B2DPi fit

(b) B2DPi data plotted against the sWeights

Figure 24: sWeights from the $\bar{B}^0 \to D^+\pi^-$ fit (Scenario B)

Figure 24b, the reweighted data, when compared to the $\bar{B}^0 \to D^+\pi^-$ data set, shows that there are positive weights in the region of the signal, negative weights in the regions where there are large amounts of combinatorial background and weights close to 0 for the regions to the left of the signal peak, where there are a lot of partially reconstructed decays..

For the BDT training input a total of 69 variables was used including PID variables, kinematic variables (p and p_T) as well as topolocigal variables [2]. However since these BDTs are supposed to select the D^+ particle, only variables for this particle and no variables from the beauty mother were chosen as input. This allows the BDT to be used on many different decays involving a D^+.
The BDT training response for the three categories from section 6.4 is shown in Figure 25:



(a) ROC curve, **Scenario A**

(b) Background and signal seperation, **Scenario A**

(c) ROC curve, **Scenario B**

(d) Background and signal seperation, **Scenario B**

(e) ROC curve, **Scenario C**

(f) Background and signal seperation, **Scenario C**

Figure 25: BDT responses for all scenarios

Figures 24 a,c and e show the ROC curves of the three different BDTs.They illustrate the relation between signal efficiency ϵ_{Signal} and background rejection $1 - \epsilon_{Background}$. For example one would reject 95% of the background by retaining roughly 85% of the signal for the BDT which belongs to scenario B (Figure 25c). As can be seen from these plots scenarios A produces the best ROC curve, meaning that this BDT rejects the most background while still getting a large signal efficiency. However this is not surprising since the BDT in scenario A only trained against the combinatorial background, while the other two BDTs also trained against misidentified decays. These are harder to seperate from the main signal since they feature a very similar topology which explains the slightly worse ROC curves.
Figures 24 b,d and f indicate that the BDTs were not overtrained since the testing and training samples lie right on top of each other for all scenarios.

8 Results for the $\bar{B}^0 \to D^+ D_s^-$ decay

The BDT can now be applied to the $\bar{B}^0 \to D^+ D_s^-$ data.
However additionally to the D^+ a BDT for the $D_s^- \to K^- \pi^- \pi^+$ decay is also needed in order to properly select the signal for this decay. This means that two BDTs will be applied to the data.
The development and training of the D_s BDT is not part of this thesis and the one used in the following sections has been developed by Dr. Marian Stahl of the LHCb group at the University of Heidelberg. It uses the channel $\bar{B}_s^0 \to D_s^- \pi^-$ as calibration channel and has been developed in much the same way as presented here for the D^+ BDT. [2]

8.1 The 2018 $\bar{B}^0 \to D^+ D_s^-$ data set

Before actually applying the BDTs to the $\bar{B}^0 \to D^+ D_s^-$ data a first look at the data is presented in this section. This decay has very similar structure to the $\bar{B}^0 \to D^+ \pi^-$ decay illustrated in Figure 8, except this time the $\bar{B}^0$ decays into two D mesons. The D^+ then again decays into a negative Kaon and two positive Pions, while the D_s^- decays into a negative Kaon, a positive Kaon and a negative Pion.

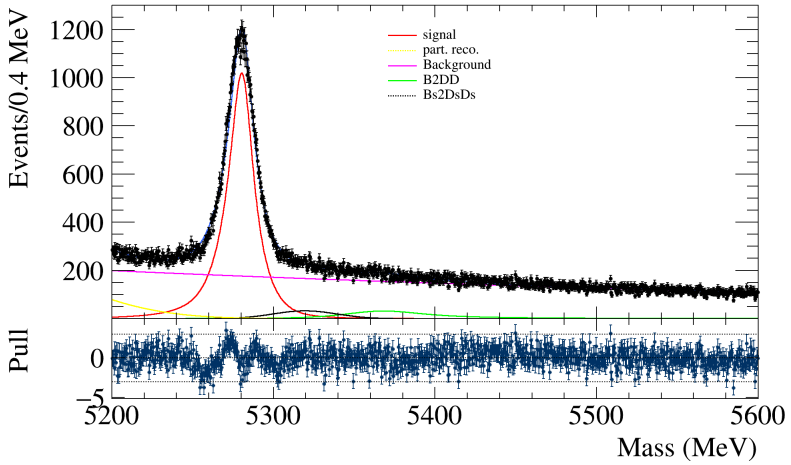
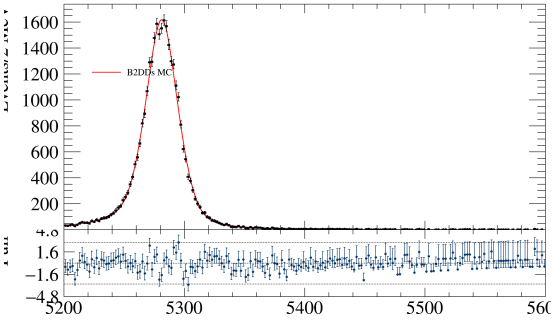


Figure 26: Fit to the full 2018 $\bar{B}^0 \to D^+ D_s^-$ data

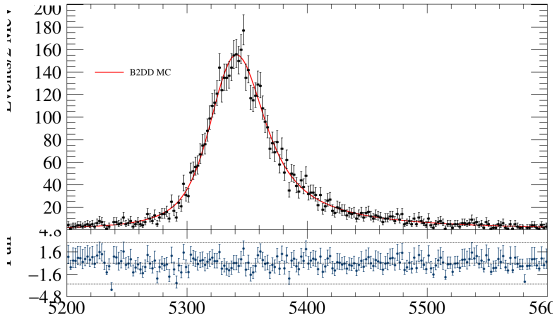
Figure 26 shows a fit to this data set. For this fit the same strategy as in the fit to the $\bar{B}^0 \to D^+\pi^-$ data in section 6 was used, i.e. the shapes for the different components used in the fit were taken from simulation and tested on the isolated component, which was positively selected from the data using PID cuts. The following components were taken into account for the fit and they describe the data sufficiently well:

- **Signal**: $\bar{B}^0 \to D^+ D_s^-$
- **Misidentified decays**: $\bar{B}^0 \to D^+ D^-$ & $B_s^0 \to D_s^+ D_s^-$
- **Partially reconstructed decays**: one PDF sufficient to describe all
- **Exponential combinatorial background**

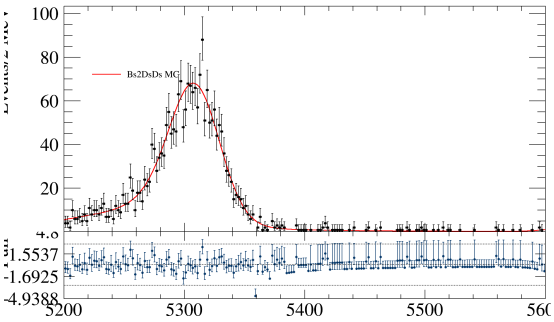
Figure 27 shows the MC simulation that was used to find the shapes of the Misidentified decays and the signal:



(a) MC fit for $\bar{B^0} \to D^+ D_s^-$



(b) MC fit for $\bar{B^0} \to D^+ D^-$



(c) MC fit for $B_s^0 \to D_s^+ D_S^-$

Figure 27: Simulation data used to find the shapes for the fit in Figure 26

The PDFs that were used here are:

$$PDF_{B2DDs} = JohnsonS_U + Gaussian \tag{14}$$

$$PDF_{B2DD} = JohnsonS_U * Gaussian \tag{15}$$

$$PDF_{Bs2DsDs} = JohnsonS_U * Gaussian \tag{16}$$

These PDFs were again chosen by trying out different shapes and settling on the PDF that best describes the data. It is notable here that the PDF used for the $\bar{B^0} \to D^+ D_s^-$ simulation is the sum of a Gaussian and the JohnsonS_U distribution as opposed to the convolution of the two for the other decays. However when looking at the shapes in the simulation it already becomes clear the $\bar{B^0} \to D^+ D_s^-$ decay shows a different shape compared to the other two decays; it is much more narrow and also more symmetrical. This is why a sum of the two distributions was chosen in this case.

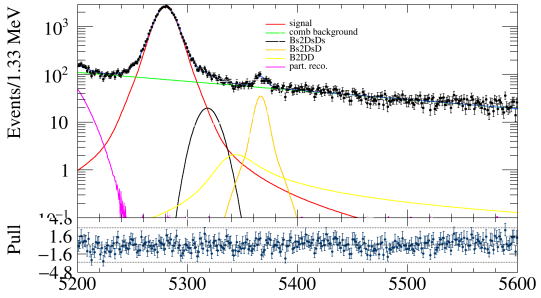
For the fit in Figure 26 a signal yield of

$$N_{Signal} = 63608 \pm 1057 \tag{17}$$

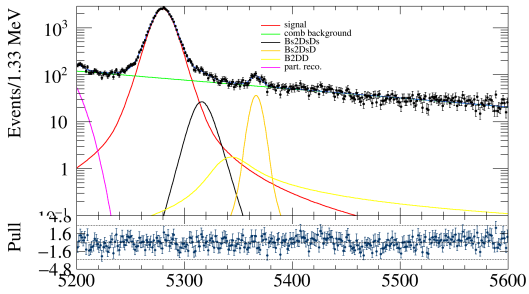
This number will be used as a reference point for the signal yields that will be found when using the PID and BDT cuts in the next sections.

8.2 Comparison of the three BDT scenarios

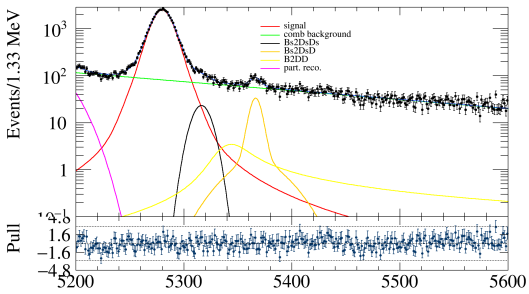
As detailed in section 6.4 three different D^+ BDTs were trained on the $\bar{B}^0 \to D^+\pi^-$ data corresponding to the three different ways of categorizing the components in the mass spectrum. In order to test which of these BDTs gives the best results they will be applied to the $\bar{B}^0 \to D^+D_s^-$ data seperately along with the D_s^- BDT. The decision which BDT gives the best results will then be based on the signal yields as well as the background rejection of the different BDTs.



(a) Fit to $\bar{B}^0 \to D^+D_s^-$ with BDT cuts, Scenario A



(b) Fit to $\bar{B}^0 \to D^+D_s^-$ with BDT cuts, Scenario B



(c) Fit to $\bar{B}^0 \to D^+D_s^-$ with BDT cuts, Scenario C

Figure 28: Comparison between the different BDT scenarios

The fits in Figure 28 feature a logarithmic y scale to better see the small background components.
The yields for the signal and the different components are listed in table 17.
When looking at these numbers it becomes clear that Scenario B seems to be the best version of the D^+ BDT. Not only does it produce the highest amount of signal yield, it also rejects the highest amount of combinatorial background. On top of that it also rejects the highest amount of misidentified background.

Component	Scenario A	Scenario B	Scenario C
$\bar{B^0} \to D^+ D_s^-$	42874 ± 207	43714 ± 199	42762 ± 216
$\bar{B^0} \to D^+ D^-$	124 ± 12	106 ± 17	204 ± 91
$B_s^0 \to D_s^+ D_s^-$	420 ± 46	336 ± 23	342 ± 49
$B_s^0 \to D_s^+ D^-$	350 ± 33	332 ± 29	359 ± 38
part. reco. decays	359 ± 55	283 ± 51	310 ± 43
comb. background	16961 ± 157	15697 ± 307	16421 ± 149

Table 17: PID cuts used for Figure 30

It is notable that scenario B proved to give the best results, despite scenario A featuring the best ROC curve as seen in section 7. This is explained by the fact that scenario B also trained against misidentified decays ($B_s^0 \to D_s^+ \pi^-$, $\Lambda_b^0 \to \Lambda_c^+ \pi^-$) and is therefore able to reject these decays as well.

The BDT cuts that were used to produce the plots in Figure 28 were:

$$D\_BDT > -0.6 \quad \& \quad D_s\_BDT > -0.6 \tag{18}$$

Additionally a cut on the D_s mass was placed. These BDT cuts are very loose in order to properly compare the background rejection of the three BDT scenarios.

For the following section scenario B will be used, when comparing the BDT to PID cuts.

8.3 Comparison of BDT and PID cuts

The D^+ (Scenario B) and D_s^- BDTs will now be applied to the data in Figure 26 with tighter cuts and the resulting mass spectrum will be fitted to find a value for the signal yield. For the fits in Figure 29 and 30 additional mass cuts on the D^+ and D_s^- mass were placed to further reduce the background. These were:

$$|m_{D_s^-} - 1968MeV| < 30MeV \quad \& \quad |m_{D^+} - 1869MeV| < 30MeV \tag{19}$$

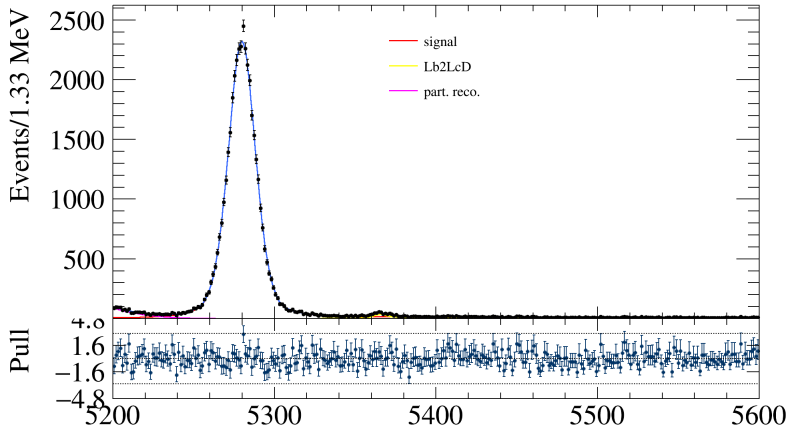


Figure 29: Fit to the data, after applying BDT cuts

Figure 29 shows the results of this fit. In this fit an additional component for $\bar{B}^0 \to D^+D^-$ was fitted. The signal yield that was found is

$$N_{Signal,BDT_{1D}} = 42688 \pm 223 \tag{20}$$

The cuts that were used here were

$$D\_BDT > -0.1 \quad \& \quad D_s\_BDT > -0.1 \tag{21}$$

and they were chosen blindly as a working point with the only requirement being that the total number of events is nearly equal to that in Figure 30 (cf. Table 19). It could therefore be argued that the performance of the BDT cuts can still be optimized for better results.
Additionally to the BDT cuts a cut on the D^+ and D_s^- masses was applied to reduce the background contributions of decays without these particles.

Furthermore a fit was produced, where only PID cuts were used to produce a clean signal.

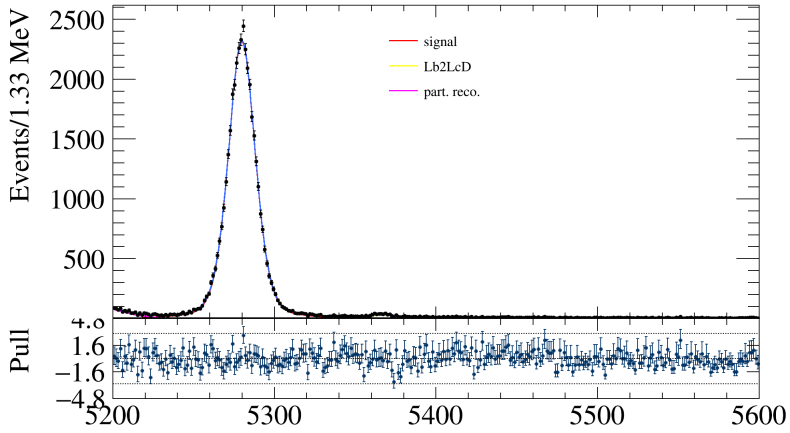


Figure 30: Fit to the data, after applying PID cuts

The cuts that were used to produce this plot were

Cut
$Ds\_Kp\_ProbNNk > 0.3$
$Ds\_Km\_ProbNNk > 0.3$
$Ds\_pi\_ProbNNpi > 0.3$
$D\_K\_ProbNNk > 0.2$
$Ds\_pi1\_ProbNNpi > 0.2$
$D\_pi\_ProbNNpi > 0.2$

Table 18: PID cuts used for Figure 30

and the first three cuts select the $D_s^- \to K^+K^-\pi^-$ decay whereas the last three cuts select the $D_s^- \to K^-\pi^+\pi^+$ decay.
In the plots 29 and 30 the cuts were generally chosen so that both plots would consist of roughly the same amount of events. This allows for a first comparison between BDT and PID selection:

Selection	N_{Total}	N_{Signal}	Signal percentage
PID cuts	42905	39181 ± 208	(91.32 ± 0.48) %
BDT cuts	43596	42688 ± 223	(97.92 ± 0.51) %

Table 19: Total number of events and signal yield for Figures 29 and 30

Table 19 illustrates the better performance of the BDT cuts compared to the PID cuts. With the BDT cuts a significantly higher signal yield could be achieved in fact an improvement of 6 % could be reached. However the PID cuts were still able to produce a very clean signal in Figure 29. It is also notable that the the signal yield for the BDT cuts is only smaller by approxamety 1000 events compared to the plot in Figure 28b (Table 17, equation 17) . This

could be explained when looking at the ROC curve for Scenario B in Figure 25c where the curve is very steep on the right end of the diagram, meaning that a relatively large change in background rejection does not effect the signal efficiency considerably. However it should also be noted that the BDT cuts away a large part of the signal, since the fit to the full data set achieved a signal yield of approximately 63000. These two points would have to be investigated in the future.
Additionally the BDT cuts proved to be easier to handle with only two variables (D_BDT and Ds_BDT) to consider compared to the 6 PID variables that had to be considered to achieve the clean signal seen in Figure 28 (compare Table 17).

9 Conclusion and Outlook

This thesis presented the development of a $D^+ \to K^-\pi^+\pi^+$ BDT as part of the D-from-B BDT package, based on the calibration channel $\bar{B}^0 \to D^+\pi^-$. The most important and time consuming step here was the exact understanding of the mass spectrum of this calibration channel to produce stable and robust fits and to ensure a good understanding of the selection.

This thesis proved to be a valueable insight into the use of D^+ BDTs and their development. For example due to the work presented here, it become apparent that not all partially reconstructed contributions to the $\bar{B}^0 \to D^+\pi^-$ mass spectrum would have to be included in the fit for the BDT training. This was used for the D-from-B BDTs that were included in the inclusive double-D stripping line for the 2018 LHCb restripping, which were based on the work presented in this thesis, where a smaller mass window of only 5150 MeV - 5800 MeV was used for the $\bar{B}^0 \to D^+\pi^-$ fit. This tighter mass window resulted in an easier and more robust fit. Additionally this thesis provided an exact understanding of the misidentified components in the $\bar{B}^0 \to D^+\pi^-$ mass spectrum, which were included as background for the BDT training.
As was shown, the BDTs presented here produced very good results and it was found that scenario B, where $B_s^0 \to D_s^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and the combinatorial background were categorised as background and only the $\bar{B}^0 \to D^+\pi^-$ component was considered as signal for the training, produced the best results. This BDT was then compared to a standard selection using PID cuts and was found to be more efficient.
Additionally the $\bar{B}^0 \to D^+D_s^-$ channel was established as a test channel for the D^+ and D_s^- BDTs. With this channel the performance of the new stripping lines can be studied and efficiency correction can be calibrated in the future. Moreover efficiency correction procedures for PID variables and D-from-B BDTs can be compared here. Should the BDTs eventually be used in the upgrade trigger for the LHCb experiment the $\bar{B}^0 \to D^+D_s^-$ channel would also become important in order to develop real-time calibration for these BDTs.

References

[1] Marian Stahl, Alessio Piucci, Florian Reiss, Nicola Skidmore, Sebastian Neubert; Identification of non-prompt charm hadrons; https://cds.cern.ch/record/2647804

[2] Marian Stahl; First observation of the decay $\Lambda_b^0 \to \Lambda_c^+ \bar{D}^{*0} K^-$ in preparation of a pentaquark search in the $\Lambda_c^+ \bar{D}^{*0}$ system at the LHCb experiment: https://www.physi.uni-heidelberg.de//Publications/Lb2LcDzbKm_v1.1.pdf

[3] Conor Fitzpatrick: A 30 MHz software trigger for the LHCb Upgrade, https://indico.cern.ch/event/786084/contributions/3315132/attachments/1800811/2937205/ConorFitzpatrick_ACAT19_v0.pdf

[4] LHCb collaboration: Letter of intent for the LHCb upgrade; https://cds.cern.ch/record/1333091?ln=de

[5] R. Aaij, S. Benson, M. De Cian, D. Dziurda, C. Fitzpatrick, V. V. Gligorov, E. Govorkova1, O. Lupton, R. Matev, S. Neubert, A. Pearce, H. Schreiner, S. Stahl, M. Vesterinen: A comprehensive real-time analysis model at the LHCb experiment, https://twiki.cern.ch/twiki/pub/LHCb/TurboSPPaper/turbosp-paper.v1.5.pdf

[6] : M. Thomson: Modern particle physics. New York: Cambridge University Press, 2013

[7] I. J. R. Aitchison and A. J. G. Hey: Gauge theories in particle physics: A practical introduction. Vol. 2: Non-Abelian gauge theories: QCD and the electroweak theory. Bristol, UK: CRC Press, 2012.

[8] LHCb collaboration: Observation of $J/\Psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\Psi K^- p$ decay; Phys. Rev. Lett. 115, 072001 (2015)

[9] T. D. Lee and C. N. Yang: Question of Parity Conservation in Weak Interactions; Phys. Rev. 104, 254 (1956)

[10] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson: Experimental Test of Parity Conservation in Beta Decay; Phys. Rev. 105, 1413 (1957)

[11] M. Thomson: Modern particle physics. New York: Cambridge University Press, 2013

[12] LHCb collaboration: Road map for selected key measurements from LHCb; https://cds.cern.ch/record/1224241/?ln=sv

[13] I. Polyakov: Heavy flavor production in the forward acceptance at the LHC; Jul 2015

[14] R. Lindner, "LHCb layout 2. LHCb schema 2." LHCb Collection., Feb 2008.

[15] P. R. Barbosa-Marinho et al., LHCb VELO (VErtex LOcator): Technical Design Report. Technical Design Report LHCb, Geneva: CERN, 2001.

[16] P. Collins, "Overview diagram showing spacing of modules along Z, and positions open and closed." LHCb Collection.

[17] S. Amato et al., LHCb magnet: Technical Design Report. Technical Design Report LHCb, Geneva: CERN, 2000.

[18] P. R. Barbosa-Marinho et al., LHCb inner tracker: Technical Design Report. Technical Design Report LHCb, Geneva: CERN, 2002.

[19] P. R. Barbosa-Marinho et al., LHCb outer tracker: Technical Design Report. Technical Design Report LHCb, Geneva: CERN, 2001.

[20] S. Amato et al., LHCb RICH: Technical Design Report. Technical Design Report LHCb, Geneva: CERN, 2000

[21] P. A. Cherenkov, "Visible radiation produced by electrons moving in a medium with velocities exceeding that of light," Phys. Rev., vol. 52, pp. 378–379, Aug 1937.

[22] C. Jones, ANN: PID Retuning for Reco14 Data, https://indico.cern.ch/event/226062/contributions/475644/attachments/371741/517276/ANNPIDRetuning-Reco14-06052013.pdf

[23] S. Amato et al., LHCb calorimeters: Technical Design Report. Technical Design Report LHCb, Geneva: CERN, 2000.

[24] P. R. Barbosa-Marinho et al., LHCb muon system: Technical Design Report. Technical Design Report LHCb, Geneva: CERN, 2001.

[25] LHCb collaboration: Trigger schemes https://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html

[26] R. Antunes-Nobrega et al., LHCb trigger system: Technical Design Report. Technical Design Report LHCb, Geneva: CERN, 2003.

[27] LHCb collaboration; Performance of the LHCb trigger and full real-time reconstruction in Run 2 https://arxiv.org/abs/1812.10790

[28] Jordy Butters, Branching fraction measurement of $B^0 \to D_s^+\pi^-$; https://indico.cern.ch/event/734620/contributions/3145068/attachments/1728443/2792639/Presentation_B2OC.pdf

[29] StrippingB02DPiD2HHHBeauty2CharmLine. http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/stripping/config/stripping21/bhadroncompleteevent/strippingb02dpid2hhhbeauty2charmline.html

[30] E. Bos et al., The Trajectory Model for Track Fitting and Alignment, Tech. Rep. CERN-LHCb-2007-008, CERN, Geneva, Mar, 2007.

[31] LoKi::KalmanFilter in LHCb Software Documentation. http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/latest_doxygen/d3/dcd/namespace_lo_ki_1_1_kalman_filter.html#details.

[32] P. Seyfert, The search for $\tau \to \mu\mu\mu$ at LHCb, PhD thesis, Physikalisches Institut, Heidelberg, Jan, 2015, CERN-THESIS-2015-021.

[33] J. H. Friedman, Greedy function approximation: A gradient boosting machine, Ann. Stat. 29 (2001), no. 5 1189.

[34] A. Hoecker et al., TMVA: Toolkit for Multivariate Data Analysis, PoS ACAT (2007) 040, arXiv:physics/0703039, latest version under http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf

[35] RapidSim: https://github.com/gcowan/RapidSim

[36] Measurement of CP observables in $B^{\pm} \to D^{*}K^{\pm}$ and $B^{\pm} \to D^{*}\pi^{\pm}$ decays; PHYS. LETT. B777 (2018) 16; ARXIV:1708.06370

[37] Particle listings π - PDG; http://pdg.lbl.gov/2014/listings/rpp2014-list-pi-plus-minus.pdf

[38] beef, S. Neubert: A generic B-fitter based on RooFit; https://gitlab.cern.ch/sneubert/beef

[39] D. Martinez Santos, F. Dupertuis; Mass distributions marginalized over per-event errors: https://arxiv.org/pdf/1312.5000.pdf

[40] Marian Stahl: RooTiltedSubbotin (code for implementation in beef); https://gitlab.cern.ch/sneubert/beef/blob/master/src/RooTiltedSubbotin.cc

[41] Tibor K. Pogany, Saralees Nadarajah; On the Characteristic Function of the Generalized Normal Distribution; http://www.maths.manchester.ac.uk/ goran/research-reports/psrr18-2009.pdf

[42] Donal Hill, Paolo Gandini; PDFs for partially reconstructed $B \to DX$ decays: RooHORNSdini and RooHILLdini; https://indico.cern.ch/event/283973/contributions/1633715/attachments/523911/722643/CTS_meeting_5_12_13.pdf#search=roohornsdini

[43] M. Pivk and F. R. Le Diberder, "SPlot: A Statistical tool to unfold data distributions," Nucl. Instrum. Meth., vol. A555, pp. 356–369, 2005.